# MAST90132: LIE ALGEBRAS

DAVID RIDOUT

ABSTRACT. The theory of Lie algebras is fundamental to the study of groups of continuous symmetries acting on vector spaces, with applications to diverse areas including geometry, number theory and the theory of differential equations. Moreover, since classical systems have conserved quantities derived from continuous symmetries, by Noether's theorem, and quantum mechanical systems are described by Hilbert spaces acted on by continuous symmetries, Lie algebras and their representations are also fundamental to modern mathematical physics. This subject develops the basic theory in a way that is (hopefully) accessible to both pure mathematics and mathematical physics students, with an emphasis on examples. The main theorems are: the classification of complex semisimple Lie algebras, in terms of Cartan matrices and Dynkin diagrams, and the classification of finite-dimensional representations of these algebras, in terms of highest weight theory.

TO THE READER. In these notes, there are many many exercises. Some are almost trivial, whilst some require serious thought. In both scenarios, they are intended to deepen your understanding of the material being introduced. They are also intended to be solved as we go, **not** at the end of semester or when they form part of a due assignment. Indeed, the material that follows will often use the exercise's result. In this spirit, every exercise is supposed to be solved using methods introduced previously in these notes or in the undergraduate curriculum. In particular, you are not supposed to use results that you have encountered in other masters-level subjects, for example Differential Topology and Geometry, or in private study. But if in doubt, simply ask!

*Date*: Last modified: November 3, 2020.

CONTENTS

## 1. Introduction

Lie theory is named after Sophus Lie (pronounced "**lee**"), a Norwegian mathematician who introduced and studied the continuous symmetries of differential equations, much as Galois introduced and studied the discrete symmetries of algebraic (polynomial) equations. Nowadays, Lie theory has expanded into a huge collection of work that aims, primarily, to model the notion of continuous symmetry in all its many incarnations and applications.

One simple class of examples consists of the rotationally invariant systems in $\mathbb{R}^3$. The symmetry of these systems is continuous because rotations in $\mathbb{R}^3$ may be parametrised by three angles with continuous ranges (two angles to choose the axis of rotation in spherical polar coordinates and then one more to specify how much to rotate). In elementary physics, this symmetry is responsible for the conservation of angular momentum. Consequently, angular momentum is modelled by a *Lie group*. For angular momentum in $\mathbb{R}^3$, this group is called $\mathsf{SO}(3)$ and we shall meet it properly soon.

Physicists come across Lie groups and their cousins, *Lie algebras*, all the time. Much of the time, one can only hope to analyse (solve) a physical system with the aid of guiding symmetry principles. Thus, Lie theory (or finite/discrete group theory) is often lurking in the shadows of fundamental physics. An excellent example of this is the awarding of the 1969 Nobel prize in Physics to Murray Gell-Mann who noticed that many of the properties of the plethora of fundamental particles being created in accelerators could be organised in terms of the representation theory of another Lie group called $\mathsf{SU}(3)$. This breakthrough led to quarks, gluons and quantum chromodynamics, a significant chunk of the standard model of fundamental particle physics. We shall meet this Lie group soon as well.

Lie theory is brought to the fore in modern physics research, especially quantum physics, because symmetry is often the only reliable tool one has to uncover order in this rather unintuitive world. The standard model, gauge theories, supersymmetry, integrability, even string theory: all rely heavily on various flavours of Lie theory and its generalisations. In fact, so too do atomic and nuclear physics, optics and even computational and spectroscopic chemistry!

But Lie theory is **mathematics**. In particular, you won't need a background in physics in order to appreciate the subject matter, although we will make plenty of remarks to relate the material to mathematical physics and will discuss a couple of applications to physics in detail. Lie theory combines practically all of the main areas of pure mathematics including algebra, analysis, geometry and topology. In many cases, the study of Lie algebras and groups has also driven (and continues to drive) research in these areas. In algebra, Lie-theoretic generalisations such as quantum groups, Kac-Moody algebras and vertex algebras continue to be strongly represented in international research. Harmonic analysis is essentially Fourier theory on (locally compact) Lie groups and ergodic theory, which combines probability, thermodynamics and chaotic dynamical systems, is most naturally formulated on compact Lie groups. The geometry of Lie groups plays a starring role in the study of many partial differential equations and their topological features represents one of the primary motivations for developing algebraic topology, homological algebra and category theory. No doubt you will see many more striking examples of the central role played by Lie theory as you learn more and more mathematics.

In this subject, a first glimpse into the beautiful world of Lie theory, we will mostly concentrate on the theory of complex semisimple Lie algebras and their representations, with some discussion of the corresponding Lie groups. This means, unfortunately, that we will not be seriously looking at the geometry and topology of the latter, nor can we delve into the important question of how to do calculus and analysis on them. The reason is the prerequisites: Lie algebras really only require a solid grounding in linear algebra and an appreciation of basic abstract algebra (which we can learn as we go). Lie groups, on the other hand, require more facility with geometry and topology than our undergraduate curriculum currently offers. That said, they are excellent examples for illustrating many aspects of modern mathematics so you will surely meet them again in other subjects.

We would hope that all pure mathematicians, and all the good applied ones (which includes all mathematical physicists), would agree that Lie theory is an essential part of both classical and modern mathematics. We trust that you will enjoy this first encounter with this beautiful theory and expect that it will leave you wanting more, more, more!

## Resources

As you'd expect for such a beautiful, important and well-studied field of mathematics, there are many good sources to learn about semisimple Lie algebras and their friends. Here are a few.

[1] *K Erdmann and M Wildon: Introduction to Lie algebras.*
This is a math textbook aimed at 3rd years at Imperial College London. It's a very good introduction, though necessarily brief in many ways. The authors also seem to have a bit of a thing for non-zero characteristics (imho).

[2] *J Fuchs and C Schweigert: Symmetries, Lie algebras and representations.*
Also introductory, but aimed more at graduate mathematical physicists who need a solid grounding in Lie theory. I like this book a lot, perhaps because it's written by two leaders in my field. While it eschews the theorem-proof style favoured elsewhere, it still maintains rigour when not inconvenient.

[3] *P Woit: Quantum theory, groups and representations: an introduction.*
A very recent textbook that aims to delve deeply into the group-theoretic formulation of quantum mechanics. Perhaps not suitable for mathematicians, but excellent bedtime reading for math physicists. Also available for free here:

`http://www.math.columbia.edu/~woit/`

[4] *J Stillwell: Naïve Lie theory.*
A beautiful book that deals with matrix Lie groups in a very concrete fashion. It lacks motivation and generality, but is well worth reading nonetheless.

[5] *B Hall: Lie groups, Lie algebras and representations: an elementary introduction.*
A fine text indeed. Occasionally unconventional but very thorough in its approach. Make sure you look at the second edition as it's quite different to the first.

[6] *W Fulton and J Harris: Representation theory.*
This is a math textbook with an extremely good (and well-deserved) reputation. It uses the theorem-proof style, but is not ashamed to ramble at times and do lots and lots of examples (often the same one more than once). That said, it requires quite a bit of mathematical maturity and there are some quite deep geometric asides throughout that one needs to recognise (and perhaps avoid).

[7] *C Curtis and I Reiner: Representation theory of finite groups and associative algebras.*
I absolutely love this old classic. Whilst it aims at people taking group representation theory (Lie algebras don't even get a mention), it's still the best reference I know for an enormous amount of general material about representations and modules.

[8] *R Carter: Lie algebras of finite and affine type.*
Another excellent math textbook. Detailed, but quite tough (and necessarily so). This would be one of the best references I know.

[9] *J-P Serre: Complex semisimple Lie algebras.*
Beautifully written and concise, this is not really an introduction. However, it has good descriptions of several important topics that are generally omitted in other sources. Very good for filling in certain gaps.

[10] *J Humphreys: Introduction to Lie algebras and representation theory.*
This is probably the most comprehensive text on finite-dimensional Lie algebras. Can be terse in places, but pretty much everything is there and the order has been carefully cultivated to make the progression as linear as possible. Notation gets a bit pedantic at times. Highly recommended, but hard work.

[11] *J Humphreys: Reflection groups and Coxeter groups.*
Another carefully cultivated text, this time reflecting (haha) on the detailed theory of a class of groups that include the Weyl groups that arise in semisimple Lie theory. Excellent, but challenging.

[12] *V Varadarajan: Lie groups, Lie algebras and their representations.*
Another comprehensive text, but even harder work than Humphreys. Good for a deeper understanding of the core material but best left to a second course.

[13] *R Moody and A Pianzola: Lie algebras with triangular decompositions.*
Sometimes generality can bring clarity. This awesome text deals with a very general class of Lie algebras, but I often find their presentation clearer than the more specialised texts. Hard work though.

[14] *N Bourbaki: Lie groups and Lie algebras.*
Speaking of hard work, this is the number one uncompromising approach to Lie theory in all its generality. Three books in total, translated from the french originals, much of this was written by the masters of the field. I regard it as a backup for when no other source has what I need! Not for the faint-hearted.

## 2. Lie groups and algebras

In this section, we shall introduce the notion of a Lie group, somewhat heuristically, and illustrate it with plenty of natural examples. We study how any given Lie group gives rise to a linearisation called a Lie algebra and use this to motivate the axioms of the latter. We also explore relationships between Lie algebras, formalised in terms of homomorphisms, and discuss the correspondence between Lie algebras over the real and complex number fields.

### 2.1. **Lie groups**

Recall that a group is just a set equipped with two operations: multiplication and inversion. The multiplication is associative, $(ab)c = a(bc)$, and inversion requires an identity $\mathbb{1}$, satisfying $a\mathbb{1} = \mathbb{1}a = a$, so that $aa^{-1} = a^{-1}a = \mathbb{1}$. Of course, every element is invertible.

A *Lie group* is a group that also carries the structure of a smooth manifold so that its operations are smooth. Rather than fuss over what this means (you should do this in a differential topology/geometry course), we shall be vague and just note that it means that it may be continuously (indeed smoothly) parametrised so that one can do calculus on it. Our aim here is not to develop the theory of Lie groups rigorously; as we said before, this is not really possible with our prerequisites. Instead, we aim to use Lie groups to motivate the introduction of Lie algebras. However, we shall also take the opportunity to quickly discuss some of the things that make Lie groups a bit more challenging (but fun).

Here are some first (very simple) examples:

**Example 1.**

- **(a)** The group $\mathsf{U}(1)$ of complex numbers of modulus 1 is a Lie group (with complex number multiplication and inversion as operations).
- **(b)** The real line $\mathbb{R}$ is a Lie group (with addition and negation as operations).
- **(c)** The integers $\mathbb{Z}$ form a group (actually a subgroup of $\mathbb{R}$), but a discrete one, so we shall not regard it as a Lie group. ▲

As for vector spaces, and so manifolds, we need to specify the "ground field". We will primarily deal with real and complex Lie groups, meaning that the ground field is $\mathbb{R}$ and $\mathbb{C}$, respectively. This is partly because these are the most common fields used in applications, but also partly because Lie theory over other fields, *eg*. finite fields, can have quite a different flavour.

Note that even though $\mathsf{U}(1)$ is defined using complex numbers, it is a real Lie group because it is parametrised by a single real parameter (the angle). Indeed, it is diffeomorphic (*ie*. isomorphic as smooth manifolds) to the circle $S^1$ which is a one-dimensional real manifold. A complex Lie group must be even-dimensional (as a real manifold).

The standard examples of Lie groups are *matrix groups*:

**Example 2.**

- **(a)** The group of invertible $n \times n$ matrices forms a Lie group, denoted by $\mathsf{GL}(n)$, called the *general linear* Lie group. When necessary, we emphasise the ground field as follows: $\mathsf{GL}(n;\mathbb{R})$, $\mathsf{GL}(n;\mathbb{C})$, *etc.* These groups have dimension $n^2$ (over their ground field).
- **(b)** The subgroup of $\mathsf{GL}(n)$ whose elements have determinant 1 forms a Lie group $\mathsf{SL}(n)$ called the *special linear* Lie group. Its dimension is $n^2 - 1$.
- **(c)** The upper-triangular $n \times n$ matrices, *ie.* those with entries satisfying $A_{ij} = 0$ for all $j < i$, form a Lie group $\mathsf{T}(n)$ provided that we insist that the diagonal entries $A_{ii}$ are non-zero. Its dimension is $\frac{1}{2}n(n+1)$. If we insist that the diagonal entries are actually all 1, then we get instead a different Lie group $\mathsf{T}'(n)$ of dimension $\frac{1}{2}n(n-1)$.
- **(d)** The orthogonal $n \times n$ matrices, *ie.* the real matrices satisfying $A^{\mathsf{T}} = A^{-1}$, form a real Lie group $\mathsf{O}(n)$ called the *orthogonal* Lie group. The subgroup $\mathsf{SO}(n)$ of matrices in $\mathsf{O}(n)$ of determinant 1 is likewise a Lie group called the *special orthogonal* Lie group. The dimension of both $\mathsf{O}(n)$ and $\mathsf{SO}(n)$ is $\frac{1}{2}n(n-1)$.
- **(e)** The unitary $n \times n$ matrices, *ie.* the complex matrices satisfying $A^{\dagger} = A^{-1}$, form a **real** Lie group $\mathsf{U}(n)$ called the *unitary* Lie group. The subgroup $\mathsf{SU}(n)$ of matrices in $\mathsf{U}(n)$ of determinant 1 is likewise a Lie group called the *special unitary* Lie group. The dimension of $\mathsf{U}(n)$ is $n^2$ and that of $\mathsf{SU}(n)$ is $n^2 - 1$.
- **(f)** Recall that orthogonal matrices $A \in \mathsf{O}(n)$ are characterised by the fact that they preserve the standard inner product on $\mathbb{R}^n$: $\langle Av, Aw \rangle = \langle v, w \rangle$. If we relax the positive-definiteness requirement, *ie.* we equip $\mathbb{R}^n$ with a non-degenerate symmetric bilinear form, then it is customary to denote $\mathbb{R}^n$ by $\mathbb{R}^{p,q}$, where $p$ and $q$ are the number of positive and negative eigenvalues of the invertible symmetric matrix $B$ representing the bilinear form ($\langle v, w \rangle = v^{\mathsf{T}} B w$). The matrices preserving such a bilinear form also form a Lie group denoted by $\mathsf{O}(p,q)$. We have $p + q = n$, $\mathsf{O}(p,q) \simeq \mathsf{O}(q,p)$ and $\mathsf{O}(n,0) = \mathsf{O}(n)$. These orthogonal groups have special counterparts as well, obtained by restricting to the matrices of determinant 1 and denoted by $\mathsf{SO}(p,q)$.
- **(g)** If we instead relax the inner product on $\mathbb{R}^n$ to a non-degenerate skew-symmetric bilinear form, *ie.* the representing matrix $J$ is invertible and antisymmetric, then the matrices preserving this form yield a Lie group $\mathsf{Sp}(n)$ called the *symplectic* Lie group. Since $J$ is antisymmetric and invertible, its eigenvalues are pure imaginary and pair up as distinct complex conjugates. It follows that $n$ must be even; we accordingly often denote the symplectic Lie group by $\mathsf{Sp}(2n)$. The dimension of $\mathsf{Sp}(2n)$ is $n(2n+1)$. We do not need special counterparts because it may be shown that matrices in $\mathsf{Sp}(2n)$ already have determinant 1. ▲

We remark that for $\mathsf{O}(p,q)$, $\mathsf{SO}(p,q)$ and $\mathsf{Sp}(2n)$, the matrices $B$ and $J$ may be chosen, without loss of generality, to have the following block-diagonal forms:

$$B = \left(\begin{array}{c|c} \mathbb{1}_p & 0_{pq} \\ \hline 0_{qp} & -\mathbb{1}_q \end{array}\right), \quad J = \left(\begin{array}{c|c} 0_{nn} & \mathbb{1}_n \\ \hline -\mathbb{1}_n & 0_{nn} \end{array}\right). \tag{2.1}$$

Here, $\mathbb{1}_k$ denotes the $k \times k$ identity matrix (and $0_{k\ell}$ is the $k \times \ell$ zero matrix).

Of course, there are variants of these Lie groups that are defined in terms of linear transformations rather than matrices. For example, $\mathsf{GL}(V)$ denotes the Lie group of invertible linear transformations from the vector space $V$ to itself. This generalisation is particularly important when $V$ is infinite-dimensional. However, we shall mostly restrict ourselves to finite dimensions in what follows, in which case there is no real difference between $\mathsf{GL}(V)$ and the matrix group $\mathsf{GL}(\dim V)$.

We remark that the Lorentz group of special relativity is basically $\mathsf{O}(1,3)$ (or $\mathsf{SO}(1,3)$ or even something called $\mathsf{SO}^+(1,3)$, depending on the author). In *Minkowski spacetime* $\mathbb{R}^{1,3}$, the bilinear form (the *metric*) is represented by the matrix $B = \operatorname{diag}\{1,-1,-1,-1\}$.

It is rather nice that matrices provide such a rich set of examples of Lie groups. However, there are many other important examples that are not so easily described. In particular, the orthogonal groups $\mathsf{O}(n)$ and $\mathsf{SO}(n)$ have "double covers", called the *pin* and *spin* groups $\mathsf{Pin}(n)$ and $\mathsf{Spin}(n)$, that are important in quantum physics. They may be constructed using matrices, but are far more naturally constructed using Clifford algebras. Moreover, there are finite-dimensional Lie groups that cannot be constructed from matrices at all. A notable example is the "universal cover" of $\mathsf{SL}(2;\mathbb{R})$ that arises in physics in connection with general relativity and/or string theory on 3-dimensional anti-de Sitter space.

**Exercise 1.** Explain why the Lie groups of Example 2 should have the dimensions indicated. [Note that this is not asking you to learn manifold theory, define the dimension of a manifold and then rigorously prove the dimension statements above. It is rather asking you to come up with plausible reasons for the dimensions being as stated.] ▼

Some Lie groups can be (usefully) understood using algebraic or geometric tools. For example, $\mathsf{O}(2)$ is the set of $2 \times 2$ real matrices that preserve lengths, *ie.* rotations and reflections in $\mathbb{R}^2$. Since reflections have determinant $-1$, the rotations alone correspond to $\mathsf{SO}(2)$. But, rotations in $\mathbb{R}^2$ are just specified by an angle so $\mathsf{SO}(2)$ should be diffeomorphic to the circle $S^1$. In particular, we have a (rather obvious) isomorphism $\mathsf{SO}(2) \xrightarrow{\simeq} \mathsf{U}(1)$ of Lie groups given by

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \longmapsto \mathsf{e}^{\mathsf{i}\theta}. \tag{2.2}$$

This means that the map and its inverse are smooth group homomorphisms (*ie.* they preserve the group operations).

Similarly, $\mathsf{SO}(3)$ describes rotations in $\mathbb{R}^3$. However, it is not diffeomorphic to the unit sphere $S^2$ because its dimension is 3. It is actually diffeomorphic to the real projective space $\mathbb{R}P^3$.

**Example 3.** Consider $\mathsf{SU}(2)$, the group of unitary $2 \times 2$ complex matrices of determinant 1. Parametrising such a matrix as $A = \left(\begin{smallmatrix} z & w \\ w' & z' \end{smallmatrix}\right)$, with $z, w, w', z' \in \mathbb{C}$, the determinant condition means that $A^{-1} = \left(\begin{smallmatrix} z' & -w \\ -w' & z \end{smallmatrix}\right)$. Unitarity then gives $z' = z^*$ and $w' = -w^*$, so

$$\mathsf{SU}(2) = \left\{ \begin{pmatrix} z & w \\ -w^* & z^* \end{pmatrix} : z, w \in \mathbb{C} \text{ and } |z|^2 + |w|^2 = 1 \right\} \tag{2.3}$$
$$= \left\{ \begin{pmatrix} a + b\mathrm{i} & c + d\mathrm{i} \\ -c + d\mathrm{i} & a - b\mathrm{i} \end{pmatrix} : a, b, c, d \in \mathbb{R} \text{ and } a^2 + b^2 + c^2 + d^2 = 1 \right\}.$$

It now follows easily that $\mathsf{SU}(2)$ is diffeomorphic to the 3-sphere $S^3$. ▲

**Exercise 2.** Show that $\mathsf{SL}(2; \mathbb{R})$ can be similarly realised as the 3-sphere in $\mathbb{R}^{2,2}$, *ie.* find a parametrisation in terms of four real numbers satisfying $a^2 + b^2 - c^2 - d^2 = 1$. ▼

Unfortunately, most other Lie groups cannot be identified with nice familiar manifolds. Nevertheless, the description of their topology is a classical (and very beautiful) subject in its own right. In this vein, we shall conclude by briefly discussing some of these topological features, specifically connectedness and compactness.

Recall that a space, *eg.* a manifold, is said to be *connected* if it cannot be written as the union of two closed subspaces. Because we are only considering matrix Lie groups, they may be embedded into a suitably high-dimensional $\mathbb{R}^n$ (or $\mathbb{C}^n$). We may therefore regard being closed as meaning that the space contains its limit points with respect to the usual metric on $\mathbb{R}^n$ (or $\mathbb{C}^n$). Similarly, a subspace of $\mathbb{R}^n$ ($\mathbb{C}^n$) is said to be *compact* if it is closed and bounded, where we may take bounded to mean that there is an absolute upper- and lower-bound on the possible entries of the matrix.

**Example 4.** It should be clear from their realisations as spheres that $\mathsf{U}(1)$ and $\mathsf{SU}(2)$ are compact and connected. ▲

**Example 5.** It is easy to see that $\mathsf{O}(n)$ is not connected: $\det\colon \mathsf{O}(n) \to \{\pm 1\}$ is a polynomial in the entries of the orthogonal matrix, hence continuous, so it must actually be constant on each connected component. $\mathsf{O}(n)$ therefore has at least two connected components. It is not quite so easy to see that $\mathsf{SO}(n)$ is connected (but it is). ▲

**Example 6.** The function $A \mapsto A^\dagger A$ is polynomial in the (real and imaginary parts of the) entries of $A$, hence continuous. Because the inverse image under a continuous function of a closed set is closed, we see that $\mathsf{U}(n)$ is the inverse image of the closed set consisting of the identity matrix $\mathbb{1}$, hence it is closed. $\mathsf{U}(n)$ is moreover bounded because

$$A^\dagger A = \mathbb{1} \quad \Rightarrow \quad \sum_{j=1}^{n} (A^\dagger)_{ij} A_{ji} = 1 \quad \Rightarrow \quad \sum_{j=1}^{n} \left|A_{ij}\right|^2 = 1, \quad \text{for all } i, \\ \Rightarrow \quad |A_{ij}| \leqslant 1, \quad \text{for all } i \text{ and } j. \tag{2.4}$$

It follows that $\mathsf{U}(n)$ is compact. ▲

**Exercise 3.** Show that $\mathsf{SL}(2;\mathbb{R})$ is not compact. Use this to conclude that $\mathsf{GL}(2;\mathbb{R})$ is neither closed nor bounded. ▼

Compact Lie groups are actually significantly easier to analyse than non-compact ones. Unfortunately, their beautiful theory will have to be left for the future. For now we turn to the main mathematical concept of this subject: Lie algebras.

## 2.2. **Lie algebras**

Let us say that an *algebra* is a vector space equipped with a bilinear multiplication that distributes over the vector space addition. If the multiplication is associative, then we have an *associative* algebra. If it has a multiplicative identity, then the algebra is *unital*. Finally, if the multiplication is commutative, then we have a commutative or *abelian* algebra. For example, $\mathbb{R}$ is a commutative associative algebra, as is $\mathbb{C}$. The set of $n \times n$ matrices is a non-commutative associative algebra. Each of these examples is unital.

Contrarily, a *Lie algebra* is a (generally) non-unital algebra $\mathfrak{g}$ with a (generally) non-associative non-abelian multiplication $[\cdot,\cdot]$ called the *Lie bracket* that satisfies

**Antisymmetry:** $[x,y] = -[y,x]$, for all $x,y \in \mathfrak{g}$;
**The Jacobi identity:** $[[x,y],z] + [[y,z],x] + [[z,x],y] = 0$, for all $x,y,z \in \mathfrak{g}$.

We write the (bilinear) multiplication as $[\cdot,\cdot]$ in order to prevent the temptation to use associativity. As usual, there is an underlying ground field that we shall generally take to be either $\mathbb{R}$ or $\mathbb{C}$, for simplicity.

Now, antisymmetry is a rather mild generalisation of commutativity, but the Jacobi identity is something of a strange replacement for associativity. Where does it come from? The answer lies in the relationship between Lie groups and algebras. Recall that we required a Lie group to be a smooth manifold so that we could do calculus on it. The Lie algebra of a Lie group $\mathsf{G}$ is, roughly speaking, what you obtain by differentiating at the group's identity $\mathbb{1}$.

More precisely, differentiating a smooth curve through $\mathbb{1} \in \mathsf{G}$ and evaluating at $\mathbb{1}$ gives a vector that is tangent to the curve at $\mathbb{1}$. The set of such vectors then forms a vector space of the same dimension as the Lie group that is tangent to the Lie group at $\mathbb{1}$. This is called the tangent space $T_{\mathbb{1}}\mathsf{G}$ of the Lie group at $\mathbb{1}$. One can make this idea precise with the tools of differential geometry, but here we shall have to rely on somewhat more heuristic methods. Specifically, an element $x \in T_{\mathbb{1}}\mathsf{G}$ should generate a smooth curve $\gamma$ in $\mathsf{G}$ with (local) parametrisation

$$\gamma_x(\varepsilon) = \mathbb{1} + \varepsilon x + O(\varepsilon^2). \tag{2.5}$$

Differentiating with respect to $\varepsilon$ and then setting $\varepsilon$ to 0 recovers $x$ from $\gamma_x(\varepsilon)$.

Group inversion therefore sends $\gamma_x(\varepsilon)$ to $\gamma_x(\varepsilon)^{-1} = \mathbb{1} - \varepsilon x + O(\varepsilon^2)$, *ie.* differentiating results in $x \mapsto -x$. Similarly, group multiplication corresponds to $(x, y) \mapsto x + y$. Neither of these is particularly exciting for the vector space $T_{\mathbb{1}}\mathsf{G}$. In particular, they do not give us an algebra structure. Moreover, they do not even capture the non-abelian nature of most Lie groups. To correct this, we will instead look at the $O(\varepsilon^2)$-terms in the *group commutator* $(g, h) \mapsto ghg^{-1}h^{-1}$.

**Exercise 4.** Show that writing

$$\gamma_x(\varepsilon) = \mathbb{1} + \varepsilon x + \varepsilon^2 x' + O(\varepsilon^3) \quad \text{and} \quad \gamma_x(\varepsilon)^{-1} = \mathbb{1} - \varepsilon x + \varepsilon^2 x'' + O(\varepsilon^3) \tag{2.6}$$

leads to the *non-linear* relation $x' + x'' = x^2$. Use this to show that

$$\gamma_x(\varepsilon)\gamma_y(\varepsilon)\gamma_x(\varepsilon)^{-1}\gamma_y(\varepsilon)^{-1} = \mathbb{1} + \varepsilon^2(xy - yx) + O(\varepsilon^3). \tag{2.7}$$

▼

**Exercise 5.** Show that if $x$ and $y$ belong to an arbitrary associative algebra $\mathsf{A}$, then replacing the product structure by the commutator

$$[x, y] = xy - yx \tag{2.8}$$

makes $\mathsf{A}$ into a Lie algebra. *ie.* show that the commutator is bilinear and satisfies both antisymmetry and the Jacobi identity. ▼

The idea then is that $T_{\mathbb{1}}\mathsf{G}$ should inherit an algebraic structure that satisfies the same axioms as the commutator does. We do not claim that this structure must be the commutator as $T_{\mathbb{1}}\mathsf{G}$ does not need to have the structure of an associative algebra in general — this proviso is actually a symptom of our heuristic approach to tangent spaces and is removed in the fully rigorous differential-geometric treatment. Nevertheless, it is straightforward to verify that $T_{\mathbb{1}}\mathsf{GL}(n)$ does inherit the matrix commutator as its Lie bracket. This will also be the case for Lie subgroups of $\mathsf{GL}(n)$. In any case, this is one way to motivate the axioms of a Lie algebra.

Enough of heuristics: let's look at some examples.

**Example 7.** An abelian Lie algebra $\mathfrak{g}$ must satisfy $[x, y] = [y, x]$, for all $x, y \in \mathfrak{g}$. However, antisymmetry requires that $[y, x] = -[x, y]$. Combining these, we conclude that an abelian Lie algebra has Lie bracket $[x, y] = 0$, for all $x, y \in \mathfrak{g}$. ▲

**Example 8.** If $\mathfrak{g}$ is a one-dimensional Lie algebra, with basis $\{x\}$ say, then the Lie bracket is completely determined by $[x, x]$. However, antisymmetry again forces this to be 0, so a one-dimensional Lie algebra is necessarily abelian. ▲

**Example 9.** There is a unique non-abelian Lie algebra $\mathfrak{g}$ with $\dim \mathfrak{g} = 2$. Let $\{x, y\}$ be a basis of $\mathfrak{g}$. Since $[x, x] = [y, y] = 0$ and $[x, y] = -[y, x]$, the Lie bracket is completely determined by $[x, y]$. We write

$$[x, y] = ax + by, \quad \text{for some scalars } a, b \text{ with } ab \neq 0. \tag{2.9}$$

Without loss of generality, we may assume that $a \neq 0$. Then, we may choose a new basis $\{x' = ax + by, y' = y/a\}$ with respect to which the Lie bracket takes the form

$$[x', y'] = [ax + by, y/a] = [ax, y/a] + [by, y] = [x, y] = ax + by = x'. \tag{2.10}$$

One can (and should) check that this Lie bracket indeed satisfies the Jacobi identity. ▲

**Example 10.** The Heisenberg uncertainty relation of quantum mechanics is encapsulated by the complex 3-dimensional Lie algebra spanned by $x$, $p$ and $\mathbb{1}$, with parameter $\hbar \in \mathbb{C}$ (called *Planck's constant*) and Lie bracket specified by

$$[x, p] = \mathrm{i}\hbar \mathbb{1} \quad \text{and} \quad [x, \mathbb{1}] = [p, \mathbb{1}] = 0. \tag{2.11}$$

This Lie algebra is often called the *Heisenberg algebra*; however, this term is also sometimes used for more general algebras. We remark that the dependence of the Heisenberg algebra on the parameter $\hbar$ is quite superficial. As long as $\hbar \neq 0$, we may rescale it to 1 by rescaling the basis element $x$ by a factor of $\hbar^{-1}$. The Heisenberg algebra with $\hbar = 0$ is of course different to those with $\hbar \neq 0$ because the former is abelian. ▲

The notion of uniqueness of a non-abelian two-dimensional Lie algebra, or of the Heisenberg algebra with $\hbar \neq 0$, presupposes that we know what it means for two Lie algebras to be the same, *ie.* isomorphic. We shall discuss this shortly. Whatever it means, however, it should be clear that the isomorphism class cannot change just because we are representing the Lie bracket in a different basis!

We continue looking at examples, in particular we describe some that arise as tangent spaces to Lie groups.

**Example 11.** The Lie group $\mathsf{U}(1)$ is realised by the elements $z \in \mathbb{C}$ that satisfy $z^* z = 1$. Differentiating this defining relation at 1 corresponds (heuristically) to expanding $z \in \mathsf{U}(1)$ as

$$z = 1 + \varepsilon x + O(\varepsilon^2) \tag{2.12}$$

and determining the first-order (in $\varepsilon$) relation that results. This gives

$$\left(1+\varepsilon x+O(\varepsilon)^2\right)^*\left(1+\varepsilon x+O(\varepsilon)^2\right)=1 \quad\Rightarrow\quad \varepsilon(x^*+x)+O(\varepsilon^2)=0, \tag{2.13}$$

*ie.* $x^*+x=0$. This forces $x$ to be pure imaginary, so we conclude that the tangent space at 1 to $\mathsf{U}(1)$ may be identified with the pure-imaginary complex numbers. The latter may of course be identified with $\mathbb{R}$. As it is one-dimensional, the Lie bracket is always zero. This abelian Lie algebra is denoted by $\mathfrak{u}(1)$. ▲

It is a very common convention to use capital letters to denote Lie groups and lowercase gothic/fraktur to denote the corresponding Lie algebras. We shall follow it here when not inconvenient.

**Example 12.** The **real** Lie group $\mathsf{SU}(2)$ consists of the **complex** $2\times 2$ matrices $A$ satisfying $A^\dagger A=\mathbb{1}$ and $\det A=1$. Expanding $A\in\mathsf{SU}(2)$ as $\mathbb{1}+\varepsilon x+O(\varepsilon^2)$, we find that the first defining relation becomes $x^\dagger+x=0$, whence we conclude that $x$ must be antihermitian. For the second defining relation, note that to first-order in $\varepsilon$, we have

$$\mathbb{1}+\varepsilon x\sim \mathsf{e}^{\varepsilon x}. \tag{2.14}$$

The corresponding relation is therefore $\operatorname{tr} x=0$ because

$$1\sim\det(\mathbb{1}+\varepsilon x)\sim\det \mathsf{e}^{\varepsilon x}=\mathsf{e}^{\varepsilon \operatorname{tr} x}\sim 1+\varepsilon\operatorname{tr} x. \tag{2.15}$$

The vector space of the Lie algebra $\mathfrak{su}(2)$ of $\mathsf{SU}(2)$ is therefore identified with the **real** vector space of $2\times 2$ **complex** antihermitian traceless matrices. As $\mathsf{SU}(2)$ is a matrix Lie group, the Lie bracket on $\mathfrak{su}(2)$ is the matrix commutator. ▲

**Exercise 6.** Explain why $\mathfrak{su}(2)$, as defined in terms of complex antihermitian traceless matrices, is a real Lie algebra and not a complex one. ▼

We should mention that the Lie group $\mathsf{SU}(2)$ is usually described in physics texts in terms of its "generators" or "infinitesimal generators". The infinitesimal qualifier should tip us off to the fact that they are here referring to something that should live in the Lie algebra $\mathfrak{su}(2)$. However, the physicists' favourite basis for these generators is given by the celebrated Pauli matrices

$$\sigma_1=\sigma_x=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \sigma_2=\sigma_y=\begin{pmatrix}0&-\mathsf{i}\\\mathsf{i}&0\end{pmatrix},\quad \sigma_3=\sigma_z=\begin{pmatrix}1&0\\0&-1\end{pmatrix}, \tag{2.16}$$

which are all quite obviously hermitian. Indeed, to get a bona fide basis of $\mathfrak{su}(2)$, one may multiply each of the Pauli matrices by $\mathsf{i}$ (or $-\mathsf{i}$). However, all of this is largely academic because physics texts will always implicitly work with a complex version ($\mathfrak{sl}(2;\mathbb{C})$ actually

— see Example 18) in which multiplying by $\mathsf{i}$ is allowed. Unfortunately, it is surprisingly common to read in such texts that the generators belong to $\mathfrak{su}(2)$ or, worse yet, $\mathsf{SU}(2)$.

But that shouldn't stop you from being correct! To help you, here's a rather long list of Lie algebras to get familiar with. Oh joy!

**Exercise 7.** Find the defining relations of the Lie algebras $\mathfrak{gl}(n)$, $\mathfrak{sl}(n)$, $\mathfrak{t}(n)$, $\mathfrak{t}'(n)$, $\mathfrak{o}(n)$, $\mathfrak{u}(n)$, $\mathfrak{su}(n)$, $\mathfrak{o}(p,q)$ and $\mathfrak{sp}(2n)$ that correspond to the Lie groups of Example 2. In each case, the Lie bracket is just the matrix commutator and you should check, when necessary, that the Lie algebras you have derived are actually closed under this Lie bracket. Show also that $\mathfrak{o}(n) = \mathfrak{so}(n)$ and $\mathfrak{o}(p,q) = \mathfrak{so}(p,q)$. ▼

You might also want to check that the dimension of each of your Lie algebras matches that of the corresponding Lie group. This must be the case because of the realisation of the former as the tangent space to the latter at $\mathbb{1}$.

Of course, there do exist (interesting) Lie algebras for which the Lie bracket is not given by a commutator.

**Exercise 8.**

- **(a)** Show that equipping $\mathbb{R}^3$ with the cross product, $[v, w] = v \times w$, results in a Lie algebra.
- **(b)** Rescale each of the Pauli matrices (2.16) by a complex number so that they are:
  - Actually elements of $\mathfrak{su}(2)$;
  - Have Lie brackets that match those of $(\mathbb{R}^3, \times)$ in the standard basis $\{i, j, k\}$.

Use this to conclude that $(\mathbb{R}^3, \times)$ and $\mathfrak{su}(2)$ are isomorphic as Lie algebras. [If you must, look at the next section to see the definition of an isomorphism of Lie algebras.] ▼

We finish with a (very!) brief exploration of how certain Lie groups can be reconstructed from their Lie algebras.

**Exercise 9.** Recall the three non-isomorphic 3-dimensional real Lie groups $\mathsf{SO}(3)$, $\mathsf{SU}(2)$ and $\mathsf{SL}(2;\mathbb{R})$ and their Lie algebras $\mathfrak{so}(3) \simeq \mathfrak{su}(2)$ and $\mathfrak{sl}(2;\mathbb{R})$. We consider, for each of these three Lie groups and algebras, the (always-convergent) matrix exponential

$$X \mapsto \mathsf{e}^X = \sum_{j=0}^{\infty} \frac{X^j}{j!}. \tag{2.17}$$

- **(a)** Show that the matrix exponential maps each of the three Lie algebras above into the corresponding Lie group.
- **(b)** Prove that the matrix exponential is not injective by finding a non-zero matrix $X$ in each of the three Lie algebras which is mapped to the identity matrix in the Lie group.

**(c)** Argue that for $X \neq 0$ sufficiently small (use whichever norm floats your boat) in each of the three Lie algebras, $\mathsf{e}^X$ is not the identity. (The matrix exponential is therefore *locally injective*.)

**(d)** For $\mathsf{SO}(3)$ and $\mathsf{SU}(2)$, one can show that their compactness implies that their matrix exponential maps are surjective. Show that the matrix exponential for $\mathsf{SL}(2;\mathbb{R})$ is not surjective by:
- Proving that $X \in \mathfrak{sl}(2;\mathbb{R})$ implies that $\operatorname{tr} \mathsf{e}^X \geqslant -2$.
- Using this to exhibit an element of $\mathsf{SL}(2;\mathbb{R})$ that is not in the image of the matrix exponential. ▼

## 2.3. **Lie algebra morphisms, subalgebras and ideals**

As promised, we shall now discuss when two Lie algebras may be regarded as "the same". Abstract algebra has all the answers, as usual, but there are (completely standard) definitions to absorb.

A *homomorphism* $\phi\colon \mathfrak{g} \to \mathfrak{h}$ of Lie algebras is a linear map that sends the Lie bracket of $\mathfrak{g}$ to that of $\mathfrak{h}$:

$$\phi([x,y]_{\mathfrak{g}}) = [\phi(x),\phi(y)]_{\mathfrak{h}}. \tag{2.18}$$

If $\mathfrak{g} = \mathfrak{h}$ in this definition, then the homomorphism $\phi$ is said to be an *endomorphism* of Lie algebras.

As usual, the homomorphism $\phi$ in (2.18) is said to be *injective* if its *kernel*

$$\ker \phi = \{x \in \mathfrak{g} : \phi(x) = 0\} \tag{2.19}$$

is 0 and it is said to be *surjective* if its *image*

$$\operatorname{im} \phi = \{y \in \mathfrak{h} : y = \phi(x) \text{ for some } x \in \mathfrak{g}\} \tag{2.20}$$

is $\mathfrak{h}$. A homomorphism that is *bijective*, meaning both injective and surjective, is an *isomorphism*. An endomorphism that is also an isomorphism is called an *automorphism*.

**Exercise 10.** Show that the non-abelian (*ie.* $\hbar \neq 0$) Heisenberg algebra of Example 10 is isomorphic to $\mathfrak{t}'(3;\mathbb{C})$. ▼

**Exercise 11.** Construct an isomorphism between $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$. Show that $\mathfrak{sl}(2;\mathbb{R})$ is not isomorphic to either. [Note that because these are real Lie algebras, any isomorphism must be real: with respect to any choice of bases, the representing matrix must have real entries.] ▼

An important remark is that the isomorphism $\mathfrak{su}(2) \simeq \mathfrak{so}(3)$ does **not** imply that the Lie groups $\mathsf{SU}(2)$ and $\mathsf{SO}(3)$ are isomorphic. In fact, they are not. One easy way to see this is to show that their centres have different sizes. Recall that the *centre* $\mathsf{Z}(\mathsf{G})$ of a Lie group $\mathsf{G}$ is just the set of elements that commute with all other elements:

$$\mathsf{Z}(\mathsf{G}) = \{g \in \mathsf{G} : gh = hg \text{ for all } h \in \mathsf{G}\}. \tag{2.21}$$

**Exercise 12.**

**(a)** Show that $\left|\mathsf{Z}(\mathsf{SU}(2))\right| = 2$ by explicitly parametrising an arbitrary matrix of $\mathsf{SU}(2)$, as in Example 3, and checking if the result commutes with $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}$, both of which are in $\mathsf{SU}(2)$.

**(b)** Show that $\left|\mathsf{Z}(\mathsf{SO}(3))\right| = 1$ by checking which matrices commute with a diagonal matrix in $\mathsf{SO}(3)$ and then checking which also commute with $\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$, both of which are in $\mathsf{SO}(3)$. ▼

A *Lie subalgebra* $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ is a vector subspace that is closed under the Lie bracket:

$$x, y \in \mathfrak{h} \quad \Rightarrow \quad [x, y] \in \mathfrak{h}. \tag{2.22}$$

A Lie subalgebra is therefore a Lie algebra in its own right. We often write the condition to be a Lie subalgebra as $[\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{h}$. A stronger concept is for $\mathfrak{h}$ to be an *ideal* of $\mathfrak{g}$, which requires that $[\mathfrak{g}, \mathfrak{h}] \subseteq \mathfrak{h}$, *ie.*

$$x \in \mathfrak{g} \text{ and } y \in \mathfrak{h} \quad \Rightarrow \quad [x, y] \in \mathfrak{h}. \tag{2.23}$$

Note that $0$ and $\mathfrak{g}$ are always ideals of $\mathfrak{g}$.

**Example 13.** $\mathfrak{t}(2)$ is a subalgebra of $\mathfrak{gl}(2)$, but is not an ideal because

$$\left[\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right] = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \notin \mathfrak{t}(2). \tag{2.24}$$ ▲

**Example 14.** $\mathfrak{sl}(n)$ is an ideal of $\mathfrak{gl}(n)$ because the trace of a commutator is always zero. ▲

**Exercise 13.** Show that if $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, then the inclusion map $\iota \colon \mathfrak{h} \to \mathfrak{g}$ is an injective homomorphism of Lie algebras. ▼

**Exercise 14.** Show that the kernel $\ker \phi$ of a Lie algebra homomorphism $\phi \colon \mathfrak{g} \to \mathfrak{h}$ is an ideal of $\mathfrak{g}$ and that the image $\operatorname{im} \phi$ of $\phi$ is a Lie subalgebra of $\mathfrak{h}$. Give an example to show that $\operatorname{im} \phi$ need not be an ideal. ▼

**Exercise 15.** Show that the *centre* $\mathfrak{z}(\mathfrak{g})$ of a Lie algebra $\mathfrak{g}$, defined by

$$\mathfrak{z}(\mathfrak{g}) = \{x \in \mathfrak{g} : [x, \mathfrak{g}] = 0\} \tag{2.25}$$

is an abelian ideal of $\mathfrak{g}$. ▼

Ideals are distinguished from subalgebras because they can be used to form quotient Lie algebras. Roughly speaking, the quotient of $\mathfrak{g}$ by an ideal $\mathfrak{h} \subseteq \mathfrak{g}$ is the Lie algebra $\mathfrak{g}/\mathfrak{h}$ that is obtained by identifying $\mathfrak{h}$ with zero. In this picture, we need to quotient by an ideal because the Lie bracket of everything with 0 must give 0. More formally, the elements of $\mathfrak{g}/\mathfrak{h}$ are the equivalence classes $\overline{x}$, for $x \in \mathfrak{g}$, defined by

$$\overline{x} = \{y \in \mathfrak{g} : y - x \in \mathfrak{h}\}, \tag{2.26}$$

the vector space structure is given by

$$\overline{ax + by} = a\overline{x} + b\overline{y}, \quad \text{for all scalars } a, b \text{ and } x, y \in \mathfrak{g}, \tag{2.27}$$

and the Lie bracket is given by

$$[\overline{x}, \overline{y}]_{\mathfrak{g}/\mathfrak{h}} = \overline{[x, y]_{\mathfrak{g}}}, \quad \text{for all } x, y \in \mathfrak{g}. \tag{2.28}$$

Choosing different representatives $x'$ and $y'$ for $\overline{x}$ and $\overline{y}$, respectively, we immediately see that $x' - x$ and $y' - y$ belong to $\mathfrak{h}$. The difference in the right-hand side of the Lie bracket (2.28) is therefore

$$\overline{[x', y']_{\mathfrak{g}}} - \overline{[x, y]_{\mathfrak{g}}} = \overline{[x' - x, y' - y]_{\mathfrak{g}}} + \overline{[x' - x, y]_{\mathfrak{g}}} + \overline{[x, y' - y]_{\mathfrak{g}}} = 0, \tag{2.29}$$

since all three brackets on the right-hand side belong to $\mathfrak{h}$, by the ideal property. The Lie bracket of $\mathfrak{g}/\mathfrak{h}$ is therefore well-defined when $\mathfrak{h}$ is an ideal of $\mathfrak{g}$.

**Exercise 16.** Show that if $\mathfrak{h}$ is an ideal of $\mathfrak{g}$, then the canonical quotient map $\pi \colon \mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$, given by $x \mapsto \overline{x}$, is a surjective homomorphism of Lie algebras. ▼

**Exercise 17.** Show that if $\phi \colon \mathfrak{g} \to \mathfrak{h}$ is a homomorphism of Lie algebras, then

$$\frac{\mathfrak{g}}{\ker \phi} \simeq \operatorname{im} \phi, \tag{2.30}$$

as Lie algebras (*cf.* Exercise 14). ▼

So ideals are somehow better than Lie subalgebras. It is therefore satisfying that the set of ideals of a given Lie algebra is also closed under the operations of intersection and sum. Given ideals $\mathfrak{h}_1$ and $\mathfrak{h}_2$ of $\mathfrak{g}$, define

$$\begin{aligned} \mathfrak{h}_1 \cap \mathfrak{h}_2 &= \{x \in \mathfrak{g} : x \in \mathfrak{h}_1 \text{ and } x \in \mathfrak{h}_2\} \\ \text{and} \quad \mathfrak{h}_1 + \mathfrak{h}_2 &= \{x_1 + x_2 \in \mathfrak{g} : x_1 \in \mathfrak{h}_1 \text{ and } x_2 \in \mathfrak{h}_2\}. \end{aligned} \tag{2.31}$$

**Exercise 18.**

**(a)** Show that both $\mathfrak{h}_1 \cap \mathfrak{h}_2$ and $\mathfrak{h}_1 + \mathfrak{h}_2$ are ideals of $\mathfrak{g}$ when $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are.

**(b)** Give examples, with $\mathfrak{g} = \mathfrak{sl}(2; \mathbb{C})$, of Lie subalgebras $\mathfrak{h}_1$ and $\mathfrak{h}_2$ whose sum $\mathfrak{h}_1 + \mathfrak{h}_2$ is not a Lie subalgebra of $\mathfrak{g}$. ▼

A very important special case is when one has two ideals $\mathfrak{h}_1$ and $\mathfrak{h}_2$ of $\mathfrak{g}$ with $\mathfrak{h}_1 \cap \mathfrak{h}_2 = 0$. Then, we say that $\mathfrak{h}_1 + \mathfrak{h}_2$ is the *direct sum* of $\mathfrak{h}_1$ and $\mathfrak{h}_2$. You know that the direct sum is important because we use a special notation to distinguish it from the ordinary sum: we write $\mathfrak{h}_1 \oplus \mathfrak{h}_2$ instead of $\mathfrak{h}_1 + \mathfrak{h}_2$.

**Exercise 19.** Show that $\mathfrak{gl}(n)$ is the direct sum of two ideals, one isomorphic to $\mathfrak{sl}(n)$ and the other to $\mathfrak{u}(1)$, if the ground field is $\mathbb{R}$, or $\mathfrak{gl}(1;\mathbb{C})$, if the ground field is $\mathbb{C}$. ▼

**Exercise 20.** Show that if $\mathfrak{g} = \mathfrak{h}_1 \oplus \mathfrak{h}_2$, then $\mathfrak{g}/\mathfrak{h}_1 \simeq \mathfrak{h}_2$. ▼

We now come to some of the most important definitions in Lie algebra theory. A Lie algebra is said to be *simple* if it is non-abelian and has no ideals except for 0 and itself. (It is traditional to regard the abelian Lie algebras $\mathfrak{u}(1)$ and $\mathfrak{gl}(1;\mathbb{C})$ as **not** simple, even though they have only two ideals, in order to simplify the statements of many results.) A Lie algebra is said to be *semisimple* if it may be written as a direct sum of its simple ideals, each of which is therefore a simple Lie algebra in its own right. Finally, a Lie algebra is said to be *reductive* if it may be written as a direct sum of its centre and its simple ideals.

At this point, it is convenient to quote a standard technical result (a special case of the Krull-Schmidt Theorem of abstract algebra) that guarantees the essential uniqueness of the decomposition of a reductive Lie algebra into its centre and simple ideals — the only ambiguity is in the ordering of the ideals. As semisimple Lie algebras are special cases of reductive ones, this uniqueness applies to them too.

**Proposition 2.1** (Krull-Schmidt)**.** *If $\mathfrak{g}$ is a reductive Lie algebra with two decompositions involving its centre and its simple ideals, eg.*

$$\mathfrak{g} \simeq \mathfrak{z}(\mathfrak{g}) \oplus \mathfrak{h}_1 \oplus \cdots \oplus \mathfrak{h}_m \quad \text{and} \quad \mathfrak{g} \simeq \mathfrak{z}(\mathfrak{g}) \oplus \mathfrak{l}_1 \oplus \cdots \oplus \mathfrak{l}_n, \tag{2.32}$$

*then $m = n$ and there is a permutation $\sigma$ of $\{1,\dots,n\}$ such that $\mathfrak{h}_i \simeq \mathfrak{l}_{\sigma(i)}$ for all $i = 1,\dots,n$.*

*Proof.* Set $\mathfrak{g}_m = \mathfrak{g}/\mathfrak{z}(\mathfrak{g})$. By Exercise 20, the two decompositions give

$$\mathfrak{g}_m \simeq \mathfrak{h}_1 \oplus \cdots \oplus \mathfrak{h}_m \simeq \mathfrak{l}_1 \oplus \cdots \oplus \mathfrak{l}_n. \tag{2.33}$$

Let $p_m \colon \mathfrak{g}_m \longrightarrow \mathfrak{h}_m$ and $q_j \colon \mathfrak{g}_m \longrightarrow \mathfrak{l}_j$, $j = 1,\dots,n$, denote the projections onto the indicated simple ideals. Then, $\sum_{j=1}^n q_j = \mathbb{1}_{\mathfrak{g}_m}$.

We consider the restrictions

$$\pi_j = p_m\big|_{\mathfrak{l}_j} \circ q_j\big|_{\mathfrak{h}_m} : \mathfrak{h}_m \longrightarrow \mathfrak{h}_m. \tag{2.34}$$

Because $\ker \pi_j$ is an ideal of the simple Lie algebra $\mathfrak{h}_m$ (Exercise 14), it is either $\mathfrak{h}_m$ or 0. In the former case, $\pi_j$ is 0 and in the latter, it is an isomorphism (since $\dim \mathfrak{h}_m < \infty$).

However,

$$\sum_{j=1}^{n} \pi_j = p_m \circ \sum_{j=1}^{n} q_j\big|_{\mathfrak{h}_m} = p_m \circ \mathbb{1}_{\mathfrak{h}_m} = \mathbb{1}_{\mathfrak{h}_m}, \tag{2.35}$$

so there must be at least one $\pi_j$, $j = 1, \dots, n$, which is an isomorphism. Pick such a $j$ and call it $\sigma(m)$.

Now, the kernel of the restriction of $q_{\sigma(m)}$ to $\mathfrak{h}_m$ is an ideal, hence it is 0 or $\mathfrak{h}_m$. But, the latter contradicts $\pi_{\sigma(m)} \neq 0$, hence it is zero and so $q_{\sigma(m)} \colon \mathfrak{h}_m \to \mathfrak{l}_{\sigma(m)}$ is injective. It follows that $\dim \mathfrak{h}_m \leqslant \dim \mathfrak{l}_{\sigma(m)}$. If this inequality were strict, then the restriction of $p_m$ to $\mathfrak{l}_{\sigma(m)}$ would necessarily have a non-zero kernel, hence this kernel would be $\mathfrak{l}_{\sigma(m)}$, again contradicting $\pi_{\sigma(m)} \neq 0$. We conclude that $\dim \mathfrak{h}_m = \dim \mathfrak{l}_{\sigma(m)}$, proving that $\mathfrak{h}_m \simeq \mathfrak{l}_{\sigma(m)}$.

We can repeat this argument with $\mathfrak{g}_{m-1} = \mathfrak{g}_m/\mathfrak{h}_m$, noting that its decomposition into the $\mathfrak{l}_j$ will now omit $\mathfrak{l}_{\sigma(m)}$. In this way, we discover that $\mathfrak{h}_{m-1} \simeq \mathfrak{l}_{\sigma(m-1)}$, for some $\sigma(m-1) \neq \sigma(m)$ in $\{1, \dots, n\}$. Continuing, we find that every $\mathfrak{h}_i$ is isomorphic to some $\mathfrak{l}_{\sigma(i)}$ and that the $\sigma$ so constructed in a permutation of $\{1, \dots, m\}$. This proves that $m \leqslant n$. However, if $m < n$, then quotienting away all of the $\mathfrak{h}_i$ would lead to the direct sum of the remaining $\mathfrak{l}_j$ being isomorphic to 0, a contradiction. We therefore conclude that $m = n$ and the proof is complete. ■

**Example 15.** By Exercise 19, the Lie algebra $\mathfrak{gl}(n)$ is not simple, because it has non-zero proper ideals, and is not semisimple, because the ideal isomorphic to $\mathfrak{u}(1)$ or $\mathfrak{gl}(1;\mathbb{C})$ obviously cannot be written as a direct sum of simple ideals. We shall see in Example 17 below that $\mathfrak{sl}(2)$ is simple, from which it follows that $\mathfrak{gl}(2)$ is reductive. In fact, $\mathfrak{gl}(n)$ is reductive for all $n \in \mathbb{Z}_{\geqslant 0}$ (see Exercise 53). ▲

**Example 16.** The real Lie algebra $\mathfrak{t}(2)$ is not simple nor semisimple because the identity matrix spans an abelian ideal isomorphic to $\mathfrak{u}(1)$. Because the trace of a commutator is zero, the traceless matrices of $\mathfrak{t}(2)$ form an ideal $\mathfrak{h}$. We therefore have $\mathfrak{t}(2) \simeq \mathfrak{u}(1) \oplus \mathfrak{h}$. It is easy to check that $\mathfrak{h}$ is non-abelian and that $e = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ spans an ideal isomorphic to $\mathfrak{u}(1)$ in $\mathfrak{h}$. However, this ideal has no complement in $\mathfrak{h}$: if it did, then $\mathfrak{h}$ would be isomorphic to $\mathfrak{u}(1) \oplus \mathfrak{u}(1)$, contradicting the fact that $\mathfrak{h}$ is non-abelian. Thus, $\mathfrak{h}$ cannot be written as a direct sum of ideals, hence it is not reductive and so neither is $\mathfrak{t}(2)$. ▲

If fact, $\mathfrak{t}(n)$ is never reductive (for $n > 1$) because it has the nasty property that it possesses an ideal $\mathfrak{h}$ with no complement: there exists an ideal $\mathfrak{h} \subset \mathfrak{t}(n)$, but $\mathfrak{t}(n) \neq \mathfrak{h} \oplus \mathfrak{l}$ for any ideal $\mathfrak{l} \subset \mathfrak{t}(n)$. This bad behaviour means that one is forced to resort to quotients to analyse the structure of $\mathfrak{t}(n)$, *ie.* instead of splitting off $\mathfrak{h}$ as a direct summand, we have to consider $\mathfrak{t}(n)/\mathfrak{h}$. Semisimplicity (and reductivity) avoids this bad behaviour completely!

**Exercise 21.** Define the *derived subalgebra* of a Lie algebra $\mathfrak{g}$ to be

$$[\mathfrak{g},\mathfrak{g}] = \operatorname{span}\{[x,y] : x,y \in \mathfrak{g}\}. \tag{2.36}$$

Show that $[\mathfrak{g},\mathfrak{g}]$ is an ideal of $\mathfrak{g}$. Deduce that if $\mathfrak{g}$ is simple, then $[\mathfrak{g},\mathfrak{g}] = \mathfrak{g}$. Conclude that the same is true for semisimple Lie algebras, but that it is false for reductive Lie algebras in general. ▼

**Example 17.** We claim that $\mathfrak{sl}(2)$ is simple. To see this, assume the opposite, *ie.* that $\mathfrak{sl}(2)$ has a non-zero proper ideal $\mathfrak{h}$. Its dimension is necessarily 1 or 2.

If $\dim \mathfrak{h} = 1$, then $\mathfrak{h}$ is spanned by a single matrix. One can use the Lie bracket to explicitly show that this is inconsistent with the property of being an ideal, but this is dull. Here is a slightly more refined argument. Consider an element $h \in \mathfrak{sl}(2)$ and a linear map $\operatorname{ad}(h) \colon \mathfrak{sl}(2) \longrightarrow \mathfrak{sl}(2)$ defined by

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \operatorname{ad}(h)\, x = [h,x], \quad \text{for } x \in \mathfrak{sl}(2). \tag{2.37}$$

For $\mathfrak{h}$ to be an ideal, any spanning element must be an eigenvector of $\operatorname{ad}(h)$. It is easy to check that this reduces the possible spanning elements to (non-zero multiples of) $h$,

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{2.38}$$

However, $h$ does not span an ideal, because *eg.* $[h,e] = 2e$, and neither do $e$ or $f$, because $[e,f] = h$. This contradiction therefore rules out a one-dimensional ideal of $\mathfrak{sl}(2)$.

If $\dim \mathfrak{h} = 2$, one can again arrive at an inconsistency using direct but very tedious arguments. However, we shall proceed by using the quotient map $\pi \colon \mathfrak{sl}(2) \longrightarrow \mathfrak{sl}(2)/\mathfrak{h}$, which is a surjective Lie algebra homomorphism by Exercise 16. Now note that

$$\begin{aligned} \pi\big([\mathfrak{sl}(2),\mathfrak{sl}(2)]\big) &= \big[\pi(\mathfrak{sl}(2)),\pi(\mathfrak{sl}(2))\big] = [\mathfrak{sl}(2)/\mathfrak{h},\mathfrak{sl}(2)/\mathfrak{h}] = 0 \\ &\Rightarrow \quad [\mathfrak{sl}(2),\mathfrak{sl}(2)] \subseteq \mathfrak{h}, \end{aligned} \tag{2.39}$$

since $\mathfrak{sl}(2)/\mathfrak{h}$ is one-dimensional, hence abelian. However, the easily checked Lie brackets $[h,e] = 2e$, $[h,f] = -2f$ and $[e,f] = h$ show that the derived subalgebra of $\mathfrak{sl}(2)$ is actually $\mathfrak{sl}(2)$, since $\{e,h,f\}$ is a spanning set. This contradiction completes the proof. ▲

**Exercise 22.** Use the isomorphism of Exercise 8 to argue geometrically that $\mathfrak{su}(2)$ has no 2-dimensional Lie subalgebra, let alone a 2-dimensional ideal. ▼

## 2.4. **Complexifications**

From here on, unless explicitly noted to the contrary, we shall be exclusively concerned with **complex** Lie algebras. The reason for this is that their study relies heavily on linear algebra and $\mathbb{C}$ is a much better field for this than $\mathbb{R}$ (*eg*. eigenvalues always exist). In many applications, in particular those involving quantum mechanics, the natural underlying ground field is anyway $\mathbb{C}$.

If we are given a real Lie algebra $\mathfrak{g}$, there is a canonical way to turn it into a complex Lie algebra, called the *complexification* of $\mathfrak{g}$, that we shall denote by $\mathfrak{g}^{\mathbb{C}}$. At the formal level, this can be summarised by the following definition: $\mathfrak{g}^{\mathbb{C}} = \mathbb{C} \otimes_{\mathbb{R}} \mathfrak{g}$. Here, the subscript on the tensor product indicates that only real scalars may pass from the first factor to the second:

$$r \otimes_{\mathbb{R}} x = 1 \otimes_{\mathbb{R}} (rx), \quad \text{for all } r \in \mathbb{R} \text{ and } x \in \mathfrak{g}. \tag{2.40}$$

Complex scalars such as $\mathsf{i}$ cannot pass through. This captures the fact that multiplying $\mathsf{i}$ and $x$ doesn't make sense when $x$ belongs to a real Lie algebra. The product of $\mathsf{i}$ and $x$ (which only makes sense in $\mathfrak{g}^{\mathbb{C}}$) is instead represented by the formal tensor product $\mathsf{i} \otimes_{\mathbb{R}} x$. Of course, one generally drops these formal tensor product symbols once one has become comfortable with this approach to the complexification.

Admittedly, this is all a bit abstract, so it may be helpful to note that, at the level of bases, complexification just amounts to allowing linear combinations with complex coefficients instead of only real ones. Obviously, the complexification of a complex Lie algebra is just the original complex Lie algebra again.

One nice feature of complexification is that it amalgamates many of the real Lie algebras that we've already met.

**Example 18.** It should be clear that the complexification of $\mathfrak{sl}(2;\mathbb{R})$ is $\mathfrak{sl}(2;\mathbb{R})^{\mathbb{C}} = \mathfrak{sl}(2;\mathbb{C})$. We therefore ask instead for the complexification of $\mathfrak{su}(2)$.

Recall that $\mathfrak{su}(2)$ consists of the traceless antihermitian complex $2\times 2$ matrices. A basis is therefore given by

$$\left\{ \begin{pmatrix} 0 & \mathsf{i} \\ \mathsf{i} & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} \mathsf{i} & 0 \\ 0 & -\mathsf{i} \end{pmatrix} \right\}. \tag{2.41}$$

In the complexification $\mathfrak{su}(2)^{\mathbb{C}}$, we are allowed to multiply these basis elements by $-\mathsf{i}$, 1 and $-\mathsf{i}$, respectively, giving a new basis of $\mathfrak{su}(2)^{\mathbb{C}}$:

$$\left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}. \tag{2.42}$$

Taking half the sum and the difference of the first two matrices, we arrive at yet another basis:

$$\left\{ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}. \tag{2.43}$$

This we recognise as the basis $\{e, f, h\}$ of $\mathfrak{sl}(2;\mathbb{C})$ that we used in Example 17. As both $\mathfrak{su}(2)^{\mathbb{C}}$ and $\mathfrak{sl}(2;\mathbb{C})$ have the matrix commutator for their Lie bracket, it follows that these complex Lie algebras are isomorphic.

We conclude that even though $\mathfrak{su}(2)$ and $\mathfrak{sl}(2;\mathbb{R})$ are non-isomorphic real Lie algebras (Exercise 11), their complexifications are isomorphic: $\mathfrak{su}(2)^{\mathbb{C}} \simeq \mathfrak{sl}(2;\mathbb{R})^{\mathbb{C}} \simeq \mathfrak{sl}(2;\mathbb{C})$. ▲

Note that if we have a real Lie algebra $\mathfrak{g}$ and a complex Lie algebra $\mathfrak{h}$ such that $\mathfrak{h}$ is (isomorphic to) the complexification $\mathfrak{g}^{\mathbb{C}}$ of $\mathfrak{g}$, then $\mathfrak{g}$ is said to be a *real form* of $\mathfrak{h}$. From the previous example, we see that $\mathfrak{su}(2)$ and $\mathfrak{sl}(2;\mathbb{R})$ are non-isomorphic real forms of $\mathfrak{sl}(2;\mathbb{C})$.

**Example 19.** Here are some further complexifications:

**(a)** $\mathfrak{gl}(n;\mathbb{R})^{\mathbb{C}} \simeq \mathfrak{u}(n)^{\mathbb{C}} \simeq \mathfrak{gl}(n;\mathbb{C})$.
**(b)** $\mathfrak{sl}(n;\mathbb{R})^{\mathbb{C}} \simeq \mathfrak{su}(n)^{\mathbb{C}} \simeq \mathfrak{sl}(n;\mathbb{C})$.
**(c)** $\mathfrak{t}(n;\mathbb{R})^{\mathbb{C}} \simeq \mathfrak{t}(n;\mathbb{C})$ and $\mathfrak{t}'(n;\mathbb{R})^{\mathbb{C}} \simeq \mathfrak{t}'(n;\mathbb{C})$.
**(d)** $\mathfrak{o}(p,q)^{\mathbb{C}} = \mathfrak{so}(p,q)^{\mathbb{C}} \simeq \mathfrak{so}(p+q;\mathbb{C}) = \{A \in \mathfrak{gl}(p+q;\mathbb{C}) : A^{\mathsf{T}} = -A \text{ and } \operatorname{tr} A = 0\}$.
**(e)** $\mathfrak{sp}(2n)^{\mathbb{C}} \simeq \mathfrak{sp}(2n;\mathbb{C}) = \{A \in \mathfrak{gl}(2n;\mathbb{C}) : A^{\mathsf{T}} J = -JA\}$.

We refer to (2.1) for the definition of $J$. ▲

**Exercise 23.** Explain why the complexifications of $\mathfrak{o}(p,q)$ and $\mathfrak{so}(p,q)$ only depend on the sum of $p$ and $q$. ▼

From now on, if we neglect to specify the field in the notation being used for a Lie algebra, then the field is understood to be $\mathbb{C}$. In particular, $\mathfrak{gl}(n)$, $\mathfrak{sl}(n)$, $\mathfrak{so}(n)$ and $\mathfrak{sp}(2n)$ will hereafter refer to $\mathfrak{gl}(n;\mathbb{C})$, $\mathfrak{sl}(n;\mathbb{C})$, $\mathfrak{so}(n;\mathbb{C})$ and $\mathfrak{sp}(2n;\mathbb{C})$, respectively.

Now, of the Lie algebras of most interest to us, the (finite-dimensional) simple complex ones, we have now met *all* of them bar five. These five are called exceptional and they are denoted as in the following table of all simple complex Lie algebras.

| $\mathfrak{g}$ | $\mathfrak{sl}(n),\ n \geqslant 2$ | $\mathfrak{so}(n), n = 3 \text{ or } n \geqslant 5$ | $\mathfrak{sp}(2n), n \geqslant 1$ | $\mathfrak{g}_2$ | $\mathfrak{f}_4$ | $\mathfrak{e}_6$ | $\mathfrak{e}_7$ | $\mathfrak{e}_8$ |
|---|---|---|---|---|---|---|---|---|
| $\dim \mathfrak{g}$ | $n^2 - 1$ | $\frac{1}{2}n(n-1)$ | $n(2n+1)$ | 14 | 52 | 78 | 133 | 248 |

Moreover, the only coincidences among the simple complex Lie algebras are

$$\mathfrak{sl}(2) \simeq \mathfrak{so}(3) \simeq \mathfrak{sp}(2), \quad \mathfrak{sl}(4) \simeq \mathfrak{so}(6) \quad \text{and} \quad \mathfrak{so}(5) \simeq \mathfrak{sp}(4). \tag{2.44}$$

We remark that $\mathfrak{sl}(1)$ and $\mathfrak{so}(1)$ are both 0, $\mathfrak{so}(2) \simeq \mathfrak{gl}(1)$ is abelian (and one-dimensional) and $\mathfrak{so}(4) \simeq \mathfrak{sl}(2) \oplus \mathfrak{sl}(2)$ is the only one that is semisimple but not simple.

We conclude with an exercise that hints at a direction that one can take to classify certain types of real Lie algebras, *eg.* semisimple ones, assuming that the corresponding complex ones have already been classified.

**Exercise 24.** An *adjoint* $^\dagger$ on a complex Lie algebra $\mathfrak{g}$ is a conjugate-linear map, *ie.*

$$(ax+by)^\dagger = a^* x^\dagger + b^* y^\dagger, \quad \text{for all } a,b \in \mathbb{C} \text{ and } x,y \in \mathfrak{g}, \tag{2.45}$$

from $\mathfrak{g}$ to itself that satisfies

$$(x^\dagger)^\dagger = x \quad \text{and} \quad [x,y]^\dagger = [y^\dagger, x^\dagger], \quad \text{for all } x,y \in \mathfrak{g}. \tag{2.46}$$

- **(a)** Show that $e^\dagger = f$ and $h^\dagger = h$ defines an adjoint on $\mathfrak{sl}(2;\mathbb{C})$. (This adjoint is suggested by the defining representation — see (2.51) below.)
- **(b)** Show that if $^\dagger$ is an adjoint on $\mathfrak{g}$, then the set

$$\mathfrak{g}_{\mathbb{R}} = \{x \in \mathfrak{g} : x^\dagger = -x\} \tag{2.47}$$

  of antihermitian elements of $\mathfrak{g}$ forms a **real** Lie algebra. It's called a *real form* of $\mathfrak{g}$.
- **(c)** Show that the real form of $\mathfrak{sl}(2;\mathbb{C})$ defined by the adjoint in a is isomorphic to $\mathfrak{su}(2)$.
- **(d)** Show that $e^\dagger = -e$, $h^\dagger = -h$ and $f^\dagger = -f$ defines another adjoint on $\mathfrak{sl}(2;\mathbb{C})$ and determine the (isomorphism class of the) corresponding real form. ▼

### 2.5. **Representations and modules**

A representation of a Lie algebra $\mathfrak{g}$ is, roughly speaking, a way to assign to each element in $\mathfrak{g}$ a matrix (or a linear transformation) so that the Lie bracket of $\mathfrak{g}$ becomes the matrix commutator. More formally, a *representation* of $\mathfrak{g}$ on a vector space $V$ is a Lie algebra homomorphism $\pi \colon \mathfrak{g} \to \mathfrak{gl}(V)$. Recall that this means that $\pi$ is a linear map satisfying

$$\pi([x,y]_{\mathfrak{g}}) = [\pi(x),\pi(y)]_{\mathfrak{gl}(V)} = \pi(x)\pi(y) - \pi(y)\pi(x), \quad \text{for all } x,y \in \mathfrak{g}. \tag{2.48}$$

The *dimension* of the representation $\pi$ is the dimension of $V$. Generally, the field underlying $V$ will match the field underlying $\mathfrak{g}$. In our case, this means that we shall assume that $V$ is complex unless otherwise noted.

**Example 20.** A representation $\pi \colon \mathfrak{g} \to \mathfrak{gl}(V)$ is said to be *trivial* if $\pi(x) = 0$ for all $x \in \mathfrak{g}$. In other words, if $\pi = 0$ (which is obviously a Lie algebra homomorphism for any choice of $V$). ▲

**Exercise 25.** Use Exercise 21 to prove that a one-dimensional representation of a semisimple Lie algebra is always trivial. Give an example of a non-trivial one-dimensional representation of a non-simple Lie algebra. ▼

**Example 21.** The *adjoint representation* of $\mathfrak{g}$ is, *cf.* (2.37), the Lie algebra homomorphism $\operatorname{ad} \colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ defined by

$$\operatorname{ad}(x)\, y = [x,y], \quad \text{for all } x,y \in \mathfrak{g}. \tag{2.49}$$

▲

**Exercise 26.** Show that ad is actually a homomorphism of Lie algebras. ▼

Each of the Lie algebras $\mathfrak{gl}(n)$, $\mathfrak{sl}(n)$, $\mathfrak{so}(n)$ and $\mathfrak{sp}(2n)$ has a *defining representation*, this being the homomorphism that maps each element of the abstract Lie algebra to the matrix that we used in Section 2.2 to define it.

**Example 22.** $\mathfrak{sl}(2)$ may be defined as the complex Lie algebra spanned by the abstract elements $e$, $h$ and $f$ with Lie bracket determined by

$$[h,e] = 2e, \quad [h,f] = -2f \quad \text{and} \quad [e,f] = h. \tag{2.50}$$

The defining representation of $\mathfrak{sl}(2)$ is then the linear map $\pi \colon \mathfrak{sl}(2) \longrightarrow \mathfrak{gl}(2)$ defined by

$$\pi(e) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \pi(h) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \pi(f) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{2.51}$$

Note that this representation of $\mathfrak{sl}(2)$ is two-dimensional while the adjoint representation of $\mathfrak{sl}(2)$ has dimension 3. The defining representation of $\mathfrak{sl}(2)$ is often also referred to as the *fundamental representation* (for reasons that we'll get to later). ▲

**Exercise 27.** Determine $\pi(e)$ so that, along with

$$\pi(h) = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix} \quad \text{and} \quad \pi(f) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{2.52}$$

$\pi$ defines a four-dimensional representation of $\mathfrak{sl}(2)$. ▼

It is common for mathematicians (and good physicists) to prefer to emphasise the vector space $V$ that $\mathfrak{g}$ acts on rather than the homomorphism $\pi$. Then, we speak not of a representation, but of a module. Formally, a *module* of a Lie algebra $\mathfrak{g}$ (a $\mathfrak{g}$-module for short) is a vector space $V$ equipped with an action of $\mathfrak{g}$ that is linear and respects the Lie bracket. Here, an action of $\mathfrak{g}$ on $V$ means a map from $\mathfrak{g} \times V$ to $V$ denoted by $x \cdot v$, with $x \in \mathfrak{g}$ and $v \in V$, or (rather more lazily) just by $xv$. Linearity means that

$$\begin{aligned} (ax+by)v &= a(xv) + b(yv), \\ x(av+bw) &= a(xv) + b(xw), \end{aligned} \qquad \text{for all } a,b \in \mathbb{C},\ x,y \in \mathfrak{g} \text{ and } v,w \in V, \tag{2.53}$$

and respecting the Lie bracket means that

$$[x,y]v = x(yv) - y(xv), \quad \text{for all } x,y \in \mathfrak{g} \text{ and } v \in V. \tag{2.54}$$

Usually, of course, we would drop the parentheses in these expressions.

Hopefully, it is clear that representations and modules are just two ways of describing the same concept. Simply put, a representation $\pi \colon \mathfrak{g} \longrightarrow \mathfrak{gl}(V)$ defines a $\mathfrak{g}$-module structure on $V$ via

$$xv = \pi(x)v, \quad \text{for all } x \in \mathfrak{g} \text{ and } v \in V, \tag{2.55}$$

and a $\mathfrak{g}$-module $V$ defines a representation $\pi$ by the same formula. The "module side" is just a little shorter to write, which probably explains its popularity.

Just as ideals are the important sub-structure for Lie algebras, it is important to consider the analogous sub-structure for modules. A *submodule* of a $\mathfrak{g}$-module $V$ is a vector subspace $W \subseteq V$ that is preserved by the action of $\mathfrak{g}$:

$$w \in W \quad \Rightarrow \quad xw \in W \quad \text{for all } x \in \mathfrak{g}. \tag{2.56}$$

Physicists often refer to submodules as "invariant subspaces" or subrepresentations. Note that submodules generalise ideals in the sense that $W$ is a submodule of the *adjoint module* of $\mathfrak{g}$, *ie.* the module $\mathfrak{g}$ corresponding to the adjoint representation, if and only if $W$ is an ideal of $\mathfrak{g}$. There is no useful analogue of Lie subalgebras for modules.

**Exercise 28.** Show that if $V$ and $W$ are submodules of some $\mathfrak{g}$-module, then so are $V \cap W$ and $V + W$. ▼

The analogy with ideals is manifest in the notion of a *quotient module*. If $W$ is a submodule of a $\mathfrak{g}$-module $V$, then the vector space quotient $V/W$ is naturally a $\mathfrak{g}$-module with the action

$$x\overline{v} = \overline{xv}, \quad \text{for all } x \in \mathfrak{g} \text{ and } \overline{v} \in V/W. \tag{2.57}$$

We recall that the equivalence class $\overline{v} \in V/W$ is defined by $\overline{v} = \{v' \in V : v - v' \in W\}$. Thus, $\overline{v} = \overline{v'}$ if $v - v' \in W$. The action on $V/W$ given above is therefore well-defined: $\overline{v} = \overline{v'}$ implies that $v - v' \in W$, hence $xv - xv' = x(v - v') \in W$ (as $W$ is a submodule), so $\overline{xv} = \overline{xv'}$.

A $\mathfrak{g}$-module $V$ is said to be *irreducible* if its only submodules are 0 and $V$ itself. A module that is not irreducible is, of course, *reducible*. Irreducibility is the analogue of simplicity for Lie algebras. (Indeed, many mathematicians use "simple" instead of "irreducible" for modules.) However, while we excluded the one-dimensional Lie algebra from being simple, we will not exclude a one-dimensional module from being irreducible.

**Example 23.** Since $\mathfrak{g}$ acts as 0 on any trivial $\mathfrak{g}$-module $V$, it follows that every subspace of $V$ is a submodule. A trivial $\mathfrak{g}$-module $V$ is therefore only simple if $\dim V = 1$. In general, a trivial $\mathfrak{g}$-module $V$ decomposes as the direct sum of $\dim V$ copies of the one-dimensional trivial $\mathfrak{g}$-module. ▲

**Example 24.** The adjoint module of a **simple** Lie algebra has no non-zero proper submodules, hence is irreducible, because a simple Lie algebra has no non-zero proper ideals. Indeed, a Lie algebra is simple if and only if its adjoint module is irreducible. ▲

The analogue of semisimplicity for Lie algebras is complete reducibility: a $\mathfrak{g}$-module is said to be *completely reducible* if it is a (finite) direct sum of simple $\mathfrak{g}$-modules. (Again, mathematicians often use "semisimple" instead of "completely reducible".) We explain the notion of a direct sum of modules, along with some other frequently encountered gadgets, in the following exercise.

**Exercise 29.** Let $\mathfrak{g}$ be a Lie algebra and let $V$ and $W$ be $\mathfrak{g}$-modules.

**(a)** Show that the vector space direct sum $V \oplus W$ becomes a $\mathfrak{g}$-module under the action

$$x(v \oplus w) = (xv) \oplus (xw), \quad \text{for all } x \in \mathfrak{g},\ v \in V \text{ and } w \in W. \tag{2.58}$$

The $\mathfrak{g}$-module $V \oplus W$ is called the *direct sum* of $V$ and $W$.

**(b)** Show that the vector space tensor product $V \otimes W$ does **not** become a $\mathfrak{g}$-module under

$$x(v \otimes w) = (xv) \otimes (xw), \quad \text{for all } x \in \mathfrak{g},\ v \in V \text{ and } w \in W. \tag{2.59}$$

**(c)** Show that

$$x(v \otimes w) = (xv) \otimes w + v \otimes (xw), \quad \text{for all } x \in \mathfrak{g},\ v \in V \text{ and } w \in W, \tag{2.60}$$

extended by linearity to all $V \otimes W$, does result in a $\mathfrak{g}$-module: the *tensor product* of $V$ and $W$.

**(d)** Show that the vector space dual $V^*$, *ie.* the space of linear functionals $f$ from $V$ to the ground field, becomes a $\mathfrak{g}$-module under the action

$$(xf)(v) = -f(xv), \quad \text{for all } x \in \mathfrak{g},\ f \in V^* \text{ and } v \in V. \tag{2.61}$$

What goes wrong if we omit the minus sign in this action? ▼

**Proposition 2.2.** *A finite-dimensional $\mathfrak{g}$-module $V$ is completely reducible if and only if every submodule $W \subseteq V$ has a complement, ie. there exists a submodule $X \subseteq V$ such that $V = W \oplus X$.*

*Proof.* We prove that complete reducibility implies that every submodule has a complement, the converse being a simple induction argument on the dimension. So suppose that $V = V_1 \oplus \cdots \oplus V_n$, where each $V_i$ is irreducible. Given $W \subseteq V$, consider the collection $\mathcal{S}$ of submodules whose intersection with $W$ is 0. Take $X$ to be maximal in $\mathcal{S}$, meaning that $W \cap X = 0$ and $X \subset Y$ implies that $W \cap Y \neq 0$. Such an $X$ must exist because the dimensions of the submodules in $\mathcal{S}$ are bounded above by $\dim V$.

We want to prove that $X$ is the desired complement of $W$. Since $W \cap X = 0$, this only requires that $W + X = V$. So suppose that $W + X \neq V$. Then, we may choose $v \notin W + X$. Writing $v$ as $v_1 + \cdots + v_n$, where $v_i \in V_i$, we conclude that some $v_j \notin W + X$. This makes $V_j \cap (W + X)$ a proper submodule of $V_j$, hence it is 0 by the irreducibility of $V_j$.

Given $V_j \cap (W + X) = 0$, we consider $w \in W \cap (X + V_j)$. We may thus write $w \in W$ as $x + v'$, where $x \in X$ and $v' \in V_j$. Then, $v' = w - x \in W + X$ and so $v' = 0$ follows from $V_j \cap (W + X) = 0$. This means that $w = x$, which forces $w = 0$, because $W \cap X = 0$, and so we conclude that $W \cap (X + V_j) = 0$. As $v_j$ is in $V_j$, but not in $W + X \supseteq X$, we have $X \subset X + V_j$ and so this contradicts $X$ being maximal among the submodules of $\mathcal{S}$. We therefore have $V = W \oplus X$, as desired. ■

**Corollary 2.3.** *A submodule $W$ of a finite-dimensional completely reducible $\mathfrak{g}$-module $V$ is also completely reducible.*

*Proof.* Let $X$ be a submodule of $W$. Then, it is also a submodule of $V$, hence there is a submodule $Y \subseteq V$ such that $V = X \oplus Y$, by Proposition 2.2. We consider $W \cap Y$ as a potential complement of $X$ in $W$. Obviously, $X \cap (W \cap Y) = W \cap (X \cap Y) = 0$. Moreover, $X \subseteq W$ implies that $X + (Y \cap W) = (X + Y) \cap W = V \cap W = W$, by the modular law for submodules (Exercise 30 below). This shows that every submodule of $W$ has a complement, hence $W$ is completely reducible, again by Proposition 2.2. ■

**Exercise 30.** Prove the *modular law* for submodules $A$, $B$ and $C$ of some module $V$:

$$A \subseteq C \quad \Rightarrow \quad A + (B \cap C) = (A + B) \cap C. \tag{2.62}$$

▼

**Exercise 31.** Prove the following version of the Krull-Schmidt theorem (*cf.* Proposition 2.1). If a completely reducible $\mathfrak{g}$-module $V$ has two decompositions as a direct sum of irreducible submodules,

$$V \simeq U_1 \oplus \cdots \oplus U_m \quad \text{and} \quad V \simeq W_1 \oplus \cdots \oplus W_n, \tag{2.63}$$

then $m = n$ and there is a permutation $\sigma$ of $\{1, \ldots, n\}$ such that $U_i \simeq W_{\sigma(i)}$ for all $i = 1, \ldots, n$. ▼

We remark that all of these complete reducibility results also have generalisations to infinite-dimensional modules. However, the proofs are somewhat more subtle because one needs certain maximal and/or minimal submodules to exist. The existence of these submodules in general is not at all obvious and is, in fact, equivalent to the infamous axiom of choice (see Curtis and Reiner, for example).

Direct sums of modules are quite common in the wild and are easily understood. However, it is not normally the case that all modules are completely reducible (see Example 28 below). If a module $V$ may be written as the direct sum of two non-zero submodules, $V = V_1 \oplus V_2$, then $V$ is said to be *decomposable*. The opposite notion of course defines an *indecomposable* module. The existence of non-completely reducible modules then implies the existence of modules that are indecomposable, but reducible. *ie.* they possess non-zero proper submodules that are never direct summands.

Finally, we have to discuss when two representations/modules are the same, *ie.* are isomorphic. A *homomorphism* of modules of a Lie algebra $\mathfrak{g}$ (a $\mathfrak{g}$-*module homomorphism* for short) is a linear map $\phi$ between two $\mathfrak{g}$-modules $V$ and $W$ that respects the two actions of $\mathfrak{g}$. Precisely, this means that if $\cdot$ and $\circ$ denote the actions of $\mathfrak{g}$ on $V$ and $W$, respectively, then $\phi\colon V \to W$ must satisfy

$$\phi(x \cdot v) = x \circ \phi(v), \quad \text{for all } x \in \mathfrak{g} \text{ and } v \in V. \tag{2.64}$$

Of course, mathematicians are a lazy bunch, so this is usually written as

$$\phi(xv) = x\phi(v), \quad \text{for all } x \in \mathfrak{g} \text{ and } v \in V, \tag{2.65}$$

leaving the reader to remember that there are secretly two actions at play here. As with Lie algebra homomorphisms (Section 2.3), a homomorphism from a $\mathfrak{g}$-module to itself (*ie.* $V = W$) is an *endomorphism*, a bijective homomorphism is an *isomorphism* and a bijective endomorphism is an automorphism.

**Exercise 32.** Suppose that $\phi\colon V \to W$ is a $\mathfrak{g}$-module homomorphism. Show that:

**(a)** $\ker \phi$ is a submodule of $V$.
**(b)** $\operatorname{im} \phi$ is a submodule of $W$.
**(c)** If $W'$ is a submodule of $W$, then the preimage

$$\phi^{-1}(W') = \{v \in V : \phi(v) \in W\} \tag{2.66}$$

is a submodule of $V$. ▼

**Exercise 33.**

**(a)** Show that if $W$ is a submodule of a $\mathfrak{g}$-module $V$, then the canonical homomorphism $\gamma\colon V \to V/W$ defined by $\gamma(v) = \overline{v}$ is a surjective $\mathfrak{g}$-module homomorphism.
**(b)** A proper submodule $W$ of $V$ is said to be *maximal* if every submodule chain of the form $W \subseteq X \subset V$ necessarily has $W = X$. Prove that $W \subset V$ is maximal if and only if $V/W$ is irreducible. ▼

In abstract algebra, there are always three isomorphism theorems that get used again and again. The first isomorphism theorem for modules is as follows (though we state it as a proposition rather than a theorem). You may wish to compare with the first isomorphism theorem for Lie algebras which you proved in Exercise 17.

**Proposition 2.4.** *If $\phi\colon V \to W$ is a $\mathfrak{g}$-module homomorphism, then*

$$\frac{V}{\ker \phi} \simeq \operatorname{im} \phi, \quad \textit{as } \mathfrak{g}\textit{-modules.} \tag{2.67}$$

*Proof.* Consider the map $\overline{\phi}\colon V/\ker\phi \to \operatorname{im}\phi$ given by $\overline{\phi}(\overline{v}) = \phi(v)$, for all $v \in V$ and so for all $\overline{v} \in V/\ker\phi$. It is well-defined because $\overline{v} = \overline{w}$ implies that $v - w \in \ker\phi$, hence $\phi(v) = \phi(w)$. It is moreover injective, as $0 = \overline{\phi}(\overline{v}) = \phi(v)$ implies that $v \in \ker\phi$, hence $\overline{v} = 0$. It is obviously surjective. Finally, $\overline{\phi}$ is a $\mathfrak{g}$-module homomorphism as $\phi$ is:

$$x\overline{\phi}(\overline{v}) = x\phi(v) = \phi(xv) = \overline{\phi}(\overline{xv}) = \overline{\phi}(x\overline{v}), \quad \text{for all } x \in \mathfrak{g} \text{ and } v \in V. \tag{2.68}$$ ■

**Exercise 34.** Prove the second and third isomorphism theorems for modules:

**2**. If $W$ and $X$ are submodules of a $\mathfrak{g}$-module $V$, then

$$\frac{W}{W \cap X} \simeq \frac{W + X}{X}, \quad \text{as } \mathfrak{g}\text{-modules.} \tag{2.69}$$

**3**. If $W$ is a submodule of $V$ and $X$ is a submodule of $W$, then

$$\frac{V/X}{W/X} \simeq \frac{V}{W}, \quad \text{as } \mathfrak{g}\text{-modules.} \tag{2.70}$$ ▼

Let's translate the definition of a $\mathfrak{g}$-module homomorphism back into the language of representations. If $\pi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ and $\rho\colon \mathfrak{g} \to \mathfrak{gl}(W)$ denote the representations on $V$ and $W$, respectively, then the condition for $\phi$ to be a $\mathfrak{g}$-module homomorphism becomes

$$\phi(\pi(x)v) = \rho(x)\phi(v), \quad \text{for all } x \in \mathfrak{g} \text{ and } v \in V. \tag{2.71}$$

This may be summarised as follows (with $\circ$ now denoting composition of maps):

$$\phi \circ \pi(x) = \rho(x) \circ \phi, \quad \text{for all } x \in \mathfrak{g}. \tag{2.72}$$

$$\begin{array}{ccc} V & \xrightarrow{\ \phi\ } & W \\ {\scriptstyle \pi(x)}\big\downarrow & & \big\downarrow{\scriptstyle \rho(x)} \\ V & \xrightarrow[\ \phi\ ]{} & W \end{array}$$

Because of this (equivalent) property, $\mathfrak{g}$-module homomorphisms are often called *intertwiners* when speaking of representations of $\mathfrak{g}$.

**Example 25.** Recall that changing the basis of a vector space $V$ is implemented by an invertible change-of-basis matrix $S$. The representing matrix $A$ of a given linear transformation from $V$ to itself then changes by conjugation: $A \mapsto SAS^{-1}$.

If $\pi$ is a representation of $\mathfrak{g}$ on $V$, then we can get a new representation $\rho$ on $V$ by "changing basis" as above:

$$\rho(x) = S\pi(x)S^{-1}, \quad \text{for all } x \in \mathfrak{g}. \tag{2.73}$$

However, $\rho$ is isomorphic to $\pi$ because $S$ is an invertible intertwiner. In other words, changing bases can't change the isomorphism class of the representation (phew!). ▲

**Example 26.** Let $\mathfrak{g}$ be the one-dimensional abelian Lie algebra and let $x \in \mathfrak{g}$ be non-zero. Then, a representation $\pi$ of $\mathfrak{g}$ on $\mathbb{C}^n$ is completely determined by the $n \times n$ matrix $A = \pi(x)$. As we've seen, conjugating by an invertible matrix $S$ doesn't change the isomorphism class of the representation, so we may assume that $A$ is in Jordan canonical form. Each Jordan block of $A$ is characterised by its eigenvalue $\lambda$ and its size $n$. Moreover, it's not too hard to now show that the representation $\pi$ is isomorphic the direct sum of the representations specified by these Jordan blocks. ▲

If the representation theory of the one-dimensional Lie algebra is governed by Jordan canonical form (so hard), it doesn't bode well for the representation theory of non-abelian Lie algebras. In general, you'd be right to be scared. Luckily, we're only going to study the representations of (semi)simple Lie algebras and these are very well behaved.

# 3. All about $\mathfrak{sl}(2)$

We have already introduced the complex Lie algebra $\mathfrak{sl}(2)$ and shown (Example 17) that it is simple. We have also seen that it has at least two inequivalent real forms: $\mathfrak{su}(2)$ and $\mathfrak{sl}(2;\mathbb{R})$. In this section, we shall study the representation theory of $\mathfrak{sl}(2)$. As we shall see in Section 4 below, this turns out to be the key to understanding the structure of all the simple (and hence semisimple and reductive) complex Lie algebras. Moreover, the representation theory of $\mathfrak{sl}(2)$ is plenty interesting in its own right and we shall briefly discuss how it arises in the classical and quantum mechanical description of angular momentum and quantum spin.

## 3.1. **Irreducible representations of** $\mathfrak{sl}(2)$

As promised, we now turn our attention to the finite-dimensional irreducible representations of $\mathfrak{sl}(2)$. We recall the basis of $\mathfrak{sl}(2)$ consisting of the matrices

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{3.1}$$

with respect to which the Lie bracket is characterised by

$$[h,e] = 2e, \quad [h,f] = -2f \quad \text{and} \quad [e,f] = h. \tag{3.2}$$

Our first port of call is to remember that we are working over $\mathbb{C}$, so the corresponding modules will be finite-dimensional complex vector spaces. (We shall generally prefer to use the language and notation of modules in what follows.)

So choose such an $\mathfrak{sl}(2)$-module. Because it is complex and finite-dimensional, any given endomorphism of the underlying vector space will have an eigenvalue. In particular, this is the case for the endomorphism representing the action of the basis element $h \in \mathfrak{sl}(2)$. Because it is finite-dimensional, there are only finitely many eigenvalues, so there is (at least) one, $\lambda$ say, whose real part is maximal. Choose an eigenvector $v_\lambda$ corresponding to the eigenvalue $\lambda$, so that

$$hv_\lambda = \lambda v_\lambda. \tag{3.3}$$

Then, $ev_\lambda$ is also an eigenvector of $h$ and its eigenvalue is $\lambda + 2$:

$$h(ev_\lambda) = [h,e]v_\lambda + e(hv_\lambda) = 2ev_\lambda + e(\lambda v_\lambda) = (\lambda+2)ev_\lambda. \tag{3.4}$$

However, this contradicts the fact that we chose the real part of $\lambda$ to be maximal among all eigenvalues. The only way out is to conclude that $ev_\lambda$ is not an $h$-eigenvector and this can only occur if it is instead zero:

$$ev_\lambda = 0. \tag{3.5}$$

Having established the action of $h$ and $e$ on $v_\lambda$, we turn to the action of $f$. The same argument as that used for $e$ now shows that, if it is not zero, $fv_\lambda$ is an $h$-eigenvector of eigenvalue $\lambda - 2$. We can repeat this argument with $v_\lambda$ replaced by $fv_\lambda$ to conclude that $f^2 v_\lambda = f(fv_\lambda)$ is, if non-zero, an $h$-eigenvector of eigenvalue $\lambda - 4$. In general, we find that $f^n v_\lambda$ is an $h$-eigenvector of eigenvalue $\lambda - 2n$, for any $n \in \mathbb{Z}_{\geqslant 0}$, again provided that it is non-zero.

Recall that eigenvectors corresponding to different eigenvalues are automatically linearly independent. If no $f^n v_\lambda$ were zero, then we would infinitely many linearly independent eigenvectors in our module, contradicting the fact that it is assumed to be finite-dimensional. We therefore conclude that there must exist $n \in \mathbb{Z}_{>0}$ such that

$$f^n v_\lambda = 0. \tag{3.6}$$

Crucially, we can even compute the minimal such $n$ using the following exercise.

**Exercise 35.** Show inductively that the endomorphisms representing $e$ and $f$ satisfy

$$[e, f^n] = n f^{n-1} \big(h - (n-1)\mathbb{1}\big), \qquad \text{hence that} \quad e f^n v_\lambda = n(\lambda + 1 - n) f^{n-1} v_\lambda, \tag{3.7}$$

for all $n \in \mathbb{Z}_{\geqslant 0}$. ▼

For the minimal $n \in \mathbb{Z}_{>0}$, we have $f^{n-1} v_\lambda \neq 0$ but

$$0 = e f^n v_\lambda = n(\lambda + 1 - n) f^{n-1} v_\lambda \quad \Rightarrow \quad n = \lambda + 1. \tag{3.8}$$

Rather amazingly, we learn from this analysis that $\lambda = n - 1 \in \mathbb{Z}_{\geqslant 0}$: the $h$-eigenvalue with maximal real part, for a finite-dimensional (complex) $\mathfrak{sl}(2)$-module, is necessarily a non-negative integer!

We can even go a step further. Along with our construction, Exercise 35 shows that the vector subspace spanned by the $f^n v_\lambda$, with $n = 0, 1, \ldots, \lambda$, is closed under the action of $e$, $h$ and $f$: it is a submodule of our $\mathfrak{sl}(2)$-module. If we further insist that our $\mathfrak{sl}(2)$-module is irreducible, then it must coincide with the span of the $f^n v_\lambda$. We learn that the maximal $h$-eigenvalue of a finite-dimensional irreducible (complex) $\mathfrak{sl}(2)$-module is a non-negative integer $\lambda$ and that the dimension of this module is $\lambda + 1$.

**Example 27.** The irreducible trivial representation of $\mathfrak{sl}(2)$ has maximal $h$-eigenvalue 0 and dimension 1. The defining representation, introduced in Example 22, has $h$ being represented by a diagonal matrix with entries 1 and $-1$. The maximal $h$-eigenvalue is thus 1 and the dimension of the representation is 2. Similarly, the adjoint representation and that appearing in Exercise 27 have maximal $h$-eigenvalues 2 and 3 and dimensions 3 and 4, respectively. ▲

A pure mathematician would no doubt interrupt the cheering now to (correctly) object that the above analysis assumes from the outset that we actually have, in our possession, a finite-dimensional irreducible $\mathfrak{sl}(2)$-module. It therefore still remains to verify that these modules actually exist. There are smart ways to do this (and we'll see one later), but for now we can fall back on brute-force methods.

**Exercise 36.** For $\lambda \in \mathbb{Z}_{\geqslant 0}$, consider the $(\lambda + 1)$-dimensional vector space spanned by vectors $w_n = f^n v_\lambda$, $n = 0, 1, \ldots, \lambda$. Compute $(\lambda + 1) \times (\lambda + 1)$-matrices $[e]$, $[h]$ and $[f]$ that represent the action of $e$, $h$ and $f$, with respect to the (ordered) basis $\{w_n\}$, and prove that they satisfy the commutation relations of $\mathfrak{sl}(2)$. ▼

This exercise demonstrates that a $(\lambda+1)$-dimensional irreducible $\mathfrak{sl}(2)$-module $\mathcal{L}_\lambda$ exists for each $\lambda \in \mathbb{Z}_{\geqslant 0}$. The preceding argument shows that we can find a basis of any $(\lambda + 1)$-dimensional irreducible $\mathfrak{sl}(2)$-module so that it is clearly $\mathcal{L}_\lambda$. In other words, the $\mathcal{L}_\lambda$, $\lambda \in \mathbb{Z}_{\geqslant 0}$, are all the finite-dimensional irreducible $\mathfrak{sl}(2)$-modules, up to isomorphism. We state this as a theorem for later reference.

**Theorem 3.1.** *There exists an irreducible* $\mathfrak{sl}(2)$*-module* $\mathcal{L}_\lambda$ *of dimension* $\lambda + 1$, *for each* $\lambda \in \mathbb{Z}_{\geqslant 0}$. *Moreover, these are all the finite-dimensional irreducible* $\mathfrak{sl}(2)$*-modules, up to isomorphism.*

Much easier than Jordan canonical form, right?

**Exercise 37.** Prove that the matrix representing the action of $h$ on any finite-dimensional $\mathfrak{sl}(2)$-module has zero trace. Use this to quickly prove that $\dim \mathcal{L}_\lambda = \lambda + 1$, for all $\lambda \in \mathbb{Z}_{\geqslant 0}$. ▼

**Exercise 38.** In this exercise, we explore whether the finite-dimensional irreducible $\mathfrak{sl}(2)$-module $\mathcal{L}_\lambda = \operatorname{span}\{f^n v_\lambda : n = 0, 1, \ldots, \lambda\}$ admits an "invariant" inner product. Such modules are said to be *unitarisable* or, once an invariant inner product has been fixed, *unitary*.

**(a)** Consider a hermitian form $\langle \cdot, \cdot \rangle$ on $\mathcal{L}_\lambda$ that satisfies

(3.9) $$\langle v_\lambda, v_\lambda \rangle = 1 \quad \text{and} \quad \langle v, xw \rangle = \langle x^\dagger v, w \rangle, \qquad \text{for all } v, w \in \mathcal{L}_\lambda \text{ and } x \in \mathfrak{sl}(2),$$

where $^\dagger$ is the adjoint defined in Exercise 24a. Show that this uniquely determines the hermitian form by computing the values $\langle f^m v_\lambda, f^n v_\lambda \rangle$.

**(b)** Conclude that this hermitian form is an inner product, *ie.* that the $\mathcal{L}_\lambda$ are unitary $\mathfrak{sl}(2)$-modules with this choice of adjoint.

**(c)** Show that replacing † by the adjoint of Exercise 24d leads to a hermitian form on $\mathcal{L}_1$ with respect to which there are (non-zero) vectors of positive, negative and zero norm. *ie.* $\mathcal{L}_1$ is not unitary with this choice of adjoint. ▼

The reason for the term "unitary" in the previous exercise stems from the fact that a given hermitian form on $\mathfrak{g}$ restricts to a bilinear form on the real form $\mathfrak{g}_{\mathbb{R}}$ (*cf.* Exercise 24) defined by the adjoint $^{\dagger}$. As the real form consists of antihermitian elements (with respect to $^{\dagger}$), exponentiating them as in Exercise 9 will give **unitary** elements of the corresponding real Lie group (again, with respect to $^{\dagger}$).

**Exercise 39.** Show that $\mathfrak{sl}(2)$ also possesses infinite-dimensional modules by verifying that the vector space $\mathbb{C}[z]$ of polynomials becomes an $\mathfrak{sl}(2)$-module when equipped with the following action (extended by linearity to all $\mathfrak{sl}(2)$) on each $p(z) \in \mathbb{C}[z]$:

$$e \cdot p(z) = -zp''(z), \quad h \cdot p(z) = -2zp'(z), \quad f \cdot p(z) = zp(z). \tag{3.10}$$

Show that $h$ has a basis of eigenvectors and determine the corresponding eigenvalues. Explain why $f$ has no eigenvectors. What are the eigenvalues and eigenvectors of $e$? ▼

## 3.2. Finite-dimensional representations of $\mathfrak{sl}(2)$

Suppose that we are trekking through the jungles of Borneo and we happen upon an irreducible $\mathfrak{sl}(2)$-module. We now know what to do: classify it by computing its maximal $h$-eigenvalue (or just by figuring out its dimension). However, what if this feral $\mathfrak{sl}(2)$-module isn't irreducible? What do we do? How bad could it be? To get a glimpse of the type of badness that awaits us in the wild world, we return (briefly) to the representation theory of the one-dimensional Lie algebra.

**Example 28.** Recall from Example 26 that a finite-dimensional module of the one-dimensional Lie algebra $\mathfrak{g}$ is completely determined by the matrix $A$ that represents the action of an arbitrarily chosen non-zero $x \in \mathfrak{g}$. We suppose that the module $V$ is two-dimensional, for simplicity, and that $A$ is diagonalisable. Then, $A$ has two linearly independent eigenvectors with eigenvalues $\lambda$ and $\mu$, respectively. The action of $x$ (*ie.* acting with $A$) then preserves each eigenspace: both eigenspaces, $W_\lambda$ and $W_\mu$ say, are submodules of $V$. The sum of the eigenspaces is clearly all of $V$ and their intersection is 0, hence we have the direct sum decomposition

$$V = W_\lambda \oplus W_\mu. \tag{3.11}$$

As both $W_\lambda$ and $W_\mu$ are one-dimensional, hence irreducible, it follows that $V$ is completely reducible (see Section 2.5).

So this is very nice, but it turns out that bad things happen if we don't have two linearly independent eigenvectors, *ie.* if $A$ is not diagonalisable. Let's take

$$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \tag{3.12}$$

for definiteness. Then, the eigenvector $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ is preserved by the action of $A$ and so spans a submodule $W_\lambda$ of $V$. However, there is no other one-dimensional submodule (it would be a second eigenspace), so we can't write $V$ as the direct sum of $W_\lambda$ and another submodule.

With this $A$, $V$ is reducible, because it has a non-zero proper submodule $W_\lambda$, but indecomposable, because it cannot be written as a direct sum of two non-zero modules. The way to understand the structure of $V$ is then to analyse the *quotient module* $V/W_\lambda$. This is the $\mathfrak{g}$-module of equivalence classes in which $W_\lambda$ is identified with 0. Since $A$ acts on $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ to give $\lambda\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$, the equivalence class of $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ is an eigenvector under the action of $A$ with eigenvalue $\lambda$. We therefore conclude that

$$\frac{V}{W_\lambda} \simeq W_\lambda, \quad \text{but} \quad V \not\simeq W_\lambda \oplus W_\lambda. \tag{3.13}$$

Reducible but indecomposable modules are scary, no? ▲

With this type of badness lurking in the representation theory of the one-dimensional Lie algebra, it is perhaps surprising (and gratifying) to learn that the (finite-dimensional) representation theory of $\mathfrak{sl}(2)$ is free of such defects: all finite-dimensional $\mathfrak{sl}(2)$-modules are completely reducible! *ie.* they are all isomorphic to a direct sum of irreducible submodules.

Proving this is a little tricky, but it does give us the opportunity to introduce a couple of important concepts. First up is the *quadratic Casimir operator* $Q$ of $\mathfrak{sl}(2)$. This is **not** an element of $\mathfrak{sl}(2)$; rather we shall define it as the following linear operator acting on an arbitrary $\mathfrak{sl}(2)$-module:

$$Q = \frac{1}{2}h^2 + ef + fe. \tag{3.14}$$

The great utility of this definition is the fact that the quadratic Casimir commutes with the action of $\mathfrak{sl}(2)$ on any module $V$.

**Exercise 40.** Recalling that the Lie bracket becomes the commutator when acting on modules, verify that $[Q, e] = [Q, h] = [Q, f] = 0$. ▼

It therefore defines an $\mathfrak{sl}(2)$-module endomorphism of $V$:

$$Q(xv) = x(Qv) \qquad \text{for all } x \in \mathfrak{sl}(2) \text{ and } v \in V. \tag{3.15}$$

So what? Well...

**Lemma 3.2.** *Let $V$ be a finite-dimensional $\mathfrak{g}$-module and $Q\colon V \to V$ be a $\mathfrak{g}$-module endomorphism. Then, the generalised eigenspaces $V_q$ of $Q$ are submodules of $V$ and*

$$V = \bigoplus_q V_q. \tag{3.16}$$

*Proof.* Suppose that $v \in V_q$ so that $(Q - q\mathbb{1})^n v = 0$ for some $n \in \mathbb{Z}_{>0}$ ($\mathbb{1}$ is the identity endomorphism on $V$). Then, $xv \in V_q$, for all $x \in \mathfrak{g}$, because

$$(Q - q\mathbb{1})^n xv = x(Q - q\mathbb{1})^n v = x0 = 0. \tag{3.17}$$

The $V_q$ are thus submodules. Linear algebra informs us that their sum is all of $V$ and their pairwise intersections are 0, so we get (3.16). ■

**Lemma 3.3** (Schur's lemma)**.** *Let $Q\colon V \to V$ be a $\mathfrak{g}$-module endomorphism, where $V$ is an irreducible finite-dimensional complex $\mathfrak{g}$-module. Then, $Q$ acts on $V$ as a multiple of the identity endomorphism.*

*Proof.* Since $V$ is finite-dimensional and complex, $Q$ has an eigenvalue $\lambda$. Consider now the $\mathfrak{g}$-module endomorphism $Q - \lambda\mathbb{1}$ on $V$. Its kernel is a submodule of $V$, by Exercise 32, hence it is either $V$ or 0 because $V$ is irreducible. However, the eigenvector corresponding to $\lambda$ belongs to $\ker(Q - \lambda\mathbb{1})$, hence the kernel is not 0. It is therefore $V$ and so we conclude that $Q - \lambda\mathbb{1} = 0$ on $V$. ■

**Exercise 41.** Show that Schur's lemma fails over $\mathbb{R}$ by constructing an irreducible real two-dimensional module $V$ for the real one-dimensional Lie algebra $\mathfrak{u}(1)$ and a $\mathfrak{u}(1)$-module endomorphism of $V$ that isn't a multiple of the identity. ▼

Recall the $(\lambda+1)$-dimensional irreducible $\mathfrak{sl}(2)$-module $\mathcal{L}_\lambda$, $\lambda \in \mathbb{Z}_{\geqslant 0}$, that we constructed in Section 3.1. Schur's lemma says that the quadratic Casimir acts as a multiple of the identity on $\mathcal{L}_\lambda$. It's easy to determine this multiple because we can just act with the quadratic Casimir on any non-zero element in $\mathcal{L}_\lambda$. In particular, the $h$-eigenvector $v_\lambda$ of eigenvalue $\lambda$ is convenient because it is annihilated by $e$:

$$\begin{aligned} Qv_\lambda &= \left(\frac{1}{2}h^2 + ef + fe\right)v_\lambda = \left(\frac{1}{2}h^2 + ef\right)v_\lambda = \left(\frac{1}{2}h^2 + [e,f]\right)v_\lambda \\ &= \left(\frac{1}{2}h^2 + h\right)v_\lambda = \left(\frac{1}{2}\lambda(\lambda+2)\right)v_\lambda. \end{aligned} \tag{3.18}$$

Notice that the multiples $\frac{1}{2}\lambda(\lambda+2)$ are different for each different $\lambda \in \mathbb{Z}_{\geqslant 0}$. The quadratic Casimir therefore distinguishes between all the finite-dimensional irreducible $\mathfrak{sl}(2)$-modules (up to isomorphism of course).

**Lemma 3.4.** *Let $V$ be a finite-dimensional $\mathfrak{sl}(2)$-module on which the quadratic Casimir $Q$ has a single eigenvalue $q$. Then, the eigenvalues of $h$ on $V$ are precisely the elements of the set $S_\lambda = \{\lambda, \lambda-2, \dots, -\lambda+2, -\lambda\}$, where $\lambda \in \mathbb{Z}_{\geqslant 0}$ is uniquely determined by $q = \frac{1}{2}\lambda(\lambda+2)$.*

*Proof.* Since $\dim V < \infty$, $h$ has an eigenvector $v \in V$. Let $\mu$ be the corresponding eigenvalue. For each $j \in \mathbb{Z}_{\geqslant 0}$, the vector $e^j v$ is either an $h$-eigenvector, with eigenvalue $\mu + 2j$, or it is 0. Indeed, $\dim V < \infty$ also implies that there exists some $n \in \mathbb{Z}_{\geqslant 0}$ such that $e^n v \neq 0$ and $e^{n+1} v = 0$. Let $\lambda' = \mu + 2n$ denote the $h$-eigenvalue of $e^n v$. Since $e$ annihilates $e^n v$, (3.18) shows that $e^n v$ is an eigenvector of $Q$ of eigenvalue $\frac{1}{2}\lambda'(\lambda' + 2)$. This eigenvalue is necessarily $q$, by hypothesis.

Now act on $v_{\lambda'} = e^n v$ with powers of $f$. Again, $\dim V < \infty$ means that there exists $m$ such that $f^m v_{\lambda'} \neq 0$ and $f^{m+1} v_{\lambda'} = 0$. The analysis of Section 3.1, particularly Equation (3.8), then shows that $\lambda' = m \in \mathbb{Z}_{\geqslant 0}$. But, $q = \frac{1}{2}\lambda'(\lambda' + 2)$ has a unique non-negative solution: $\lambda' = \lambda$. The $f^j v_\lambda$, $j = 0, 1, \dots, \lambda$, are therefore $h$-eigenvectors whose eigenvalues fill out the set $S_\lambda$.

It remains to show that the arbitrary eigenvalue $\mu$ also lies in $S_\lambda$. Assume then that $\mu \notin S_\lambda$. Since $\mu = \lambda - 2n$, $n \in \mathbb{Z}_{\geqslant 0}$, this may only happen if $\mu < -\lambda$. Once again, $\dim V < \infty$ gives the existence of $\ell \in \mathbb{Z}_{\geqslant 0}$ such that $f^\ell v \neq 0$ but $f^{\ell+1} v = 0$. Let $\nu = \mu - 2\ell$ be the $h$-eigenvalue of $f^\ell v$, so that $-\nu = -\mu + 2\ell > \lambda$. Then, $f^\ell v$ is a $Q$-eigenvector:

$$Qf^\ell v = \left(\frac{1}{2}h^2 + fe\right) f^\ell v = \left(\frac{1}{2}h^2 - h\right) f^\ell v = \frac{1}{2}\nu(\nu - 2) f^\ell v. \tag{3.19}$$

However, its $Q$-eigenvalue therefore satisfies

$$q = \frac{1}{2}(-\nu)(-\nu + 2) > \frac{1}{2}\lambda(\lambda + 2) = q, \tag{3.20}$$

which is the desired contradiction. ■

We're now ready to prove the complete reducibility of all finite-dimensional $\mathfrak{sl}(2)$-modules.

**Theorem 3.5** (Weyl)**.** *Every finite-dimensional (complex) $\mathfrak{sl}(2)$-module is completely reducible.*

*Proof.* Let $V$ be a finite-dimensional $\mathfrak{sl}(2)$-module. By Lemma 3.2, it decomposes as a direct sum of the generalised eigenspaces $V_q$ of the quadratic Casimir $Q$ (here $q$ is the eigenvalue of $Q$). We therefore need only show that the $V_q$ are completely reducible. If $V_q = 0$, then this is trivially the case, so assume that $V_q \neq 0$.

Choose a **maximal** completely reducible submodule $M$ of $V_q$. $M$ is therefore isomorphic to the direct sum of a finite number of copies of $\mathcal{L}_\lambda$, where $\lambda \in \mathbb{Z}_{\geqslant 0}$ is uniquely determined by $q = \frac{1}{2}\lambda(\lambda + 2)$. As $h$ acts diagonalisably on each $\mathcal{L}_\lambda$, it acts diagonalisably on $M$. Our aim is to prove that $M = V_q$. We shall do so by contradiction.

So suppose that $M \neq V_q$, hence that $V_q/M \neq 0$. By Lemma 3.4, the eigenvalues of $h$ on $V_q/M$ are then precisely those in $S_\lambda = \{\lambda, \lambda-2, \dots, -\lambda+2, -\lambda\}$. Since $\dim(V_q/M) < \infty$, there is an $h$-eigenvector $\overline{v} \in V_q/M$ of eigenvalue $\lambda$ (here, $v$ denotes an arbitrary preimage in $V_q$). Since $\overline{(h-\lambda\mathbb{1})v} = (h-\lambda\mathbb{1})\overline{v} = 0$, we must have $(h-\lambda\mathbb{1})v = w$ for some $w \in M$. By projecting onto the generalised eigenspace corresponding to eigenvalue $\lambda$, we may assume that both $v$ and $w$ belong to this eigenspace.

Suppose that $w$ is not zero, *ie.* that $h$ has a non-trivial Jordan block. Because $\lambda + 2$ is not an $h$-eigenvalue on $M$, again by Lemma 3.4, it follows that $ew$ is zero. As $h$ acts diagonalisably on $M$, it follows that the $f^n w$, with $n = 0, 1, \dots, \lambda$, span a submodule of $M$ isomorphic to $\mathcal{L}_\lambda$. In particular, $f^\lambda w \neq 0$. Similarly, we must have $ev = f^{\lambda+1} v = 0$ because $\lambda + 2$ and $-\lambda - 2$ are not $h$-eigenvalues on $V_q$. Using Exercise 35, we now get

$$0 = ef^{\lambda+1}v = [e, f^{\lambda+1}]v = (\lambda+1)f^\lambda(h-\lambda\mathbb{1})v = (\lambda+1)f^\lambda w \neq 0, \tag{3.21}$$

a contradiction.

Our supposition that $w \neq 0$ must be false, hence $w = 0$ and so $v$ is an eigenvector of $h$ with eigenvalue $\lambda$ that is annihilated by $e$. As $\overline{v} \neq 0$, we must have $v \notin M$. Hence the submodule $N \simeq \mathcal{L}_\lambda$ spanned by the $f^n v$, with $n = 0, 1, \dots, \lambda$, has zero intersection with $M$. Therefore, $M \oplus N$ is a completely reducible submodule of $V_q$, contradicting the maximality of $M$. Our supposition that $M \neq V_q$ must be false, hence $M = V_q$ and so $V_q$ is completely reducible as required. ■

To summarise, we have shown that every finite-dimensional $\mathfrak{sl}(2)$-module $V$ is completely reducible, *ie.* isomorphic to a direct sum of irreducible submodules:

$$V \simeq \bigoplus_{\lambda \in \mathbb{Z}_{\geqslant 0}} m_\lambda \mathcal{L}_\lambda. \tag{3.22}$$

Here, the multiplicities $m_\lambda$ give the number of copies of $\mathcal{L}_\lambda$ in this decomposition. Of course, we expect that such decompositions as direct sums are (essentially) unique. This is guaranteed by the Krull-Schmidt theorem for modules (Exercise 31).

**Exercise 42.** Unfortunately, infinite-dimensional $\mathfrak{sl}(2)$-modules can suffer from the badness of not being completely reducible. Verify this by showing that the $\mathfrak{sl}(2)$-module of Exercise 39 is reducible, but not completely reducible. ▼

### 3.3. An application to quantum mechanics

We've sorted out the finite-dimensional representation theory of $\mathfrak{sl}(2)$. Let's now take a moment to appreciate how it applies to quantum mechanics through the intrinsic spin of a particle.

Recall from Example 10 that the position and momentum of a quantum system are modelled by operators $x$ and $p_x$, respectively, that satisfy $[x, p_x] = \mathrm{i}\hbar\mathbb{1}$, where $\hbar \neq 0$ and

$\mathbb{1}$ is central. We called this Lie algebra the Heisenberg algebra and it acts on the Hilbert space (*ie.* a vector space with extra-nice structure) of quantum states. It follows that the quantum state space is a module for the Heisenberg algebra. We shall assume that $\mathbb{1}$ acts on the quantum state space as the identity operator.

**Exercise 43.** Use traces to show that a finite-dimensional representation of the $\hbar \neq 0$ Heisenberg algebra may not have $\mathbb{1}$ represented by the identity operator. It follows that the quantum state space is necessarily infinite-dimensional. ▼

In three-dimensional space, we need similar operators $y$, $z$, $p_y$ and $p_z$, so we take the direct sum of three copies of the Heisenberg algebra. As the three central operators $\mathbb{1}_x$, $\mathbb{1}_y$ and $\mathbb{1}_z$ are all assumed to act as the identity, we may identify them if we wish (this amounts to quotienting the direct sum by the ideal spanned by $\mathbb{1}_x - \mathbb{1}_y$ and $\mathbb{1}_y - \mathbb{1}_z$). In classical mechanics, angular momentum is defined to be the cross product of the position and momentum vectors, hence the quantised angular momentum operators are given by

$$J_x = yp_z - zp_y, \quad J_y = zp_x - xp_z, \quad J_z = xp_y - yp_x. \tag{3.23}$$

These operators span a representation of $\mathfrak{sl}(2)$ on the quantum state space.

**Exercise 44.** Show that

$$[J_x, J_y] = \mathrm{i}\hbar J_z, \quad [J_y, J_z] = \mathrm{i}\hbar J_x \quad \text{and} \quad [J_z, J_x] = \mathrm{i}\hbar J_y. \tag{3.24}$$

Consequently, verify that

$$e = \frac{1}{\hbar}(J_x + \mathrm{i}J_y), \quad h = \frac{2}{\hbar}J_z \quad \text{and} \quad f = \frac{1}{\hbar}(J_x - \mathrm{i}J_y) \tag{3.25}$$

satisfy the commutation relations of $\mathfrak{sl}(2)$. ▼

As far as quantum spin is concerned, elementary particles are modelled by **irreducible** $\mathfrak{sl}(2)$-modules. There is a subtlety to mention here. Exercise 43 shows that the quantum state space is infinite-dimensional, so it is possible that the quantum state space is not completely reducible (*cf.* Exercise 42) or that it decomposes as a direct sum of irreducible $\mathfrak{sl}(2)$-modules, some of which are infinite-dimensional. However, it turns out that this is not the case and one can consistently restrict attention to the finite-dimensional irreducibles that we classified in Section 3.1. We shall not prove this here, but interested parties can look up such wonderful things as spherical harmonics and the Peter–Weyl theorem.

Having addressed (ignored) this subtlety, we can get on with things. Each (non-zero) vector (called the *quantum state*) in the module is a possible description of the particle's spin in the $z$-direction (say) is given by the eigenvalue of $J_z$ on its quantum state. Note that it follows from Theorem 3.1 that this eigenvalue belongs to a **discrete** set: they are always $\frac{1}{2}\hbar$ times an integer. We therefore say that the $z$-component of the particle's spin is **quantised**. Of course, there is nothing special about the $z$-direction here.

It is common to describe a particle as having a given spin. This does not mean that the $z$-component is fixed. Rather, the *spin* of the elementary particle is defined to be half of the maximal eigenvalue of $h$. Spin is therefore a non-negative half-integer. The celebrated spin-statistics theorem (which we won't prove) states that bosonic particles (*eg*. photons, gravitons, the Higgs) always have integer spins (1, 2 and 0, respectively) while fermionic particles (*eg*. protons, neutrons, electrons) always have non-integer spins (all spin-$\frac{1}{2}$).

**Exercise 45.** Recall the quadratic Casimir of $\mathfrak{sl}(2)$, defined in (3.14). Show that this has a quantum mechanical interpretation, up to an unimportant proportionality constant, as the *total angular momentum* operator

$$|J|^2 = J_x^2 + J_y^2 + J_z^2. \tag{3.26}$$

Recall that the Casimir commutes with the action of $\mathfrak{sl}(2)$. What does this mean for the quantum states of a given irreducible representation? [*Hint: what does Schur say?*] ▼

It would be remiss not to mention now that physicists rarely, if ever, mention the complex Lie algebra $\mathfrak{sl}(2)$ in the context of quantum spin. Instead, they invariably speak of the real Lie group $\mathsf{SU}(2)$. What gives? Spin is all about rotating, so you'd think that if any Lie group was going to be relevant, it would be the rotation group in three dimensions: $\mathsf{SO}(3)$.

If we only cared about the classical world, then $\mathsf{SO}(3)$ would be the group we're after. But quantum mechanics has to be different and one way in which this difference manifests is in the representation theories of Lie groups, as we shall now explain.

We saw in Exercises 11 and 12 that the real Lie algebras $\mathfrak{so}(3)$ and $\mathfrak{su}(2)$ were isomorphic, but that the corresponding real Lie groups were distinguished by the number of elements in their centres: 1 for $\mathsf{SO}(3)$ and 2 for $\mathsf{SU}(2)$. We also saw, in Exercise 9, that the matrix exponential (in the defining representation) maps each of these Lie algebras onto its corresponding Lie group. Recalling the standard basis $\{e, h, f\}$ of the complexification $\mathfrak{sl}(2)$, we see that (in the defining representation) $\pi i h = \begin{pmatrix} \pi i & 0 \\ 0 & -\pi i \end{pmatrix}$ is antihermitian and traceless, hence belongs to $\mathfrak{su}(2)$. Moreover, its matrix exponential $\mathsf{e}^{\pi i h} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ is clearly the non-trivial central element of $\mathsf{SU}(2)$.

**Exercise 46.** Use the explicit isomorphism you constructed in Exercise 11 to write $\pi i h \in \mathfrak{su}(2)$ in terms of a $3 \times 3$ matrix in the defining representation of $\mathfrak{so}(3)$. Verify that $\mathsf{e}^{\pi i h}$ is the identity matrix in this representation. ▼

This exercise shows that $\mathsf{e}^{\pi i h}$ is central in $\mathsf{SO}(3)$, hence is the group unit. Now comes the important bit: the group unit must always be represented by the identity matrix on any

$\mathsf{SO}(3)$-module. As our irreducible $\mathfrak{sl}(2)$-modules $\mathcal{L}_\lambda$, $\lambda \in \mathbb{Z}_{\geqslant 0}$, obviously define $\mathfrak{so}(3)$-modules (just restrict the representation homomorphism to $\mathfrak{so}(3) \simeq \mathfrak{su}(2) \hookrightarrow \mathfrak{sl}(2;\mathbb{C})$), we can now explicitly test whether they define $\mathsf{SO}(3)$-modules or not. This is easy because we know that $h$ can be diagonalised on $\mathcal{L}_\lambda$ with eigenvalues $\lambda, \lambda-2, \dots, -\lambda+2, -\lambda$. It follows that $\mathsf{e}^{\pi \mathsf{i} h}$ is the identity matrix, if $\lambda$ is even, and is minus the identity, if $\lambda$ is odd. We conclude that the $\mathcal{L}_\lambda$, with $\lambda$ odd (*ie.* non-integer spin), are not representations of $\mathsf{SO}(3)$.

As quantum mechanics needs half-integer spin particles (fermions are pretty common in everyday life after all), $\mathsf{SO}(3)$ is not the correct group to describe quantum-mechanical spin. $\mathsf{SU}(2)$ is a good alternative because it has a non-trivial central element that can be represented by minus the identity. It is possible to mathematically prove that $\mathsf{SU}(2)$ is in fact the correct choice. However, we shall not do so here — interested parties should search for the keywords "covering group" and "central extension".

We shall instead recall that the spin $\sigma$ is given as half the $h$-eigenvalue, hence the central element that we have been studying is $\mathsf{e}^{2\pi \mathsf{i} \sigma}$. The action of this element on a quantum state has the physical interpretation of rotating the corresponding particle by $2\pi$ around the origin, *ie.* back to where it started. If the spin is integer, as for bosons, then the state is invariant under a $2\pi$-rotation. If however the spin is not an integer, as for fermions, then the state picks up a minus sign upon $2\pi$-rotations. It is only invariant under a $4\pi$-rotation. This distinction between quantum states of bosonic and fermionic particles is absolutely fundamental in physics and boils down, as we have seen, to some nifty Lie theory.

Let us now turn to a second nifty application of Lie theory to quantum mechanics. This also involves quantum spins, but now asks what happens if we want to consider a system of two (or more) elementary particles, *eg.* a hydrogen atom (proton + electron) or a deuterium nucleus (proton + neutron). The laws of physics tell us that the quantum spin of the composite systems should be modelled as the tensor product of the representations that model the elementary particles. We recall from Exercise 29 that the tensor product of two $\mathfrak{sl}(2)$-modules is again an $\mathfrak{sl}(2)$-module.

Consider the tensor product of the irreducible $\mathfrak{sl}(2)$-module $\mathcal{L}_1$ with itself. Since spin is half the maximal $h$-eigenvalue, this describes a composite system of two spin-$\frac{1}{2}$ fermions (*eg.* a hydrogen atom or a deuterium nucleus). Recall that $\dim \mathcal{L}_1 = 2$, hence $\dim(\mathcal{L}_1 \otimes \mathcal{L}_1) = 4$. As a finite-dimensional $\mathfrak{sl}(2)$-module, $\mathcal{L}_1 \otimes \mathcal{L}_1$ therefore decomposes as a direct sum of irreducible $\mathfrak{sl}(2)$-modules, by Weyl's theorem. There are therefore five possibilities for this decomposition:

$$\mathcal{L}_3, \quad \mathcal{L}_0 \oplus \mathcal{L}_2, \quad \mathcal{L}_1 \oplus \mathcal{L}_1, \quad \mathcal{L}_0 \oplus \mathcal{L}_0 \oplus \mathcal{L}_1 \quad \text{and} \quad \mathcal{L}_0 \oplus \mathcal{L}_0 \oplus \mathcal{L}_0 \oplus \mathcal{L}_0. \tag{3.27}$$

How can we tell which it is? The easiest way is to let $v_1$ denote the eigenvector of maximal $h$-eigenvalue (1) in $\mathcal{L}_1$ and compute the $h$-eigenvalue of $v_1 \otimes v_1 \in \mathcal{L}_1 \otimes \mathcal{L}_1$:

$$h(v_1 \otimes v_1) = (hv_1) \otimes v_1 + v_1 \otimes (hv_1) = v_1 \otimes v_1 + v_1 \otimes v_1 = 2(v_1 \otimes v_1). \tag{3.28}$$

As the only one of the five possibilities above to contain an $h$-eigenvector of eigenvalue 2 is the second, it follows that

$$\mathcal{L}_1 \otimes \mathcal{L}_1 \simeq \mathcal{L}_0 \oplus \mathcal{L}_2. \tag{3.29}$$

In physics language, this means that a composite system of two spin-$\frac{1}{2}$ fermions is equivalent to a linear superposition of a spin-0 and a spin-1 system. In particular, two fermions behave collectively like a boson.

This calculation is easy to generalise, though you should be careful with your arguments!

**Exercise 47.**

- **(a)** Let $\Gamma_V$ denote the set of $h$-eigenvalues of the finite-dimensional $\mathfrak{sl}(2)$-module $V$. Show that

$$\Gamma_{V \oplus W} = \Gamma_V \cup \Gamma_W \quad \text{and} \quad \Gamma_{V \otimes W} = \Gamma_V + \Gamma_W. \tag{3.30}$$

- **(b)** Use this to (carefully!) decompose the tensor product $\mathcal{L}_1 \otimes \mathcal{L}_\mu$, for any $\mu \in \mathbb{Z}_{\geqslant 0}$.
- **(c)** Now use the fact that the tensor product of $\mathfrak{sl}(2)$-modules is commutative and associative, along with the fact that it distributes over direct sums, to deduce the well-known tensor product formula for irreducible finite-dimensional $\mathfrak{sl}(2)$-modules:

$$\mathcal{L}_\lambda \otimes \mathcal{L}_\mu \simeq \mathcal{L}_{|\lambda-\mu|} \oplus \mathcal{L}_{|\lambda-\mu|+2} \oplus \cdots \oplus \mathcal{L}_{\lambda+\mu-2} \oplus \mathcal{L}_{\lambda+\mu}, \quad \text{for all } \lambda, \mu \in \mathbb{Z}_{\geqslant 0}. \tag{3.31}$$ ▼

(3.31) is stated (without proof!) in pretty much every single quantum mechanics text in terms of the spins $j_1 = \frac{1}{2}\lambda$ and $j_2 = \frac{1}{2}\mu$. More precisely, they state that the quantum spin (or angular momentum) $j$ of a state in the composite system must satisfy

$$|j_1 - j_2| \leqslant j \leqslant j_1 + j_2 \tag{3.32}$$

(and $j$ must be an integer or half-integer according as to whether $j_1 + j_2$ is an integer or half-integer). This is often referred to as "addition of angular momenta" or the "Clebsch–Gordan rule for quantum spin" (or something like that).

# 4. Semisimple Lie algebras

We now return to the world of finite-dimensional semisimple complex Lie algebras, with the aim of developing tools to understand them in considerable detail. These include the Killing form, as introduced by Cartan, and the Cartan matrix, which was introduced by Killing. We shall also need to study Cartan subalgebras which, thankfully, were actually introduced by Cartan. This leads to the notions of roots and root systems whose properties ultimately lead to a means of classifying the simple complex Lie algebras. In this section, the ground field is always assumed to be $\mathbb{C}$ unless specified to the contrary.

## 4.1. **The Killing form**

Recall that in (2.37), we defined (a special example of) the linear map $\mathrm{ad}(x)\colon \mathfrak{g} \to \mathfrak{g}$ by

$$\mathrm{ad}(x)\, y = [x, y], \quad \text{for all } y \in \mathfrak{g}. \tag{4.1}$$

We can use this to define a surprisingly useful bilinear form on the Lie algebra $\mathfrak{g}$, the *Killing form*, by

$$\kappa(x, y) = \mathrm{tr}\big[\mathrm{ad}(x)\,\mathrm{ad}(y)\big], \quad \text{for all } x, y \in \mathfrak{g}. \tag{4.2}$$

Note that we are allowed to take the product of the linear maps $\mathrm{ad}(x)$ and $\mathrm{ad}(y)$.

**Exercise 48.** Show that the Killing form is symmetric and invariant:

$$\kappa(x, y) = \kappa(y, x) \quad \text{and} \quad \kappa([x, y], z) = \kappa(x, [y, z]), \qquad \text{for all } x, y, z \in \mathfrak{g}. \tag{4.3}$$ ▼

This latter condition may be written in the form $\kappa(-\,\mathrm{ad}(y)\, x, z) = \kappa(x, \mathrm{ad}(y)\, z)$ which indicates that $\mathrm{ad}(y)$ is antihermitian with respect to the Killing form. Recall that being antihermitian is the Lie-algebraic analogue of being unitary at the group level. This is the reason for the name "invariance": the group action preserves the form.

**Exercise 49.** Given a basis $\{t_i\}$ of a Lie algebra $\mathfrak{g}$, we define the *structure constants* $c_{ijk}$ of $\mathfrak{g}$ (with respect to this basis) by

$$[t_i, t_j] = \sum_k c_{ijk} t_k. \tag{4.4}$$

Prove that if the $t_i$ are orthonormal with respect to the Killing form, *ie.* $\kappa(t_i, t_j) = \delta_{ij}$, then the structure constants are totally antisymmetric:

$$c_{ijk} = c_{jki} = c_{kij} = -c_{jik} = -c_{ikj} = -c_{kji}. \tag{4.5}$$ ▼

**Example 29.** If $\mathfrak{g}$ is abelian, then $\mathrm{ad}(x)\, y = [x, y] = 0$ for all $y \in \mathfrak{g}$. Thus, $\mathrm{ad}(x) = 0$ for all $x \in \mathfrak{g}$ and the Killing form $\kappa$ is identically zero. ▲

**Example 30.** A basis of $\mathfrak{t}(2)$ is given by the matrices

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{4.6}$$

and the Lie bracket in this basis is characterised by

$$[h, e] = 2e, \quad [h, \mathbb{1}] = 0 \quad \text{and} \quad [e, \mathbb{1}] = 0. \tag{4.7}$$

With respect to this basis, we have matrix representatives

$$\mathrm{ad}(\mathbb{1}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathrm{ad}(h) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{and} \quad \mathrm{ad}(e) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -2 & 0 \end{pmatrix}. \tag{4.8}$$

The matrix representing the Killing form is thus

$$\kappa = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.9}$$

▲

**Example 31.** Recall the basis $\{e, h, f\}$ of $\mathfrak{sl}(2)$ given in (3.1) and the explicit Lie brackets given in (3.2). With respect to this basis, we determine the following matrix representatives:

$$\mathrm{ad}(e) = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathrm{ad}(h) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad \text{and} \quad \mathrm{ad}(f) = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}. \tag{4.10}$$

The matrix representing the Killing form is thus

$$\kappa = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 0 \end{pmatrix}. \tag{4.11}$$

▲

We know that the Killing form is always symmetric and invariant. It turns out that an important question to ask is when it is non-degenerate? Let us define the *kernel* (or *radical*) of a bilinear form $B$ on $\mathfrak{g}$ by

$$\ker B = \{x \in \mathfrak{g} : B(x, y) = 0 \text{ for all } y \in \mathfrak{g}\}. \tag{4.12}$$

Obviously, the centre $\mathfrak{z}(\mathfrak{g})$ of $\mathfrak{g}$ (*cf.* Exercise 15) is in the kernel $\ker \kappa$ of the Killing form. However, the kernel may be strictly larger as Example 30 shows.

**Exercise 50.** Let $B$ be an invariant bilinear form $B$ on $\mathfrak{g}$ (such as the Killing form). Show that:

**(a)** $\ker B$ is an ideal of $\mathfrak{g}$.

**(b)** If $\mathfrak{h}$ is an ideal of $\mathfrak{g}$, then its orthogonal complement with respect to $B$, defined by

$$\mathfrak{h}^\perp = \{x \in \mathfrak{g} : B(x, y) = 0 \text{ for all } y \in \mathfrak{h}\}, \tag{4.13}$$

is also an ideal of $\mathfrak{g}$. ▼

It follows from Exercise 50a that if $\mathfrak{g}$ is simple, then either $\ker \kappa$ is all of $\mathfrak{g}$, meaning that $\kappa$ is identically zero, or $\ker \kappa$ is 0, meaning that $\kappa$ is non-degenerate. In fact, one can show that $\kappa = 0$ is impossible for simple $\mathfrak{g}$, though we shall not do so here. It follows easily from this that semisimple Lie algebras have non-degenerate Killing forms.

In fact, this can be strengthened to the following important statement, known as *Cartan's criterion for semisimplicity*.

**Theorem 4.1.** *A finite-dimensional complex Lie algebra $\mathfrak{g}$ is semisimple if and only if its Killing form is non-degenerate.*

The proof requires a lengthy digression through the theory of solvable Lie algebras and may be found in most standard textbooks.

**Exercise 51.** Suppose that $\mathfrak{g} \simeq \bigoplus_i \mathfrak{g}_i$ is semisimple with each ideal $\mathfrak{g}_i$ simple. Prove that $[\mathfrak{g}_i, \mathfrak{g}_j] = 0$ and $\kappa(\mathfrak{g}_i, \mathfrak{g}_j) = 0$, whenever $i \neq j$. ▼

**Exercise 52.** Prove that the complex Lie algebras $\mathfrak{so}(4)$ and $\mathfrak{sp}(4)$ are semisimple by choosing bases and computing the determinant of the matrix representing the Killing form. You may use a computer to perform all the boring linear algebra as long as you explain the method in detail. ▼

**Example 32.** We compute explicitly the Killing form of $\mathfrak{gl}(n)$. Let $E_{ij}$ denote the elementary $n \times n$ matrix whose entries are all 0 except for the $(i, j)$-th one, which is 1. Then, any $A \in \mathfrak{gl}(n)$ has the expansion

$$A = \sum_{i,j=1}^{n} a_{ij} E_{ij}, \quad \text{for some } a_{ij} \in \mathbb{C}. \tag{4.14}$$

For $A, B \in \mathfrak{gl}(n)$, we compute

$$\begin{aligned} \operatorname{ad}(A)\operatorname{ad}(B)\, E_{rs} &= [A, [B, E_{rs}]] = \sum_{i,j,k,\ell=1}^{n} a_{ij} b_{k\ell} [E_{ij}, [E_{k\ell}, E_{rs}]] \\ &= \sum_{i,j,k,\ell=1}^{n} a_{ij} b_{k\ell} [E_{ij}, \delta_{\ell r} E_{ks} - \delta_{sk} E_{r\ell}] \\ &= \sum_{i,j,k,\ell=1}^{n} a_{ij} b_{k\ell} \big(\delta_{\ell r}\delta_{jk} E_{is} - \delta_{\ell r}\delta_{si} E_{kj} - \delta_{sk}\delta_{jr} E_{i\ell} + \delta_{sk}\delta_{\ell i} E_{rj}\big) \end{aligned} \tag{4.15}$$

$$= \sum_{i,j=1}^{n} a_{ij} b_{jr} E_{is} - \sum_{j,k=1}^{n} a_{sj} b_{kr} E_{kj} - \sum_{i,\ell=1}^{n} a_{ir} b_{s\ell} E_{i\ell} + \sum_{j,\ell=1}^{n} a_{\ell j} b_{s\ell} E_{rj}.$$

The Killing form entry $\kappa(A,B)$ is the trace of $\mathrm{ad}(A)\,\mathrm{ad}(B)$, so we must extract the coefficient of $E_{rs}$:

$$\sum_{j=1}^{n} a_{rj} b_{jr} - a_{ss} b_{rr} - a_{rr} b_{ss} + \sum_{\ell=1}^{n} a_{\ell s} b_{s\ell} = (AB)_{rr} - a_{ss} b_{rr} - a_{rr} b_{ss} + (BA)_{ss}. \tag{4.16}$$

Summing over the basis elements $E_{rs}$, $r,s = 1, \ldots, n$, therefore gives

$$\kappa(A,B) = n\,\mathrm{tr}(AB) - \mathrm{tr}\,A\,\mathrm{tr}\,B - \mathrm{tr}\,A\,\mathrm{tr}\,B + n\,\mathrm{tr}(BA) = 2n\,\mathrm{tr}(AB) - 2\,\mathrm{tr}\,A\,\mathrm{tr}\,B. \tag{4.17}$$ ▲

**Exercise 53.** Recall from Exercise 19 that $\mathfrak{gl}(n) \simeq \mathfrak{sl}(n) \oplus \mathfrak{gl}(1)$.

**(a)** Using (4.17), argue carefully that the Killing form of $\mathfrak{sl}(n)$ is given by

$$\kappa(A,B) = 2n\,\mathrm{tr}(AB). \tag{4.18}$$

If you think that this is obvious, check your understanding by comparing the Killing forms of $\mathfrak{t}(2)$, computed in Example 30, and $\mathfrak{gl}(2)$.

**(b)** Show that the basis of $\mathfrak{sl}(n)$ given by

$$\{H_i = E_{ii} - E_{i+1\,i+1} : 1 \leqslant i \leqslant n-1\} \cup \{E_{ij} : 1 \leqslant i \neq j \leqslant n\} \tag{4.19}$$

block-diagonalises the matrix of the Killing form.

**(c)** Show that $\{E_{ij}, E_{ji}\}$ always forms an invertible $2 \times 2$ block in this diagonalisation.

**(d)** Write down the matrix for the $\{H_i\}$ block and show by induction on $n$ that it is also invertible.

This proves that the Killing form of $\mathfrak{sl}(n)$ is non-degenerate, hence that $\mathfrak{sl}(n)$ is semisimple. By Exercise 19, it now follows that $\mathfrak{gl}(n)$ is reductive. ▼

Recall that the matrices $A$ and $B$ in (4.18) may be regarded as the representatives of the corresponding abstract elements $a$ and $b$ in the defining representation $\pi$ of $\mathfrak{sl}(n)$ (see Section 2.5). In other words, the Killing form is proportional to the *trace form* given by

$$\kappa_\pi(A,B) = \mathrm{tr}\big[\pi(a)\pi(b)\big]. \tag{4.20}$$

It turns out that the trace form, with respect to any (finite-dimensional) representation, of a **simple** Lie algebra is always proportional to the Killing form. The fact that this holds for the defining representation of $\mathfrak{sl}(n)$ should give us hope that this Lie algebra is simple. (Recall that we proved that $\mathfrak{sl}(2)$ was simple in Example 17).

**Exercise 54.** Any bilinear form $B(\cdot,\cdot)$ on a finite-dimensional Lie algebra $\mathfrak{g}$ defines a linear map $\phi_B \colon \mathfrak{g} \to \mathfrak{g}^*$ by

$$\phi_B(x)(y) = B(x,y), \qquad \text{for all } x,y \in \mathfrak{g}. \tag{4.21}$$

**(a)** Why is $\phi_B(x)$ a **linear** functional? Why is $\phi_B$ a **linear** map?

**(b)** Recalling Exercise 29, show that if $B$ is invariant, then $\phi_B$ is a $\mathfrak{g}$-module homomorphism from the adjoint module to its dual.

**(c)** Show that if $B$ is non-degenerate, then $\phi_B$ is an isomorphism.

**(d)** Use Schur's lemma (Lemma 3.3) to conclude that every invariant bilinear form on a **simple** Lie algebra $\mathfrak{g}$ is proportional to the Killing form.

**(e)** Check that $\kappa_\pi(x, y) = \text{tr}\big[\pi(x)\pi(y)\big]$, $x, y \in \mathfrak{g}$, defines an invariant bilinear form on $\mathfrak{g}$ for any representation $\pi$. Give an example showing that this form may be degenerate, even when $\mathfrak{g}$ is simple. ▼

## 4.2. **Cartan subalgebras**

The basis (4.19) of $\mathfrak{sl}(n)$ has a number of nice properties. First, it naturally defines a decomposition of $\mathfrak{sl}(n)$ into the direct sum of three Lie subalgebras (not ideals):

- An abelian subalgebra $\mathfrak{sl}(n)_0$ consisting of the (traceless) diagonal matrices (the $H_i$).
- A subalgebra $\mathfrak{sl}(n)_+$ consisting of the strictly upper-triangular matrices (the $E_{ij}$, $i < j$).
- A subalgebra $\mathfrak{sl}(n)_-$ consisting of the strictly lower-triangular matrices (the $E_{ij}$, $i > j$).

The nice properties include the following:

- $\mathfrak{sl}(n)_0$ is maximal in the sense that it is not properly contained in any other abelian subalgebra of $\mathfrak{sl}(n)$.
- Both $\mathfrak{sl}(n)_+$ and $\mathfrak{sl}(n)_-$ are preserved by the adjoint action of $\mathfrak{sl}(n)_0$ and these actions are diagonalisable.
- $\mathfrak{sl}(n)_+$ and $\mathfrak{sl}(n)_-$ are isomorphic as Lie algebras (take $A \mapsto -A^\mathsf{T}$ say).

This decomposition generalises to many other classes of Lie algebras including the semisimple and reductive ones. It is known as a triangular decomposition:

$$\mathfrak{g} = \mathfrak{g}_+ \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_-. \tag{4.22}$$

The nice properties likewise generalise, though they take somewhat more complicated forms depending on which class of Lie algebras we consider. For semisimple Lie algebras, we have the following:

- $\mathfrak{g}_0$ is a maximal ad-diagonalisable subalgebra in that it consists of elements $x \in \mathfrak{g}$ for which $\text{ad}(x)$ is diagonalisable and it is not properly contained in any other ad-diagonalisable subalgebra.
- Both $\mathfrak{g}_+$ and $\mathfrak{g}_-$ are preserved by the adjoint action of $\mathfrak{g}_0$.
- $\mathfrak{g}_+$ and $\mathfrak{g}_-$ are isomorphic as Lie algebras.

We mention that Humphreys uses the term "toral subalgebra" instead of ad-diagonalisable subalgebra. If $\mathfrak{g}$ is semisimple, then we shall refer to any maximal ad-diagonalisable subalgebra as a *Cartan subalgebra*.

Here are two important facts about Cartan subalgebras of semisimple Lie algebras. We will not prove either of them.

**Proposition 4.2.**

**(a)** *Every non-zero semisimple Lie algebra* $\mathfrak{g}$ *has a non-zero Cartan subalgebra.*
**(b)** *Cartan subalgebras are not unique, but any two Cartan subalgebras of a semisimple Lie algebra* $\mathfrak{g}$ *are related by an automorphism of* $\mathfrak{g}$.

The fact that all Cartan subalgebras $\mathfrak{g}_0 \subset \mathfrak{g}$ are related by automorphisms means that they all have the same dimension. This dimension is called the *rank* of $\mathfrak{g}$:

$$\text{rank}\,\mathfrak{g} = \dim \mathfrak{g}_0. \tag{4.23}$$

It is an invariant of $\mathfrak{g}$, meaning that two semisimple Lie algebras with different ranks cannot be isomorphic.

One important fact (that we can prove) gives a hint as to why ad-diagonalisable subalgebras are generally important.

**Proposition 4.3.** *Every* ad*-diagonalisable subalgebra* $\mathfrak{h}$ *of a Lie algebra* $\mathfrak{g}$ *is abelian.*

*Proof.* Suppose that this were false. As $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, this would mean that there exists $x \in \mathfrak{h}$ such that $\text{ad}(x) \colon \mathfrak{h} \to \mathfrak{h}$ has an eigenvector $y \in \mathfrak{h}$ with a non-zero eigenvalue:

$$[x, y] = \lambda y, \quad \text{for some } \lambda \neq 0. \tag{4.24}$$

It follows that $\text{ad}(y)\, x = -\lambda y \neq 0$. Since $\text{ad}(y)$ is diagonalisable, $x$ may be written as a sum of (linearly independent) eigenvectors $x_i$ of $\text{ad}(y)$ and some of the corresponding eigenvalues $\lambda_i$ must be non-zero:

$$x = \sum_i x_i \quad \Rightarrow \quad 0 \neq \text{ad}(y)\, x = \sum_i \lambda_i x_i, \quad \text{with some } \lambda_i \neq 0. \tag{4.25}$$

But, this would imply that $\text{ad}(y)^2\, x = \sum_i \lambda_i^2 x_i \neq 0$. However,

$$\text{ad}(y)^2\, x = [y, [y, x]] = -[y, [x, y]] - [x, [y, y]] = -[y, \lambda y] = 0, \tag{4.26}$$

so we have a contradiction. ∎

**Corollary 4.4.** *A maximal abelian subalgebra that is also* ad*-diagonalisable is thus a maximal* ad*-diagonalisable subalgebra.*

We illustrate the idea of maximal ad-diagonalisable subalgebras and Cartan subalgebras with some familiar examples.

**Example 33.** We saw in Exercise 53 that $\mathfrak{sl}(n)$ is semisimple. Here, we shall show that the elements $H_i$, $1 \leqslant i \leqslant n-1$, of the basis (4.19) span a Cartan subalgebra $\mathfrak{sl}(n)_0$ of $\mathfrak{sl}(n)$. It will follow that the rank of $\mathfrak{sl}(n)$ is $n-1$.

First note that each $\mathrm{ad}(H_i)$ is diagonalisable. In fact, the basis (4.19) consists of its eigenvectors:

$$[H_i, H_j] = 0, \qquad [H_i, E_{jk}] = \left(\delta_{ij} - \delta_{i+1\,j} - \delta_{ik} + \delta_{i+1\,k}\right) E_{jk}, \quad \text{for } j \neq k. \tag{4.27}$$

Since the $H_i$-eigenvalues corresponding to the $E_{jk}$ never simultaneously vanish, the space $\mathfrak{sl}(n)_0$ spanned by the $H_i$ is maximal abelian. It is therefore maximal ad-diagonalisable, by Corollary 4.4, hence it is a Cartan subalgebra of $\mathfrak{sl}(n)$. ▲

**Exercise 55.** It is important to realise that a maximal abelian subalgebra of a semisimple Lie algebra need not be a Cartan subalgebra. Show that the subspace of $\mathfrak{sl}(2m)$ spanned by the $E_{ij}$ with $i = 1, \dots, m$ and $j = m+1, \dots, 2m$ is a maximal abelian subalgebra, but that it is not a Cartan subalgebra because it contains a non-ad-diagonalisable element. ▼

**Exercise 56.** In Exercise 7, you (should have) showed that $\mathfrak{sp}(2n)$ consists of the $2n \times 2n$ matrices $X$ that satisfy

$$X^{\mathsf{T}} J + J X = 0, \tag{4.28}$$

where $J$ is invertible and antisymmetric. If we take $X$ and $J$ to have the $n \times n$-block forms

$$X = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \quad \text{and} \quad J = \left(\begin{array}{c|c} 0 & \mathbb{1} \\ \hline -\mathbb{1} & 0 \end{array}\right), \tag{4.29}$$

*cf.* (2.1), then the defining relation of $\mathfrak{sp}(2n)$ becomes

$$A^{\mathsf{T}} = -D, \quad B^{\mathsf{T}} = B, \quad C^{\mathsf{T}} = C. \tag{4.30}$$

As with $\mathfrak{sl}(n)$, the abelian subalgebra $\mathfrak{h}$ of diagonal matrices of $\mathfrak{sp}(2n)$ is a good candidate for a maximal ad-diagonalisable subalgebra.

**(a)** Show that

$$\{H_i = E_{ii} - E_{n+i\,n+i} : 1 \leqslant i \leqslant n\} \tag{4.31}$$

is a basis for $\mathfrak{h}$.

**(b)** Show that

$$\begin{aligned} &\{H_i : 1 \leqslant i \leqslant n\} \cup \left\{E_{i\,n+j} + E_{j\,n+i}, E_{n+i\,j} + E_{n+j\,i} : 1 \leqslant i < j \leqslant n\right\} \\ &\quad \cup \left\{E_{ij} - E_{n+j\,n+i} : 1 \leqslant i \neq j \leqslant n\right\} \cup \{E_{i\,n+i}, E_{n+i\,i} : 1 \leqslant i \leqslant n\} \end{aligned} \tag{4.32}$$

is a basis of $\mathfrak{sp}(2n)$ that consists of $\mathrm{ad}(H_i)$-eigenvectors (for all $i$). What are the corresponding eigenvalues?

**(c)** Conclude that $\mathfrak{h}$ is maximal abelian, hence maximal ad-diagonalisable. ▼

We shall not refer to $\mathfrak{h}$ as a Cartan subalgebra of $\mathfrak{sp}(2n)$ because we do not know that the latter is semisimple (yet).

The same tricks work for $\mathfrak{so}(n)$, but with two provisos. The first is that orthogonality at the group level translates into antisymmetry at the algebra level. The only diagonal matrix in the real orthogonal Lie algebras is thus the zero matrix and the same is therefore true for their complexifications $\mathfrak{so}(n)$. This makes it hard to guess a maximal ad-diagonalisable subalgebra. We may surmount this by changing the matrix representing the non-degenerate symmetric bilinear form from the $n \times n$ identity matrix to another invertible symmetric matrix. The leads us to the second proviso: the matrix we choose should have a different form for $n$ even and $n$ odd. This is not a bug, but a feature: the structure of $\mathfrak{so}(n)$ indeed does depend on the parity of $n$.

**Exercise 57.** Accordingly, $\mathfrak{so}(n)$ may be taken to consist of the complex $n \times n$ matrices $X$ satisfying

$$X^{\mathsf{T}} K + KX = 0, \tag{4.33}$$

where $K$ has the block form

$$K = \left(\begin{array}{c|c} 0 & \mathbb{1}_m \\ \hline \mathbb{1}_m & 0 \end{array}\right) \quad (n = 2m) \qquad \text{or} \qquad K = \left(\begin{array}{c|c|c} 0 & \mathbb{1}_m & 0 \\ \hline \mathbb{1}_m & 0 & 0 \\ \hline 0 & 0 & 1 \end{array}\right) \quad (n = 2m+1) \tag{4.34}$$

and $\mathbb{1}_m$ denotes the $m \times m$ identity matrix.

- **(a)** For each parity of $n$, write down an appropriate block form for $X$ and determine the conditions on the blocks imposed by the defining relation (4.33).
- **(b)** Use these conditions to write down a basis $\{H_i\}$ for the abelian subalgebra $\mathfrak{h}$ of diagonal matrices of $\mathfrak{so}(n)$.
- **(c)** Find a basis of $\mathfrak{so}(n)$ that consists of simultaneous eigenvectors of the $H_i$.
- **(d)** By examining the corresponding eigenvalues, conclude that $\mathfrak{h}$ is a maximal ad-diagonalisable subalgebra of $\mathfrak{so}(n)$. ▼

Again, we shall have to wait until we show that the $\mathfrak{so}(n)$ are semisimple before we can refer to $\mathfrak{h}$ as a Cartan subalgebra. Actually, there is one orthogonal Lie algebra that is not semisimple: $\mathfrak{so}(2)$ is one-dimensional, hence abelian, so it cannot be semisimple.

At this point, we have (candidate) Cartan subalgebras for the complex Lie algebras $\mathfrak{sl}(n)$, $\mathfrak{sp}(2n)$ and $\mathfrak{so}(n)$, as well as explicit formulae for the adjoint actions of these subalgebras. This turns out to be the key to understanding the structure theory of (finite-dimensional complex) semisimple Lie algebras and will lead to their classification. Here, we note that the Lie algebras $\mathfrak{sl}(n)$, $\mathfrak{sp}(2n)$ and $\mathfrak{so}(n)$ are referred to as the *classical* Lie algebras, presumably because they are so easy to define that they were studied in classical times.

### 4.3. **Roots**

In the last section, we studied ad-diagonalisable subalgebras $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$, noting in Proposition 4.3 that they were automatically abelian. The reason why we care about ad-diagonalisability is the following well-known result from linear algebra that is too commonly omitted from standard courses.

**Proposition 4.5.** *If two endomorphisms of a vector space commute and are are diagonalisable, then they may be simultaneously diagonalised. That is, there exists a basis such that the representing matrices of the endomorphisms are both diagonal.*

*Proof.* Let $A$ and $B$ be the endomorphisms and let $V$ be the vector space. As $A$ is diagonalisable, $V$ decomposes into the direct sum of the eigenspaces $V_\lambda$ of $A$ (here, $\lambda \in \mathbb{C}$ is the corresponding eigenvalue). Since $A$ and $B$ commute, each $v \in V_\lambda$ satisfies

$$A(Bv) = B(Av) = B(\lambda v) = \lambda Bv, \qquad \text{hence} \qquad Bv \in V_\lambda. \tag{4.35}$$

Thus $B$ defines an endomorphism $B_\lambda$ of $V_\lambda$. If $B_\lambda$ were non-diagonalisable, then it would have a generalised eigenvector in $V_\lambda$ and this would define a generalised eigenvector for $B$ in $V$, contradicting the fact that $B$ is diagonalisable. Thus, $B_\lambda$ is also diagonalisable, so $V_\lambda$ decomposes into the direct sum of the eigenspaces of $B_\lambda$. Taking the union, over all $\lambda$, of the corresponding eigenbases therefore results in a basis of $V$ on which both $A$ and $B$ act diagonally, as required. ∎

The point is therefore that the endomorphisms $\mathrm{ad}(H)$, $H \in \mathfrak{h}$, may be simultaneously diagonalised when acting on the vector space $\mathfrak{g}$. Taking $\mathfrak{h}$ to be maximal among ad-diagonalisable subalgebras just means that the eigenvalues of the $\mathrm{ad}(H)$ carry the maximal amount of information possible about $\mathfrak{g}$.

Because of this simultaneous diagonalisation, we have to generalise the notion of an eigenvalue of a single endomorphism. Let $e \in \mathfrak{g}$ denote a simultaneous eigenvector of the $\mathrm{ad}(H)$, $H \in \mathfrak{h}$, where $\mathfrak{h}$ is a maximal ad-diagonalisable subalgebra of $\mathfrak{g}$. This means that

$$[H, e] = \mathrm{ad}(H)e = \alpha(H)e, \tag{4.36}$$

where the eigenvalue $\alpha(H)$ is given by some function $\alpha \colon \mathfrak{h} \to \mathbb{C}$. Because Lie brackets are bilinear, it follows that the function $\alpha$ depends linearly on $H \in \mathfrak{h}$. In other words, it is a linear functional on $\mathfrak{h}$, *ie.* $\alpha \in \mathfrak{h}^*$. If $\alpha \neq 0$, then it is called a *root* of $\mathfrak{g}$.

The simultaneous eigenvectors $e_\alpha$ corresponding to a root $\alpha \in \mathfrak{h}^*$, *ie.* the non-zero elements of $\mathfrak{g}$ satisfying $\mathrm{ad}(H)e_\alpha = \alpha(H)e_\alpha$, are the *root vectors* corresponding to $\alpha$. Along with 0, they form the *root space* $\mathfrak{g}_\alpha$ corresponding to the root $\alpha$. We have

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha, \tag{4.37}$$

where $\Delta \subset \mathfrak{h}^*$ denotes the set of all roots (the *root system*) of $\mathfrak{g}$ and $\mathfrak{g}_0$ is the eigenspace corresponding to (the non-root) 0. In fact, it turns out that $\mathfrak{g}_0 = \mathfrak{h}$ (though this is not entirely obvious). For convenience, we set $\mathfrak{g}_\alpha = 0$ for all $\alpha \notin \Delta \cup \{0\}$.

It is worthwhile noting at this point that the precise eigenspaces appearing in the decomposition (4.37) obviously depend on the choice of maximal ad-diagonalisable subalgebra $\mathfrak{h}$. When $\mathfrak{g}$ is semisimple however, Proposition 4.2 may be used to show that different choices lead to eigenspaces that are related by the action of an automorphism of $\mathfrak{g}$. In this sense, (4.37) and $\Delta$ are as independent of this choice as possible. The root system is thus an invariant of $\mathfrak{g}$. Indeed, it is even a **complete** invariant in the following sense.

**Theorem 4.6.** *If two complex finite-dimensional semisimple Lie algebras have isomorphic root systems, then they are isomorphic.*

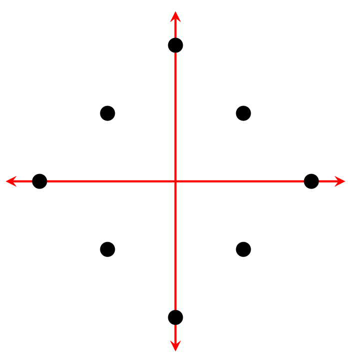
We will not prove this important result, but rather regard it as a signpost to indicate that we're on the right track. Here, two root systems are said to be isomorphic if there is an **orthogonal** linear transformation mapping one onto the other. (Here, the notion of orthogonality is the usual one, but we'll need to introduce an appropriate inner product on $\mathfrak{g}_0^*$ — see Proposition 4.15 below.)

**Example 34.** A Cartan subalgebra of $\mathfrak{sl}(2)$ is given by the span $\mathfrak{sl}(2)_0$ of $h = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. A basis of eigenvectors of $\mathrm{ad}(h)$ is given by $h$, $e = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and $f = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$. The eigenvalues are 0, 2 and $-2$, respectively. $\mathfrak{sl}(2)$ therefore has just two roots $\alpha$ and $-\alpha$, where $\alpha \in \mathfrak{sl}(2)_0^*$ is defined by $\alpha(h) = 2$. ▲

**Example 35.** A maximal ad-diagonalisable subalgebra $\mathfrak{sp}(4)_0 \subset \mathfrak{sp}(4)$ is given by the span of $H_1 = E_{11} - E_{33}$ and $H_2 = E_{22} - E_{44}$ (*cf.* Exercise 56). The roots $\alpha \in \mathfrak{sp}(4)_0^*$ are therefore defined by the values $(\alpha(H_1), \alpha(H_2))$. We find the roots and root vectors to be as follows.

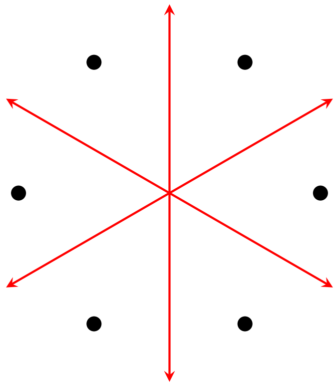
| | | | |
|---|---|---|---|
| $E_{14} + E_{23} : (1,1)$ | $E_{12} - E_{43} : (1,-1)$ | $E_{13} : (2,0)$ | $E_{24} : (0,2)$ |
| $E_{41} + E_{32} : (-1,-1)$ | $E_{21} - E_{34} : (-1,1)$ | $E_{31} : (-2,0)$ | $E_{42} : (0,-2)$ |

We plot the root system of $\mathfrak{sp}(4)$ in the real plane $\mathbb{R}^2 \subset \mathfrak{sp}(4)_0^*$ like so:



and marvel at its symmetry and elegance. ▲

**Exercise 58.** Complete a similar table for the roots and root vectors of $\mathfrak{sl}(3)$ with respect to the basis $\{H_1 = E_{11} - E_{22}, H_2 = E_{22} - E_{33}\}$ of the Cartan subalgebra $\mathfrak{sl}(3)_0$ established in Example 33. Plot a picture of the roots and weep for the missing symmetry and elegance. However, don't give up all hope just yet. Instead, show that if you use a triangular, rather than square, grid to plot the roots, symmetry and elegance is restored.



(To appreciate why, see Exercise 61 below.) ▼

We are now in a position to combine the properties of the Killing form $\kappa$ with those of the decomposition (4.37) induced from a given maximal ad-diagonalisable subalgebra $\mathfrak{g}_0$.

**Lemma 4.7.**

**(a)** $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$, *for all* $\alpha, \beta \in \mathfrak{g}_0^*$.

**(b)** *Given* $\alpha, \beta \in \mathfrak{g}_0^*$, *we have* $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ *unless* $\beta = -\alpha$.

**(c)** *If* $\kappa$ *is non-degenerate on* $\mathfrak{g}$, *then it is also non-degenerate when restricted to* $\mathfrak{g}_0$.

*Proof.* Let $e_\alpha \in \mathfrak{g}_\alpha$, $e_\beta \in \mathfrak{g}_\beta$ and $H \in \mathfrak{g}_0$. Then, the Jacobi identity and antisymmetry give

$$\begin{aligned} \text{ad}(H)[e_\alpha, e_\beta] &= [H, [e_\alpha, e_\beta]] = -[e_\alpha, [e_\beta, H]] - [e_\beta, [H, e_\alpha]] \\ &= \beta(H)[e_\alpha, e_\beta] - \alpha(H)[e_\beta, e_\alpha] = (\alpha + \beta)(H)[e_\alpha, e_\beta]. \end{aligned} \tag{4.38}$$

This shows that $[e_\alpha, e_\beta] \in \mathfrak{g}_{\alpha+\beta}$, proving a.

For b, first note that the invariance of the Killing form gives

$$\begin{aligned} \alpha(H)\kappa(e_\alpha, e_\beta) &= \kappa([H, e_\alpha], e_\beta) = -\kappa([e_\alpha, H], e_\beta) = -\kappa(e_\alpha, [H, e_\beta]) \\ &= -\beta(H)\kappa(e_\alpha, e_\beta). \end{aligned} \tag{4.39}$$

On the other hand, $\alpha + \beta \neq 0$ implies that there exists $H \in \mathfrak{g}_0$ such that $(\alpha + \beta)(H) \neq 0$. For such $H$, the above calculation clearly requires that $\kappa(e_\alpha, e_\beta) = 0$.

It follows immediately that $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}_0) = 0$, for all $\alpha \in \Delta$. Thus, for $\kappa$ to be non-degenerate on $\mathfrak{g}$, it must be non-degenerate on $\mathfrak{g}_0$. This is c. ∎

**Exercise 59.** Let $\mathfrak{g}$ be semisimple and let $\Delta$ denote its root system.

**(a)** Show that $\alpha \in \Delta$ if and only if $-\alpha \in \Delta$.

**(b)** Show that $\Delta$ is a spanning set of $\mathfrak{g}_0^*$. ▼

Now is a good time to tidy up the outstanding detail of proving that $\mathfrak{g} = \mathfrak{sp}(2n)$ and $\mathfrak{so}(n)$, excepting $\mathfrak{so}(2)$, are semisimple. Your results from Exercises 56 and 57 (should) show that the diagonal matrices of $\mathfrak{sp}(2n)$ and $\mathfrak{so}(n)$ form maximal ad-diagonalisable subalgebras, that the roots come in pairs $\alpha$ and $-\alpha$, and that the root spaces are all 1-dimensional. It therefore follows from Lemma 4.7 that the matrix representing the Killing form may be block-diagonalised with one $n \times n$ block corresponding to $\mathfrak{g}_0$ and many $2 \times 2$ blocks corresponding to pairs of root vectors.

Each of these $2 \times 2$ blocks has zeroes on the diagonals, by Lemma 4.7. The (manifestly symmetric) Killing form will therefore be degenerate if the two off-diagonal elements are zero. Unlike the situation for $\mathfrak{sl}(n)$ (Exercise 53), computing these entries explicitly is not particularly straightforward. Here is an alternative argument.

Suppose that we have already checked that the restriction of $\kappa$ to $\mathfrak{g}_0$ is non-degenerate and that the Lie bracket $[e_\alpha, e_{-\alpha}]$ is non-zero for each pair $\alpha$ and $-\alpha$ of roots. If the Killing form were degenerate, then one of the $2 \times 2$ blocks, that corresponding to $e_\alpha$ and $e_{-\alpha}$ say, would have to be identically 0. But, these root vectors would then belong to $\ker \kappa$, hence so would $[e_\alpha, e_{-\alpha}] \in \mathfrak{g}_0$, by Exercise 50. However, the latter being non-zero contradicts the non-degeneracy of $\kappa$ on $\mathfrak{g}_0$. This contradiction proves that $\kappa$ is non-degenerate.

**Exercise 60.** Complete the proof that $\mathfrak{sp}(2n)$ and $\mathfrak{so}(n)$, but not $\mathfrak{so}(2)$, are semisimple by using your results from Exercises 56 and 57 to:

**(a)** Demonstrate that $[e_\alpha, e_{-\alpha}] \neq 0$, for each pair of roots $\pm\alpha$.

**(b)** Compute explicitly the $n \times n$ block of $\kappa$ corresponding to the diagonal matrices and thereby show that it is invertible.

What happens to your general calculations when you restrict to $\mathfrak{so}(2)$? ▼

For semisimple Lie algebras $\mathfrak{g}$, we can in particular use the non-degeneracy of $\kappa$ on the Cartan subalgebra $\mathfrak{g}_0$ (*ie.* Lemma 4.7c) to conveniently identify $\mathfrak{g}_0$ with its dual. To this end, we introduce a vector space isomorphism $\imath \colon \mathfrak{g}_0 \to \mathfrak{g}_0^*$ by specifying that for any $H \in \mathfrak{g}_0$, the linear functional $\imath(H) \in \mathfrak{g}_0^*$ is defined to act on $H' \in \mathfrak{g}_0$ by

$$\imath(H)(H') = \kappa(H, H'). \tag{4.40}$$

This identification also allows us to lift the Killing form to a bilinear form $(\cdot, \cdot)$ on $\mathfrak{g}_0^*$:

$$(\lambda, \mu) = \lambda\big(\imath^{-1}(\mu)\big) = \kappa(\imath^{-1}(\lambda), \imath^{-1}(\mu)), \quad \text{for all } \lambda, \mu \in \mathfrak{g}_0^*. \tag{4.41}$$

It is common, even in introductory textbooks, for $\mathfrak{g}_0$ to be identified with its dual (meaning that $\imath$ is regarded as the identity map). As this can lead to some unfortunate confusion, we shall not do so here.

**Example 36.** Recall from Example 34 that $\mathfrak{sl}(2)$ has two roots $\pm\alpha$ and that $\alpha(h) = 2$. From Example 31, we have $\iota(h)(h) = \kappa(h,h) = 8$. Since $\{h\}$ is a basis of $\mathfrak{sl}(2)_0$, comparing therefore gives $\iota(h) = 4\alpha$, *ie.* $\iota^{-1}(\alpha) = \frac{1}{4}h$. The bilinear form on $\mathfrak{sl}(2)_0^*$ is then given by

$$\|\alpha\|^2 = (\alpha,\alpha) = \alpha\big(\iota^{-1}(\alpha)\big) = \frac{1}{4}\alpha(h) = \frac{1}{2}. \tag{4.42}$$

▲

**Example 37.** In Exercise 60, you explicitly computed the restriction of the Killing form of $\mathfrak{sp}(4)$ to its Cartan subalgebra in order to show that $\mathfrak{sp}(4)$ is semisimple. Comparing with the trace form in the defining representation, you should see that

$$\kappa(H,H') = 6\,\mathrm{tr}(HH'), \qquad \text{for all } H,H' \in \mathfrak{sp}(4)_0. \tag{4.43}$$

Referring back to Example 35, let $\alpha$ be the root whose eigenvalues $(\alpha(H_1),\alpha(H_2))$ are $(2,0)$ and, similarly, let $\beta$ be the root whose eigenvalues are $(-1,1)$. Then, the roots of $\mathfrak{sp}(4)$ are $\pm\alpha$, $\pm\beta$, $\pm(\alpha+\beta)$ and $\pm(\alpha+2\beta)$. We compare these eigenvalues with

$$\begin{aligned}
\iota(H_1)(H_1) &= \kappa(H_1,H_1) = 6\,\mathrm{tr}(E_{11}-E_{33})^2 = 12,\\
\iota(H_1)(H_2) &= \kappa(H_1,H_2) = 6\,\mathrm{tr}(E_{11}-E_{33})(E_{22}-E_{44}) = 0,\\
\iota(H_2)(H_2) &= \kappa(H_2,H_2) = 6\,\mathrm{tr}(E_{22}-E_{44})^2 = 12,
\end{aligned} \tag{4.44}$$

thereby concluding that $\alpha = \frac{1}{6}\iota(H_1)$ and $\beta = \frac{1}{12}\iota(-H_1+H_2)$.

With this, we can compute the values taken by the bilinear form on the roots, *eg.*

$$\begin{aligned}
\|\alpha\|^2 &= \kappa(\iota^{-1}(\alpha),\iota^{-1}(\alpha)) = \frac{1}{36}\kappa(H_1,H_1) = \frac{1}{3},\\
(\alpha,\beta) &= \kappa(\iota^{-1}(\alpha),\iota^{-1}(\beta)) = \frac{1}{72}\kappa(H_1,-H_1+H_2) = -\frac{1}{6},\\
\|\beta\|^2 &= \kappa(\iota^{-1}(\beta),\iota^{-1}(\beta)) = \frac{1}{144}\kappa(-H_1+H_2,-H_1+H_2) = \frac{1}{6}.
\end{aligned} \tag{4.45}$$

We can even push our luck and try to compute the *angle* $\phi_{\alpha\beta}$ between the roots $\alpha$ and $\beta$ using the cosine rule:

$$\cos\phi_{\alpha\beta} = \frac{(\alpha,\beta)}{\|\alpha\|\|\beta\|} = \frac{-\frac{1}{6}}{\sqrt{\frac{1}{3}\cdot\frac{1}{6}}} = -\frac{1}{\sqrt{2}}. \tag{4.46}$$

We conclude that this angle should be 135°. And indeed, this precisely matches the angle drawn in the picture in Example 35. ▲

**Exercise 61.** Repeat the analysis of the previous example for $\mathfrak{sl}(3)$ to show that the angles between the roots are correctly captured by drawing the root system on a triangular grid, as in the picture in Exercise 58. ▼

**Exercise 62.** Repeat to draw the root systems of $\mathfrak{so}(4)$ and $\mathfrak{so}(5)$ (with correct angles!) using your results from Exercises 57 and 60. Notice anything? ▼

Of course, we don't (yet) know whether it even makes sense to talk about angles between roots — there might be roots with zero norm-squared (deadly for the cosine rule) or even negative norm-squared (not deadly but weird). However, these investigations look so promising that surely there's something important going on here.

## 4.4. **Coroots**

We will now show that every root (along with its negative) of a semisimple Lie algebra $\mathfrak{g}$ gives rise to a subalgebra isomorphic to $\mathfrak{sl}(2)$. This observation turns out to be the key to understanding $\mathfrak{g}$ because it follows that we can regard it as a module for each of these $\mathfrak{sl}(2)$-subalgebras, *ie.* we can apply our knowledge of $\mathfrak{sl}(2)$ representation theory (Sections 3.1 and 3.2) to constrain the possibilities for a general semisimple Lie algebra $\mathfrak{g}$.

**Proposition 4.8.** *Let $\alpha$ be a root of a semisimple Lie algebra $\mathfrak{g}$. If $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_{-\alpha}$, then we have*

$$[x, y] = \kappa(x, y)\iota^{-1}(\alpha). \tag{4.47}$$

*In particular, $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ is one-dimensional.*

*Proof.* Take an arbitrary $H \in \mathfrak{g}_0$, noting that $[x, y] \in \mathfrak{g}_0$ as well. Then,

$$\begin{aligned} \kappa(H, [x, y]) &= \kappa([H, x], y) = \alpha(H)\kappa(x, y) = \kappa(\iota^{-1}(\alpha), H)\kappa(x, y) \\ &= \kappa(H, \kappa(x, y)\iota^{-1}(\alpha)), \end{aligned} \tag{4.48}$$

hence $[x, y] - \kappa(x, y)\iota^{-1}(\alpha) \in \mathfrak{g}_0$ is orthogonal to $\mathfrak{g}_0$. (4.47) now follows because $\kappa$ is non-degenerate on $\mathfrak{g}_0$, by Lemma 4.7c. The non-degeneracy of $\kappa$ on $\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}$ likewise shows that $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ cannot be 0 because $\iota^{-1}(\alpha) \neq 0$ in (4.47). ∎

**Lemma 4.9.** *Suppose that $X$, $Y$ and $Z$ are matrices satisfying the commutation relations of the Heisenberg algebra (see Example 10):*

$$[X, Y] = Z \quad \text{and} \quad [X, Z] = [Y, Z] = 0. \tag{4.49}$$

*Then, $Z$ is nilpotent.*

*Proof.* Given such matrices $X$, $Y$ and $Z$, it is clear that $\operatorname{tr} Z = \operatorname{tr}(XY - YX) = 0$ by cyclicity. However, combining cyclicity with the fact that $Z$ commutes with both $X$ and $Y$ gives

$$\operatorname{tr} Z^k = \operatorname{tr}\Big((XY - YX)Z^{k-1}\Big) = \operatorname{tr}(XYZ^{k-1}) - \operatorname{tr}(XZ^{k-1}Y) = 0, \tag{4.50}$$

for all $k \in \mathbb{Z}_{\geqslant 1}$. It follows that summing any positive-integer power of the eigenvalues of $Z$ will always give 0, hence each eigenvalue must be in fact 0. ∎

**Exercise 63.** Justify the statement in the last line of the proof of Lemma 4.9 as follows. Suppose that $Z$ is not nilpotent, so that it has some **distinct non-zero** eigenvalues $\lambda_1, \dots, \lambda_n$ whose (algebraic) multiplicities are $m_1, \dots, m_n$, respectively.

**(a)** Write the equations $\operatorname{tr} Z^k = 0$, for suitable $k$, in terms of the eigenvalues and multiplicities, thereby showing that they may be put in the form

$$\Lambda \mathrm{m} = 0, \tag{4.51}$$

where $\Lambda$ is an $n \times n$ matrix of powers of eigenvalues, m is an $n \times 1$ column vector of multiplicities and 0 is the $n \times 1$ column vector of zeroes.

**(b)** Show that $\Lambda$ is invertible. [*Hint: relate it to the $n \times n$ Vandermonde determinant.*]

**(c)** Explain why this justifies the statement in the proof. ▼

**Proposition 4.10.** *Every root $\alpha$ of a semisimple Lie algebra $\mathfrak{g}$ satisfies $\|\alpha\|^2 \neq 0$.*

*Proof.* Suppose that some $\alpha \in \Delta$ indeed has $\|\alpha\|^2 = (\alpha, \alpha) = 0$. Then, $\alpha\big(\iota^{-1}(\alpha)\big) = 0$ and so $[\iota^{-1}(\alpha), \mathfrak{g}_{\pm\alpha}] = 0$. By non-degeneracy of $\kappa$, we may choose $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_{-\alpha}$ such that $\kappa(x, y) \neq 0$. Set $z = \kappa(x, y)\iota^{-1}(\alpha) \in \mathfrak{g}_0$ and note that this is non-zero. Then, we have

$$[x, y] = z \quad \text{and} \quad [x, z] = [y, z] = 0, . \tag{4.52}$$

by (4.47). Take $X = \operatorname{ad}(x)$, $Y = \operatorname{ad}(y)$ and $Z = \operatorname{ad}(z)$. As ad is a finite-dimensional representation of $\mathfrak{g}$ (Example 21), $X$, $Y$ and $Z$ are linear transformations satisfying (4.49). $Z$ is therefore nilpotent, by Lemma 4.9. However, $z \in \mathfrak{g}_0$ is automatically ad-diagonalisable, *ie.* $Z$ is diagonalisable. The only possibility is then that $Z = 0$. But, $\operatorname{ad}(z) = 0$ implies that $z = 0$ because $\mathfrak{g}$ has zero centre (otherwise the centre would be a non-semisimple ideal). However, this contradicts the fact that we chose $z$ to be non-zero. ■

Note that this rules out roots with norm-squared equal to zero. We can therefore sensibly talk about the angle between two roots, although it isn't clear yet if this angle will have all the properties we might expect based on our experience with flat euclidean space. We will of course come back to this point soon.

**Theorem 4.11.** *Let $\mathfrak{g}$ be a semisimple Lie algebra. Given any non-zero $e_\alpha \in \mathfrak{g}_\alpha$, choose $f_\alpha \in \mathfrak{g}_{-\alpha}$ so that*

$$\kappa(e_\alpha, f_\alpha) = \frac{2}{\|\alpha\|^2} \tag{4.53}$$

*and set*

$$h_\alpha = \frac{2\iota^{-1}(\alpha)}{\|\alpha\|^2} \in \mathfrak{g}_0. \tag{4.54}$$

*Then, there is an injective homomorphism $i_\alpha \colon \mathfrak{sl}(2) \hookrightarrow \mathfrak{g}$ given by*

$$e \mapsto e_\alpha, \quad f \mapsto f_\alpha, \quad h \mapsto h_\alpha. \tag{4.55}$$

*In other words,* $\mathrm{span}\{e_\alpha, h_\alpha, f_\alpha\}$ *is a subalgebra of* $\mathfrak{g}$ *isomorphic to* $\mathfrak{sl}(2)$.

*Proof.* We first note that the choice of $f_\alpha$ is possible because of Proposition 4.10 and the non-degeneracy of $\kappa$. Next, note that defining $h_\alpha$ as above gives $[e_\alpha, f_\alpha] = h_\alpha$, by Proposition 4.8. It therefore only remains to check the Lie bracket of $h_\alpha$ with the root vectors, *eg*.

$$[h_\alpha, e_\alpha] = \frac{2}{\|\alpha\|^2}[\iota^{-1}(\alpha), e_\alpha] = \frac{2}{\|\alpha\|^2}\alpha\Big(\iota^{-1}(\alpha)\Big)e_\alpha = 2e_\alpha. \tag{4.56}$$ ■

The element $h_\alpha \in \mathfrak{g}_0$ defined in (4.54) is called the *coroot* associated to the root $\alpha \in \Delta$. A popular alternative notation for $h_\alpha$ is $\alpha^\vee$ and so we shall denote the set of coroots (the *coroot system*) of $\mathfrak{g}$ by $\Delta^\vee$. One has to be careful with notation however. For example, it is important to note that even if $\alpha$, $\beta$ and $\alpha + \beta$ are all roots, we need **not** have

$$(\alpha + \beta)^\vee = \alpha^\vee + \beta^\vee. \tag{4.57}$$

We shall see an example shortly. However, we do have $(-\alpha)^\vee = -\alpha^\vee$:

$$(-\alpha)^\vee = \frac{2\iota^{-1}(-\alpha)}{\|-\alpha\|^2} = -\frac{2\iota^{-1}(\alpha)}{\|\alpha\|^2} = -\alpha^\vee. \tag{4.58}$$

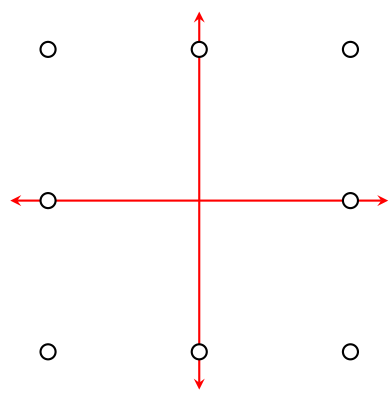
**Example 38.** Recall from Example 36 that the root $\alpha$ of $\mathfrak{sl}(2)$ satisfying $\alpha(h) = 2$ has $\iota^{-1}(\alpha) = \frac{1}{4}h$ and $\|\alpha\|^2 = \frac{1}{2}$. It follows that the corresponding coroot is

$$\alpha^\vee = \frac{2\iota^{-1}(\alpha)}{\|\alpha\|^2} = \frac{2 \cdot \frac{1}{4}h}{\frac{1}{2}} = h. \tag{4.59}$$ ▲

**Example 39.** Using the $\mathfrak{sp}(4)$ data of Example 37, we easily compute the following four coroots directly from the definition:

$$\alpha^\vee = H_1, \quad \beta^\vee = -H_1 + H_2, \quad (\alpha + \beta)^\vee = H_1 + H_2 \quad \text{and} \quad (\alpha + 2\beta)^\vee = H_2. \tag{4.60}$$

Note that taking coroots seems to swap the relative norms-squared: $\mathfrak{sp}(4)$ has *long* roots (whose norm-squared is $\frac{1}{3}$) and *short* roots (whose norm-squared is $\frac{1}{6}$), but the corresponding coroots are short and long, respectively, as measured by the Killing form.



Note also that this example gives counterexamples to (4.57). ▲

**Exercise 64.** Use the results of Exercises 61 and 62 to draw the coroot systems of $\mathfrak{sl}(3)$, $\mathfrak{so}(4)$ and $\mathfrak{so}(5)$. ▼

We close out this section by noting that it is possible to do something about the mildly annoying fact that the numerology encountered in the above examples involves "large" numbers. By way of example, take $\mathfrak{sp}(4)$ where we computed that $\kappa(H_i, H_j) = 12\delta_{ij}$. The culprit here is obviously the proportionality factor of 6 in (4.43) between the trace in the adjoint representation and the trace in the defining representation (*cf.* Exercise 54).

Obviously, the important properties of the Killing form, namely that it is symmetric, bilinear and invariant, will not change if we rescale it by a non-zero factor. If $\mathfrak{g}$ is semisimple, then such a rescaling will not affect its non-degeneracy either. However, such a rescaling will result in a rescaling of the isomorphism $\iota$ and hence a rescaling of the bilinear form $(\cdot,\cdot)$ on $\mathfrak{g}_0^*$.

**Exercise 65.** Show that rescaling $\kappa$ by a factor of $a \in \mathbb{C}^\times$, *ie.* setting $\widetilde{\kappa} = a\kappa$, leads to rescaling $\iota$ by $a$ but $(\cdot,\cdot)$ by $a^{-1}$. What happens to the coroots? ▼

The definitions (4.53) and (4.54) suggest that it would be natural to rescale the Killing form so that the factors $\frac{2}{\|\alpha\|^2}$, $\alpha \in \Delta$, are close to 1. Indeed, the most commonly observed convention used in the literature sets

$$\|\alpha\|^2 = 2, \tag{4.61}$$

where $\alpha$ is any *long* root of $\mathfrak{g}$ (meaning one whose norm-squared is maximal among all roots). From Examples 36 and 37, we see that this would correspond to rescaling $\kappa$ by $\frac{1}{4}$, for $\mathfrak{sl}(2)$, and by $\frac{1}{6}$, for $\mathfrak{sp}(4)$. Note that in both cases, this rescaled Killing form is nothing more than the trace form in the defining representation, lending further weight to the naturality of this convention.

## 4.5. **The geometry of root systems**

By Theorem 4.11, each root $\alpha$ of the semisimple Lie algebra $\mathfrak{g}$ gives rise to an $\mathfrak{sl}(2)$-subalgebra of $\mathfrak{g}$. If $i_\alpha \colon \mathfrak{sl}(2) \hookrightarrow \mathfrak{g}$ denotes the inclusion whose image is this subalgebra, then $\mathfrak{g}$ becomes a finite-dimensional $\mathfrak{sl}(2)$-module under

$$x \cdot y = [i_\alpha(x), y], \qquad \text{for all } x \in \mathfrak{sl}(2) \text{ and } y \in \mathfrak{g}. \tag{4.62}$$

Equivalently, the $\mathfrak{sl}(2)$-representation is obtained by composing $i_\alpha$ and the adjoint representation of $\mathfrak{g}$:

$$\mathfrak{sl}(2) \xrightarrow{i_\alpha} \mathfrak{g} \xrightarrow{\text{ad}} \mathfrak{gl}(\mathfrak{g}). \tag{4.63}$$

Of course, it is very common to omit all explicit references to the inclusions $i_\alpha$. Either way, the point is that we understand the finite-dimensional representation theory of $\mathfrak{sl}(2)$ extremely well (Sections 3.1 and 3.2). It is now time to put that understanding to use.

First, recall that $\mathfrak{g}$ being a finite-dimensional $\mathfrak{sl}(2)$-module means that it is isomorphic to a direct sum of finite-dimensional irreducible $\mathfrak{sl}(2)$-modules, by Weyl's theorem (Theorem 3.5). Second, every one of these finite-dimensional irreducible $\mathfrak{sl}(2)$-modules is isomorphic to some $\mathcal{L}_\lambda$, with $\lambda \in \mathbb{Z}_{\geqslant 0}$, by Theorem 3.1. Third, the eigenvalues of $h \in \mathfrak{sl}(2)$ on any given $\mathcal{L}_\lambda$ are always integers, being either all even or all odd according as to the parity of $\lambda$ (Section 3.1). On the other hand, Theorem 4.11 says that $i_\alpha$ sends $h$ to the coroot $h_\alpha = \alpha^\vee \in \mathfrak{g}$. We therefore conclude that the eigenvalues of $\operatorname{ad}(h_\alpha)$ on $\mathfrak{g}$ are also **integers**.

We're going to have to work harder to get more information about these integers, but it will be worth it. We split the bulk of the work up into the following two propositions.

**Proposition 4.12.** *Let $\alpha$ be a root of a semisimple Lie algebra $\mathfrak{g}$. Then,* $\dim \mathfrak{g}_\alpha = 1$. *Moreover, $k\alpha$ is not a root of $\mathfrak{g}$ for any $k \in \mathbb{C}$ except $k = \pm 1$.*

*Proof.* Consider first the subspace $\mathfrak{g}_{k\alpha}$, where $k \in \mathbb{C}$. If it is non-zero, then it is an eigenspace of $\operatorname{ad}(h_\alpha)$ corresponding to the eigenvalue

$$k\alpha(h_\alpha) = \frac{2k}{\|\alpha\|^2}\alpha\Big(\iota^{-1}(\alpha)\Big) = 2k. \tag{4.64}$$

As this eigenvalue must be an integer, it follows that $\mathfrak{g}_{k\alpha} = 0$ unless $k \in \frac{1}{2}\mathbb{Z}$.

Note that the adjoint action of $e_\alpha$ maps $\mathfrak{g}_{k\alpha}$ into $\mathfrak{g}_{(k+1)\alpha}$:

$$\begin{aligned} \operatorname{ad}(H)\operatorname{ad}(e_\alpha)\,\mathfrak{g}_{k\alpha} &= \big(\operatorname{ad}(e_\alpha)\operatorname{ad}(H) + \operatorname{ad}([H,e_\alpha])\big)\mathfrak{g}_{k\alpha} \\ &= \operatorname{ad}(e_\alpha)\big(\operatorname{ad}(H) + \alpha(H)\big)\mathfrak{g}_{k\alpha} \\ &= (k+1)\alpha(H)\operatorname{ad}(e_\alpha)\,\mathfrak{g}_{k\alpha}. \end{aligned} \tag{4.65}$$

Similarly, the adjoint action of $f_\alpha$ maps $\mathfrak{g}_{k\alpha}$ into $\mathfrak{g}_{(k-1)\alpha}$. The subspace

$$V = \bigoplus_{k\in\mathbb{Z}} \mathfrak{g}_{k\alpha} \tag{4.66}$$

of $\mathfrak{g}$ is therefore a module of the $\mathfrak{sl}(2)$ subalgebra spanned by $e_\alpha$, $h_\alpha$ and $f_\alpha$. It therefore decomposes as a direct sum of irreducible $\mathfrak{sl}(2)$-modules $\mathcal{L}_\lambda$. Because only integer $k$ were used in its construction, only the $\mathcal{L}_\lambda$ with $\lambda \in 2\mathbb{Z}_{\geqslant 0}$ may appear in this decomposition.

There are some easily found submodules of $V$. First, we have a copy of the adjoint $\mathfrak{sl}(2)$-module $\mathcal{L}_2$ spanned by $e_\alpha \in \mathfrak{g}_\alpha$, $h_\alpha \in \mathfrak{g}_0$ and $f_\alpha \in \mathfrak{g}_{-\alpha}$. Second, we have the trivial $\mathfrak{sl}(2)$-module $\mathcal{L}_0$ spanned by any $H \in \mathfrak{g}_0$ that is orthogonal to $h_\alpha$ (with respect to the

Killing form). That this span is indeed a module follows from $\operatorname{ad}(h_\alpha)H = [h_\alpha, H] = 0$,

$$\operatorname{ad}(f_\alpha)H = [H, f_\alpha] = \alpha(H)f_\alpha = \kappa(\iota^{-1}(\alpha), H)f_\alpha = \frac{\|\alpha\|^2}{2}\kappa(h_\alpha, H)f_\alpha = 0 \tag{4.67}$$

and the obviously similar calculation with $e_\alpha$ instead of $f_\alpha$. Because $\kappa$ is non-degenerate, the orthogonal complement of $h_\alpha$ in $\mathfrak{g}_0$ is $(r-1)$-dimensional, where $r = \operatorname{rank}\mathfrak{g}$. We therefore get $r-1$ copies of the trivial module $\mathcal{L}_0$.

These easy observations establish that $\mathcal{L}_2 \oplus (r-1)\mathcal{L}_0 \subseteq V$. However, if $V$ had any other irreducible direct summand $\mathcal{L}_\lambda$ that we had not already found, then it would have an $\operatorname{ad}(h_\alpha)$-eigenspace corresponding to eigenvalue 0 (Section 3.1) because $\lambda \in 2\mathbb{Z}_{\geqslant 0}$. However, the dimension of this eigenspace in $V$ is $r$, hence is already accounted for by $\mathcal{L}_2$ and the $r-1$ copies of $\mathcal{L}_0$. We therefore conclude that $V$ has no other irreducible submodules:

$$V \simeq \mathcal{L}_2 \oplus (r-1)\mathcal{L}_0. \tag{4.68}$$

In other words, this establishes that $\mathfrak{g}_\alpha = \operatorname{span}\{e_\alpha\}$, $\mathfrak{g}_{-\alpha} = \operatorname{span}\{f_\alpha\}$ and $\mathfrak{g}_{k\alpha} = 0$ for any integer $k$ except 1, $-1$ and 0. In particular, this proves that $\dim\mathfrak{g}_\alpha = 1$ for all $\alpha \in \Delta$.

It only remains to show that $\mathfrak{g}_{k\alpha} = 0$ for all $k \in \mathbb{Z} + \frac{1}{2}$. Here, the essential point is that the above argument shows, in particular, that "twice a root is never a root". To see how to exploit this mantra, form the $\mathfrak{sl}(2)$-submodule

$$W = \bigoplus_{k \in \mathbb{Z}+\frac{1}{2}} \mathfrak{g}_{k\alpha} \tag{4.69}$$

whose direct sum decomposition only involves the $\mathcal{L}_\lambda$ with $\lambda \in 2\mathbb{Z}_{\geqslant 0}+1$. From Section 3.1, we know that any such $\mathcal{L}_\lambda$ will have an $\operatorname{ad}(h_\alpha)$-eigenspace of eigenvalue 1. If $W$ were not zero, then we would therefore have $\mathfrak{g}_{\alpha/2} \neq 0$. But then, $\frac{1}{2}\alpha$ would be a root along with $\alpha$, contradicting our mantra. The only way out is to conclude that $W = 0$. ■

**Proposition 4.13.** *Let $\alpha$ and $\beta \neq \pm\alpha$ be roots of a semisimple Lie algebra $\mathfrak{g}$. Then*:

- **(a)** *$\beta(h_\alpha) \in \mathbb{Z}$ and $\beta - \beta(h_\alpha)\alpha$ is a root of $\mathfrak{g}$.*
- **(b)** *There exist $p, q \in \mathbb{Z}_{\geqslant 0}$ with $p - q = \beta(h_\alpha)$ such that $\beta + k\alpha$ is a root of $\mathfrak{g}$ if and only if $k$ is an integer satisfying $-p \leqslant k \leqslant q$.*
- **(c)** *If $\alpha + \beta$ is a root, then $[e_\alpha, e_\beta] = N_{\alpha\beta}e_{\alpha+\beta}$ for some scalar $N_{\alpha\beta} \neq 0$.*

*Proof.* This time, we consider the subspace

$$V_\beta = \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}_{\beta+k\alpha} \tag{4.70}$$

of $\mathfrak{g}$, noting that this is clearly a finite-dimensional $\mathfrak{sl}(2)$-submodule for the subalgebra spanned by $e_\alpha$, $h_\alpha$ and $f_\alpha$. The $\operatorname{ad}(h_\alpha)$-eigenvalues of $V_\beta$ are thus integers. In particular, the eigenvalue of $\mathfrak{g}_{\beta+k\alpha}$ is $\beta(h_\alpha) + 2k$, assuming that $\mathfrak{g}_{\beta+k\alpha} \neq 0$. As $\beta \in \Delta$, $k = 0$ gives an example in which this eigenvalue must be an integer: $\beta(h_\alpha) \in \mathbb{Z}$.

Since Proposition 4.12 ensures that $\beta + k\alpha \neq 0$, we see that $\mathfrak{g}_{\beta+k\alpha}$ is one-dimensional if $\beta + k\alpha \in \Delta$ (and is zero-dimensional otherwise). Furthermore, we have seen that the $\text{ad}(h_\alpha)$-eigenvalues of $V_\beta$ all differ by even integers. The $\mathfrak{sl}(2)$-module $V_\beta$ must therefore be irreducible. It must then have a minimal $\text{ad}(h_\alpha)$-eigenvalue $(\beta - p\alpha)(h_\alpha)$, with $p \in \mathbb{Z}_{\geqslant 0}$, and a maximal one $(\beta + q\alpha)(h_\alpha)$, with $q \in \mathbb{Z}_{\geqslant 0}$, satisfying

$$(\beta - p\alpha)(h_\alpha) = -(\beta + q\alpha)(h_\alpha) \quad \Longleftrightarrow \quad \beta(h_\alpha) = p - q. \tag{4.71}$$

Moreover, each of the "intermediate" eigenvalues is also present, hence $\beta + k\alpha$ is a root for all $k = -p, \dots, q$. Since $-p \leqslant -p + q \leqslant q$, we have the special case

$$\beta - \beta(h_\alpha)\alpha = \beta + (-p + q)\alpha \in \Delta. \tag{4.72}$$

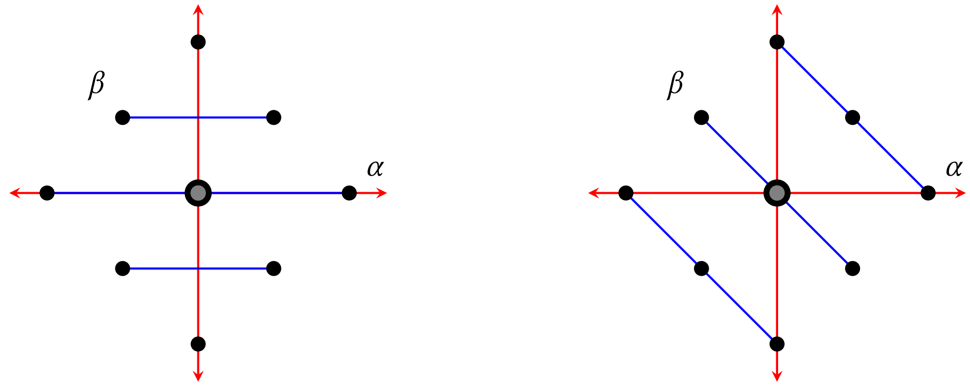
We didn't actually need the irreducibility of $V_\beta$ above, but we'll use it for the final part. So suppose that $[e_\alpha, e_\beta] = 0$. Then, $e_\beta$ is annihilated by the action of $\text{ad}(e_\alpha)$ so it must be the $\text{ad}(h_\alpha)$-eigenvector of maximal eigenvalue in the $\mathfrak{sl}(2)$-module $V_\beta$, because the latter is irreducible. However, this is a contradiction if $\alpha + \beta \in \Delta$ because then $e_{\alpha+\beta} \in \mathfrak{g}_{\beta+\alpha} \subset V_\beta$ and the $\text{ad}(h_\alpha)$-eigenvalue of $e_{\alpha+\beta}$ is clearly strictly greater than that of $e_\beta$. ■

We remark that a collection of roots of the form $\{\beta + k\alpha : k = -p, \dots, q\} \subseteq \Delta$ is sometimes called a *root string* of $\mathfrak{g}$. As noted in the proof of Proposition 4.13, they correspond to irreducible $\mathfrak{sl}(2)$-submodules of $\mathfrak{g}$.

**Example 40.** Take $\mathfrak{g} = \mathfrak{sp}(4)$ and recall the roots $\alpha$ and $\beta$ introduced in Example 37. Since $\beta(h_\alpha) = \beta(H_1) = -1$ (Example 39), Proposition 4.13 says that $\beta - \beta(h_\alpha)\alpha = \alpha + \beta$ is also a root of $\mathfrak{sp}(4)$ (which is true). We therefore have a root string $\{\beta, \beta + \alpha\}$. We might call it the $\alpha$-root string through $\beta$ (and $\beta + \alpha$). Similarly, the $\alpha$-root string through the root $\alpha + 2\beta$ is just $\{\alpha + 2\beta\}$ because $(\alpha + 2\beta)(h_\alpha) = 0$.

It is worthwhile to also look at examples of $\beta$-root strings. In particular, $\alpha(h_\beta) = \alpha(-H_1 + H_2) = -2$, so $\alpha + 2\beta \in \Delta$ and the $\beta$-root string through $\alpha$ is $\{\alpha, \alpha + \beta, \alpha + 2\beta\}$. However, we've been cheating slightly in these illustrations, because we don't actually know that these root strings aren't in fact longer than we've deduced. This is exemplified by considering the $\beta$-root string through $\alpha + \beta$: as $(\alpha + \beta)(h_\beta) = 0$, we would naturally conclude that the string is just $\{\alpha + \beta\}$. However, the previous case shows that the correct root string is in fact longer.

So there's still some work to be done, what a surprise. Until then, we can amuse ourselves by drawing pretty pictures of root strings (in blue) of $\mathfrak{sp}(4)$.




The node at the origin here is meant to remind you that even though 0 is not a root, $\mathfrak{sp}(4)_0$ is not zero (and is in fact two-dimensional). The blue lines through the origins here are not, strictly speaking, root strings in the sense of Proposition 4.13, *ie.* corresponding to some $V_\gamma$. But, we may regard them as effective root strings in the sense that they correspond to the submodules $V$ analysed in Proposition 4.12. ▲

Proposition 4.13 shows that $\beta(\alpha^\vee) \in \mathbb{Z}$, for all roots $\alpha$ and $\beta$. More is true. Because we now know that root spaces are one-dimensional (Proposition 4.12), we have

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\gamma\in\Delta} \mathbb{C}e_\gamma. \tag{4.73}$$

With no multiplicities to worry about, the Killing form is easily evaluated on the Cartan subalgebra $\mathfrak{g}_0$. In particular, we see that

$$\kappa(\alpha^\vee, \beta^\vee) = \sum_{\gamma\in\Delta} \gamma(\alpha^\vee)\gamma(\beta^\vee) \tag{4.74}$$

is an integer. It follows that

$$\|\alpha\|^2 = \kappa(\iota^{-1}(\alpha), \iota^{-1}(\alpha)) = \frac{\|\alpha\|^4}{4}\kappa(\alpha^\vee, \alpha^\vee) \quad \Rightarrow \quad \|\alpha\|^2 = \frac{4}{\kappa(\alpha^\vee, \alpha^\vee)}, \tag{4.75}$$

a rational number. Moreover,

$$(\alpha, \beta) = \kappa(\iota^{-1}(\alpha), \iota^{-1}(\beta)) = \frac{\|\alpha\|^2\|\beta\|^2}{4}\kappa(\alpha^\vee, \beta^\vee) = \frac{4\kappa(\alpha^\vee, \beta^\vee)}{\kappa(\alpha^\vee, \alpha^\vee)\kappa(\beta^\vee, \beta^\vee)} \tag{4.76}$$

is also rational, for any $\alpha, \beta \in \Delta$.

This rationality goes even further. Recall that the root system $\Delta$ is a spanning set for $\mathfrak{g}_0^*$ (Exercise 59). We may therefore choose a basis $\Pi = \{\alpha_1, \dots, \alpha_r\}$ of $\mathfrak{g}_0^*$, where $r = \operatorname{rank} \mathfrak{g}$, in which each of the basis elements is a root.

**Lemma 4.14.** *Every root is a rational linear combination of the roots in* $\Pi$.

*Proof.* Expand a root $\beta \in \Delta$ in this basis. The result is

$$\beta = \sum_{j=1}^{r} b_j \alpha_j \quad \Rightarrow \quad (\alpha_i, \beta) = \sum_{j=1}^{r} b_j (\alpha_i, \alpha_j), \tag{4.77}$$

where $i$ is any integer from 1 to $r$ and the $b_j$ are (*a priori*) complex numbers. However, the matrix form of these equations is

$$\begin{pmatrix} (\alpha_1, \beta) \\ \vdots \\ (\alpha_r, \beta) \end{pmatrix} = \begin{pmatrix} (\alpha_1, \alpha_1) & \cdots & (\alpha_1, \alpha_r) \\ \vdots & \ddots & \vdots \\ (\alpha_r, \alpha_1) & \cdots & (\alpha_r, \alpha_r) \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix}. \tag{4.78}$$

The matrix here is precisely the representing matrix, with respect to the basis $\Pi$, for the bilinear form $(\cdot, \cdot)$. As this form is non-degenerate, the matrix is invertible. Moreover, the entries of this matrix are rational and so the entries of the inverse matrix will be rational as well (by the row-reduction algorithm). The $b_i$ are thus rational linear combinations of the $(\alpha_i, \beta)$, hence are rational themselves. ∎

This then motivates the introduction of the **rational** vector space

$$\mathsf{R} = \operatorname{span}_{\mathbb{Q}} \Delta \subset \mathfrak{g}_0^*. \tag{4.79}$$

By Lemma 4.14, this vector space would not change if we replaced $\Delta$ by $\Pi$ in the definition. In particular, $\dim_{\mathbb{Q}} \mathsf{R} = r = \operatorname{rank} \mathfrak{g}$.

Of course, we could also take the real span instead of the rational one. Note that this real space $\mathsf{R} \otimes_{\mathbb{Q}} \mathbb{R}$ is the ambient space of the root diagrams that we have been drawing throughout. The next result shows that our discussion of lengths and angles in these diagrams is perfectly justified.

**Proposition 4.15.** $\mathsf{R}$ *is a rational inner product space, ie. the non-degenerate invariant symmetric bilinear form* $(\cdot, \cdot)$ *is positive-definite on* $\mathsf{R}$.

*Proof.* Given $\beta \in \mathsf{R} \subset \mathfrak{g}_0^*$, one has

$$\|\beta\|^2 = \kappa(\iota^{-1}(\beta), \iota^{-1}(\beta)) = \sum_{\alpha \in \Delta} \alpha(\iota^{-1}(\beta))^2 = \sum_{\alpha \in \Delta} (\alpha, \beta)^2, \tag{4.80}$$

as in (4.74). Since $(\alpha, \beta) \in \mathbb{Q}$, for all $\alpha \in \Delta$, this proves that $\|\beta\|^2$ is non-negative. Moreover, $\|\beta\|^2 = 0$ can only be achieved if $(\alpha, \beta) = 0$ for all $\alpha \in \Delta$. As the roots form a basis of $\mathfrak{g}_0^*$, $\beta$ is then in the kernel of $(\cdot, \cdot)$. But, this form is non-degenerate, so $\beta = 0$. ∎

Recall that the angle $\phi_{\alpha\beta}$ between two non-zero elements $\alpha, \beta \in \mathsf{R}$ is determined by the cosine rule:

$$\cos \phi_{\alpha\beta} = \frac{(\alpha, \beta)}{\|\alpha\| \|\beta\|}. \tag{4.81}$$

The fact that $\beta(\alpha^\vee) \in \mathbb{Z}$ (Proposition 4.13) now leads to a very strong constraint on the angle $\phi_{\alpha\beta}$. First, note that

$$\beta(\alpha^\vee) = \frac{2}{\|\alpha\|^2}\beta\left(\iota^{-1}(\alpha)\right) = \frac{2(\beta,\alpha)}{\|\alpha\|^2} = \frac{2\|\beta\|}{\|\alpha\|}\cos\phi_{\alpha\beta}. \tag{4.82}$$

Multiplying by $\alpha(\beta^\vee)$ therefore gives

$$\alpha(\beta^\vee)\beta(\alpha^\vee) = 4\cos^2\phi_{\alpha\beta}. \tag{4.83}$$

The left-hand side is an integer, while the right-hand side is real number lying between 0 and 4. We conclude that both sides may only take values in the finite set $\{0, 1, 2, 3, 4\}$.

**Proposition 4.16.** *Let $\alpha$ and $\beta$ be roots of a semisimple Lie algebra $\mathfrak{g}$. Then, up to exchanging $\alpha$ and $\beta$, one of the following possibilities must occur.*

| $\alpha(\beta^\vee)$ | 0 | 1 | −1 | 1 | −1 | 1 | −1 | 2 | −2 |
|---|---|---|---|---|---|---|---|---|---|
| $\beta(\alpha^\vee)$ | 0 | 1 | −1 | 2 | −2 | 3 | −3 | 2 | −2 |
| $\phi_{\alpha\beta}$ | 90° | 60° | 120° | 45° | 135° | 30° | 150° | 0° | 180° |
| $\|\beta\|^2/\|\alpha\|^2$ | – | 1 | 1 | 2 | 2 | 3 | 3 | 1 | 1 |

**Exercise 66.** Prove Proposition 4.16, paying close attention to any missing cases. ▼

**Exercise 67.** Let $\alpha$ and $\beta \neq \pm\alpha$ be roots of a semisimple Lie algebra $\mathfrak{g}$. Use Propositions 4.13 and 4.16 to prove the following statements:

**(a)** If the angle between $\alpha$ and $\beta$ is obtuse, *ie.* $\phi_{\alpha\beta} > 90°$, then $\alpha + \beta$ is a root.
**(b)** If the angle between $\alpha$ and $\beta$ is acute, *ie.* $\phi_{\alpha\beta} < 90°$, then $\alpha - \beta$ is a root. ▼

Recall from Exercise 51 that the simple ideals $\mathfrak{g}_i$ of a semisimple Lie algebra $\mathfrak{g}$ are orthogonal with respect to the Killing form: $\kappa(\mathfrak{g}_i, \mathfrak{g}_j) = 0$ if $i \neq j$. This will then obviously hold for the Cartan subalgebras $(\mathfrak{g}_i)_0$ as well, hence their duals will be orthogonal with respect to $(\cdot,\cdot)$. In particular, roots coming from different simple ideals will be automatically orthogonal.

There is a converse to this conclusion. Suppose that a root system $\Delta$ is the (necessarily disjoint) union of two non-empty subsets of roots $\Delta_1$ and $\Delta_2$ whose elements are mutually orthogonal: $(\Delta_1, \Delta_2) = 0$. Then, $\alpha + \beta \notin \Delta$ for all $\alpha \in \Delta_1$ and $\beta \in \Delta_2$ because this sum is not orthogonal to $\Delta_1$ nor to $\Delta_2$. It follows that $e_\alpha$ and $e_\beta$ commute. It is easy to check that $h_\alpha$ and $e_\beta$ likewise commute, as do $e_\alpha$ and $h_\beta$. Moreover, the $h_\alpha$ with $\alpha \in \Delta_1$ and the $h_\beta$ with $\beta \in \Delta_2$ are orthogonal with respect to $\kappa$, so their spans have zero intersection. This proves that $\mathfrak{g}$ admits the following (non-trivial) decomposition into **ideals**:

$$\mathfrak{g} = \text{span}\{e_\alpha, h_\alpha : \alpha \in \Delta_1\} \oplus \text{span}\{e_\beta, h_\beta : \beta \in \Delta_2\}. \tag{4.84}$$

In other words, $\mathfrak{g}$ is not simple.

This motivates the definition that a root system $\Delta$ is *reducible* if it may be written as the union of two non-empty subsets of roots $\Delta_1$ and $\Delta_2$ satisfying $(\Delta_1, \Delta_2) = 0$. If $\Delta$ cannot be written in such a fashion, then it is *irreducible*. Furthermore, the above argument establishes the following result.

**Proposition 4.17.** *A semisimple Lie algebra is simple if and only if its root system is irreducible.*

We know that $\mathfrak{sl}(2)$ is simple (Example 17) but, up to now, we haven't established the simplicity of any other examples. With this new tool however, it is easy to prove simplicity.

**Exercise 68.** Use Proposition 4.17 and the explicit descriptions of their root systems to decide which of $\mathfrak{sl}(3)$, $\mathfrak{sp}(4)$, $\mathfrak{so}(4)$ and $\mathfrak{so}(5)$ are simple Lie algebras. ▼

**Exercise 69.** Show that there can be at most two different lengths among the roots of a simple Lie algebra. ▼

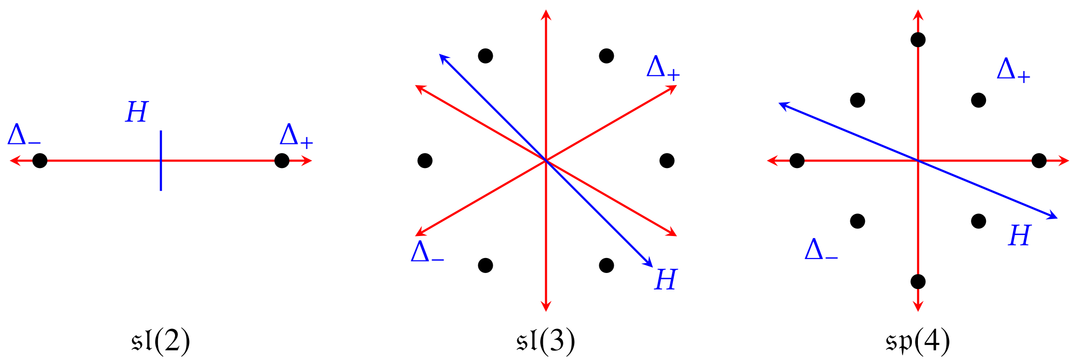
### 4.6. **Simple roots**

As $\alpha \in \Delta$ implies that $-\alpha \in \Delta$ (Exercise 59), we can partition the root system into sets of *positive* and *negative* roots:

$$\Delta = \Delta_+ \cup \Delta_-. \tag{4.85}$$

A natural way of doing this is to choose a codimension-1 hyperplane $H \subset \mathsf{R}$ through the origin and declare that the roots on one side of $H$ are positive whilst those on the other side are negative. Better yet, we may choose $\eta \in \mathsf{R}$ normal to $H$ and then define a root $\alpha \in \Delta$ to be positive or negative according to the sign of $(\eta, \alpha)$. For this to work, we have to ensure that no root lies in $H$. However, this is easy to do as the set of roots is finite.

Here are some examples of partitions into positive and negative roots.




Of course, these are not the only such partitions and so the splitting of the root system into positive and negative roots depends on a choice. However, just as with the choice of Cartan subalgebra, it turns out that any two choices of partition are related by an automorphism of the root system and so the choice doesn't really matter at all. Here, an automorphism of a root system $\Delta$ is just an invertible linear map that preserves $\Delta$.

Fix a choice of positive roots. An important concept now arises: that of a positive root that cannot be expressed as the sum of two positive roots. Such a root is said to be *simple*. In a sense, simple roots are the "smallest" positive roots — they are the closest positive roots to the partitioning hyperplane $H$. The utility of this concept stems from the fact that all positive roots must be expressible as a linear combination of simple roots with non-negative integer coefficients!

**Lemma 4.18.** *The angle between two (different) simple roots is never acute. Equivalently, any two simple roots $\alpha_1$ and $\alpha_2 \neq \alpha_1$ satisfy $(\alpha_1, \alpha_2) \leqslant 0$.*

*Proof.* If the angle $\phi_{\alpha_1\alpha_2}$ were acute, then $\alpha_1 - \alpha_2$ would be a root by Exercise 67. If this root were positive, then $\alpha_1 = (\alpha_1 - \alpha_2) + \alpha_2$ would be a sum of two positive roots. However, if this root were negative, then $\alpha_2 = -(\alpha_1 - \alpha_2) + \alpha_1$ would likewise be a sum of two positive roots. Both possibilities contradict the fact that $\alpha_1$ and $\alpha_2$ are simple. ■

**Proposition 4.19.** *Let $\Pi = \{\alpha_1, \dots, \alpha_r\}$ be a maximal set of simple roots of a semisimple Lie algebra $\mathfrak{g}$. Then, $\Pi$ is a basis of $\mathsf{R}$, so $r = \operatorname{rank}\mathfrak{g}$, and every positive root $\beta \in \Delta_+$ has the form*

$$\beta = \sum_{i=1}^{r} n_i \alpha_i, \quad \textit{for some } n_i \in \mathbb{Z}_{\geqslant 0}. \tag{4.86}$$

*Proof.* We first show that every positive root is an $\mathbb{Z}_{\geqslant 0}$-linear combination of roots in $\Pi$. Suppose that there are positive roots that cannot be expressed as an $\mathbb{Z}_{\geqslant 0}$-linear combination of roots in $\Pi$. Then, there is one such, $\alpha$ say, for which $(\eta, \alpha)$ is minimal, where $\eta$ is any fixed normal to the partitioning hyperplane $H$. Now obviously $\alpha$ is not in $\Pi$, so it is not simple ($\Pi$ is a maximal set of simple roots). We therefore have $\alpha = \beta_1 + \beta_2$ for some $\beta_1, \beta_2 \in \Delta_+$. As $(\eta, \beta_i) > 0$, it follows that $(\eta, \beta_i) < (\eta, \alpha)$, for $i = 1, 2$. The minimality of $\alpha$ now means that both $\beta_1$ and $\beta_2$ are $\mathbb{Z}_{\geqslant 0}$-linear combination of roots in $\Pi$. But, this means $\alpha$ is too. This contradiction means that no such $\alpha$ exists, hence all positive roots are $\mathbb{Z}_{\geqslant 0}$-linear combinations of roots in $\Pi$.

Since the roots certainly span $\mathsf{R}$, this shows that $\Pi$ is also a spanning set. It remains to show linear independence. So, suppose that there exist $a_i \in \mathbb{Q}$ such that

$$\sum_{i=1}^{r} a_i \alpha_i = 0. \tag{4.87}$$

Split this into two sums, one with strictly positive coefficients and one with strictly negative ones. We then have

$$\sum_{i\,:\,a_i>0} a_i\alpha_i = \sum_{j\,:\,a_j<0} (-a_j)\alpha_j. \tag{4.88}$$

Call the common value of these sums $\beta$. Now, Lemma 4.18 forces $(\alpha_i, \alpha_j) \leqslant 0$, because we must have $i \neq j$, hence

$$\|\beta\|^2 = \sum_{i\,:\,a_i>0} \sum_{j\,:\,a_j<0} a_i(-a_j)(\alpha_i, \alpha_j) \leqslant 0. \tag{4.89}$$

By positive-definiteness (Proposition 4.15), $\|\beta\|^2 = 0$ and so $\beta = 0$. We therefore find that

$$\sum_{i\,:\,a_i>0} a_i(\eta, \alpha_i) = 0 \quad \text{and} \quad \sum_{j\,:\,a_j<0} (-a_j)(\eta, \alpha_j) = 0, \tag{4.90}$$

as they are both secretly just $(\eta, \beta)$. But, $(\eta, \alpha_i) > 0$ because simple roots are positive. So the above equations are impossible if there are any $a_i > 0$ and any $a_i < 0$, respectively. The conclusion is that in fact all $a_i = 0$ and that the $\alpha_i$ are linearly independent. ■

**Example 41.** Throughout, we have been illustrating the general theory with the classical Lie algebras $\mathfrak{sl}(n)$, $\mathfrak{so}(n)$ and $\mathfrak{sp}(2n)$. The Cartan subalgebra was always chosen to consist of diagonal matrices and the root vectors then naturally split into two subsets under the transpose operation (in the defining representation). This gives us a natural choice of positive and negative roots, hence simple roots.

We tabulate the results as follows.

**(a)** $\mathfrak{g} = \mathfrak{sl}(r+1)$:
Cartan basis: $H_k = E_{kk} - E_{k+1\,k+1}$, for $1 \leqslant k \leqslant r$.
Simple root vectors: $E_{i\,i+1}$, for $1 \leqslant i \leqslant r$.
Positive root vectors: $E_{ij}$, for $1 \leqslant i < j \leqslant r+1$.
$\mathrm{ad}(H_k)$-eigenvalues: $\delta_{ik} - \delta_{jk} - \delta_{i\,k+1} + \delta_{j\,k+1}$, for $1 \leqslant i < j \leqslant r+1$ and $1 \leqslant k \leqslant r$.

**(b)** $\mathfrak{g} = \mathfrak{so}(2r+1)$:
Cartan basis: $H_k = E_{kk} - E_{r+k\,r+k}$, for $1 \leqslant k \leqslant r$.
Simple root vectors:
- $E_{i\,i+1} - E_{r+i+1\,r+i}$, for $1 \leqslant i \leqslant r-1$.
- $E_{r\,2r+1} - E_{2r+1\,2r}$.

Positive root vectors:
- $E_{ij} - E_{r+j\,r+i}$, for $1 \leqslant i < j \leqslant r$.
- $E_{i\,r+j} - E_{j\,r+i}$, for $1 \leqslant i < j \leqslant r$.
- $E_{i\,2r+1} - E_{2r+1\,r+i}$, for $1 \leqslant i \leqslant r$.

$\mathrm{ad}(H_k)$-eigenvalues:
- $\delta_{ik} - \delta_{jk}$, for $1 \leqslant i < j \leqslant r$ and $1 \leqslant k \leqslant r$.
- $\delta_{ik} + \delta_{jk}$, for $1 \leqslant i < j \leqslant r$ and $1 \leqslant k \leqslant r$.
- $\delta_{ik}$, for $1 \leqslant i \leqslant r$ and $1 \leqslant k \leqslant r$.

**(c)** $\mathfrak{g} = \mathfrak{sp}(2r)$:

Cartan basis: $H_k = E_{kk} - E_{r+k\,r+k}$, for $1 \leqslant k \leqslant r$.

Simple root vectors:
- $E_{i\,i+1} - E_{r+i+1\,r+i}$, for $1 \leqslant i \leqslant r-1$.
- $E_{r\,2r}$.

Positive root vectors:
- $E_{ij} - E_{r+j\,r+i}$, for $1 \leqslant i < j \leqslant r$.
- $E_{i\,r+j} + E_{j\,r+i}$, for $1 \leqslant i \leqslant j \leqslant r$.

$\mathrm{ad}(H_k)$-eigenvalues:
- $\delta_{ik} - \delta_{jk}$, for $1 \leqslant i < j \leqslant r$ and $1 \leqslant k \leqslant r$.
- $\delta_{ik} + \delta_{jk}$, for $1 \leqslant i \leqslant j \leqslant r$ and $1 \leqslant k \leqslant r$.

**(d)** $\mathfrak{g} = \mathfrak{so}(2r)$:

Cartan basis: $H_k = E_{kk} - E_{r+k\,r+k}$, for $1 \leqslant k \leqslant r$.

Simple root vectors:
- $E_{i\,i+1} - E_{r+i+1\,r+i}$, for $1 \leqslant i \leqslant r-1$.
- $E_{r-1\,2r} - E_{r\,2r-1}$.

Positive root vectors:
- $E_{ij} - E_{r+j\,r+i}$, for $1 \leqslant i < j \leqslant r$.
- $E_{i\,r+j} - E_{j\,r+i}$, for $1 \leqslant i < j \leqslant r$.

$\mathrm{ad}(H_k)$-eigenvalues:
- $\delta_{ik} - \delta_{jk}$, for $1 \leqslant i < j \leqslant r$ and $1 \leqslant k \leqslant r$.
- $\delta_{ik} + \delta_{jk}$, for $1 \leqslant i < j \leqslant r$ and $1 \leqslant k \leqslant r$. ▲

**Exercise 70.** A natural basis of the Cartan subalgebra is afforded by the *simple coroots*, *ie.* the coroots $\alpha_i^\vee$, $i = 1, \ldots, \operatorname{rank} \mathfrak{g}$.

**(a)** Why do the simple coroots form a basis of $\mathfrak{g}_0$?

**(b)** Show that the transpose of any of the positive root vectors $e_\alpha$ given in Example 41 for the classical Lie algebras is proportional to the negative root vector $f_\alpha$.

**(c)** For each of the classical Lie algebras, compute this constant of proportionality for every simple root and thereby deduce formulae for the simple coroots in terms of the Cartan basis elements $H_k$ of Example 41. ▼

**Exercise 71.** Draw, in $\mathsf{R} \simeq \mathbb{R}^3$, the root systems of $\mathfrak{sl}(4)$, $\mathfrak{sp}(6)$ and $\mathfrak{so}(7)$ (making sure that lengths and angles are roughly correct). [This is a good exercise to do with a computer, especially if you can rotate the drawing around!] ▼

## 4.7. **Cartan matrices**

Now that we know what simple roots are, we can define the *Cartan matrix* of $\mathfrak{g}$. Given a set of simple roots $\Pi = \{\alpha_1, \ldots, \alpha_r\}$, $r = \operatorname{rank} \mathfrak{g}$, this is the $r \times r$ matrix $\mathsf{A}$ whose entries are the following integers:

$$\mathsf{A}_{ij} = \alpha_j(\alpha_i^\vee) = \frac{2(\alpha_i, \alpha_j)}{\|\alpha_i\|^2}, \quad \text{for all } i, j = 1, \ldots, r. \tag{4.91}$$

Some people find it more natural to swap $i$ and $j$ in this definition (*ie.* their Cartan matrix is the transpose of ours). The convention used above was chosen so that

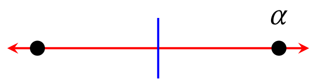
$$[h_{\alpha_i}, e_{\alpha_j}] = \mathsf{A}_{ij} e_{\alpha_j}. \tag{4.92}$$

Note that the Cartan matrix is independent of the normalisation used for the Killing form!

**Example 42.**

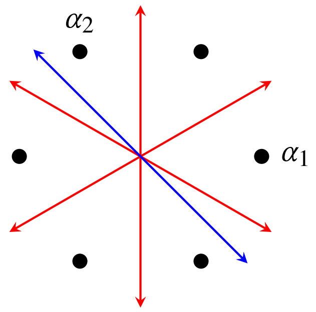
**(a)** $\mathfrak{sl}(2)$ has rank 1, with $\Pi = \{\alpha\}$, $e_\alpha = e = E_{12}$ and $\alpha^\vee = h$ (Example 38). Since $\alpha(h) = 2$, we have

$$\mathsf{A} = \begin{pmatrix} 2 \end{pmatrix}.$$



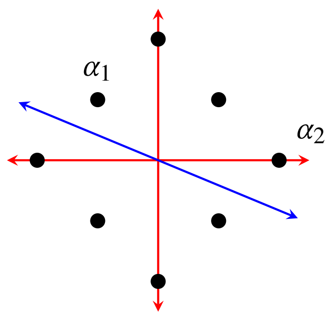

**(b)** $\mathfrak{sl}(3)$ has rank 2, with $\Pi = \{\alpha_1, \alpha_2\}$, $e_{\alpha_1} = E_{12}$, $e_{\alpha_2} = E_{23}$, $\alpha_1^\vee = H_1$ and $\alpha_2^\vee = H_2$ (in the conventions of Example 41 and Exercise 70). The Cartan matrix is thus

$$\mathsf{A} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$




**(c)** $\mathfrak{sp}(4)$ has rank 2, with $\Pi = \{\alpha_1, \alpha_2\}$, $e_{\alpha_1} = E_{12} - E_{43}$, $e_{\alpha_2} = E_{24}$, $\alpha_1^\vee = H_1 - H_2$ and $\alpha_2^\vee = H_2$ (in the conventions of Example 41 and Exercise 70). We therefore get

$$\mathsf{A} = \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}.$$




▲

**Exercise 72.** Use Example 41 and Exercise 70 to show that the Cartan matrices of $\mathfrak{sl}(r+1)$, $r \geqslant 2$; $\mathfrak{so}(2r+1)$, $r \geqslant 3$; $\mathfrak{sp}(2r)$, $r \geqslant 2$; and $\mathfrak{so}(2r)$, $r \geqslant 4$, are given by

$$\mathsf{A}^{\mathfrak{sl}(r+1)} = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}, \quad \mathsf{A}^{\mathfrak{so}(2r+1)} = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & \cdots & 0 & -2 & 2 \end{pmatrix},$$

$$\mathsf{A}^{\mathfrak{sp}(2r)} = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 2 & -2 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}, \quad \mathsf{A}^{\mathfrak{so}(2r)} = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 2 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 0 & 2 \end{pmatrix}. \tag{4.93}$$

What are the Cartan matrices of $\mathfrak{so}(4)$, $\mathfrak{so}(5)$ and $\mathfrak{so}(6)$. ▼

**Exercise 73.** For $\mathfrak{sl}(3)$, $\mathfrak{sp}(4)$, $\mathfrak{so}(4)$ and $\mathfrak{so}(5)$, choose a different hyperplane (to define a different set of simple roots) and show that the Cartan matrix doesn't change (except perhaps for a reordering of its rows and columns). ▼

Indeed, up to a reordering of the rows and columns, *ie.* a permutation of the simple roots, the Cartan matrix is independent of the choice of hyperplane. It is an invariant of the semisimple Lie algebra. Better yet, it is a complete invariant.

**Theorem 4.20.** *If two complex finite-dimensional semisimple Lie algebras have the same Cartan matrix, up to permuting the simple roots, then they are isomorphic.*

We shall not prove this important result. Instead, we shall show how one can recover the root system $\Delta$ of a semisimple Lie algebra $\mathfrak{g}$ from its Cartan matrix $\mathsf{A}$. Knowing that $\mathsf{A}$ determines $\Delta$ makes it plausible that it also determines $\mathfrak{g}$, up to isomorphism.

First, it is clearly sufficient to determine the positive roots $\Delta_+$ from a set $\Pi = \{\alpha_1, \dots, \alpha_r\}$ of simple roots. About the latter, we only know the Cartan matrix entries $\mathsf{A}_{ij} = \alpha_j(\alpha_i^\vee)$. However, this information immediately tells us the angle between every pair of simple roots and their relative lengths, by Proposition 4.16. So, choose a simple root $\alpha_i \in \Pi$ and a positive root $\beta \neq \alpha_i$, the latter being known explicitly as a linear combination of simple roots. The key here is Proposition 4.13 which says that $\beta + k\alpha_i \in \Delta_+$ if and only if $-p \leqslant k \leqslant q$, where $p, q \in \mathbb{Z}_{\geqslant 0}$ satisfy $p - q = \beta(\alpha_i^\vee)$. The point is that we can determine $p$, and thus $q$, if we know all the positive roots that are "lower" than $\beta$. This ordering on $\Delta_+$ is given by the *height* function which assigns to

$$\beta = \sum_{i=1}^{r} m_i\alpha_i, \quad \text{the positive integer } \sum_{i=1}^{r} m_i. \tag{4.94}$$

Rather than try to explain this abstractly, we illustrate the procedure with two examples.

**Example 43.** The Cartan matrix $\mathsf{A}$ of $\mathfrak{sp}(4)$ was given in Example 42c. With the simple roots $\alpha_1$ and $\alpha_2$, we know that $\alpha_1 + k\alpha_2$ is a root for $k = 0$ but not for $k = -1$. In the language of Proposition 4.13, this tells us that $p = 0$, hence $q = -\alpha_1(\alpha_2^\vee) = -\mathsf{A}_{21} = 1$. That is, $\alpha_1 + \alpha_2$ is a root, but $\alpha_1 + k\alpha_2$ is not a root for $k \geqslant 2$.

We can also consider $\alpha_2 + k\alpha_1$, again noting that this is a root for $k = 0$ but not for $k = -1$. With $p$ identified as 0, we now get $q = -\alpha_2(\alpha_1^\vee) = -\mathsf{A}_{12} = 2$, hence we conclude that $\alpha_2 + \alpha_1$ and $\alpha_2 + 2\alpha_1$ are roots while $\alpha_2 + k\alpha_1$ is not if $k \geqslant 3$.

The next step would be to consider whether there are new roots to find by considering Proposition 4.13 with $\alpha$ a simple root and $\beta$ a root of height 2, *ie.* $\beta = \alpha_1 + \alpha_2$. This doesn't lead to any new roots, so we try the root of height 3: $\beta = 2\alpha_1 + \alpha_2$. This also gives no new roots. We have therefore found all the positive roots of $\mathfrak{sp}(4)$. ▲

**Example 44.** The Cartan matrix of $\mathfrak{so}(7)$ is

$$\mathsf{A} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -2 & 2 \end{pmatrix}. \tag{4.95}$$

Again, taking $i \neq j$, we see that $\alpha_j + k\alpha_i \in \Delta$ for $k = 0$, but not for $k = -1$, hence $p = 0$ and $q = -\mathsf{A}_{ij}$. In this way, we learn that $\alpha_1 + \alpha_2$, $\alpha_2 + \alpha_3$ and $\alpha_2 + 2\alpha_3$ are roots and that $\alpha_1 + \alpha_3$ is not.

Considering $\beta + k\alpha_i$, where $\beta$ is one of the height-2 roots already found, we conclude that $\alpha_1 + \alpha_2 + \alpha_3$ and $\alpha_1 + \alpha_2 + 2\alpha_3$ are roots. For example, $\beta = \alpha_1 + \alpha_2$ and $i = 3$ give $p = 0$, hence

$$q = -(\alpha_1 + \alpha_2)(\alpha_3^\vee) = -\mathsf{A}_{31} - \mathsf{A}_{32} = 2. \tag{4.96}$$

Repeating with $\beta$ a height-3 root gives no new roots, but the height-4 root $\alpha_1 + \alpha_2 + 2\alpha_3$ with $i = 2$ yields $p = 0$ (since $\alpha_1 + 2\alpha_3 \notin \Delta$), hence

$$q = -(\alpha_1 + \alpha_2 + 2\alpha_3)(\alpha_2^\vee) = -\mathsf{A}_{21} - \mathsf{A}_{22} - 2\mathsf{A}_{23} = 1. \tag{4.97}$$

Thus, $\alpha_1 + 2\alpha_2 + 2\alpha_3 \in \Delta$. One can check exhaustively that this is the last root to find. Alternatively, one can note that we have found 9 positive roots and that this must be all by dimension counting:

$$|\Delta_+| = \frac{\dim \mathfrak{g} - \dim \mathfrak{g}_0}{2} = \frac{21 - 3}{2} = 9. \tag{4.98}$$ ▲

**Exercise 74.** Determine the root systems of $\mathfrak{sl}(4)$ and $\mathfrak{so}(8)$ from their Cartan matrices. ▼

Theorem 4.20 reduces the classification of (complex finite-dimensional) semisimple Lie algebras to the classification of Cartan matrices (up to permutations). These matrices have the following properties.

**Proposition 4.21.** *The Cartan matrix* $\mathsf{A}$ *of a semisimple Lie algebra* $\mathfrak{g}$ *satisfies*:

- **(a)** $\mathsf{A}_{ii} = 2$.
- **(b)** $\mathsf{A}_{ij} \in \{0, -1, -2, -3\}$, *for all* $i \neq j$.
- **(c)** $\mathsf{A}_{ij} < -1$ *implies that* $\mathsf{A}_{ji} = -1$.
- **(d)** $\mathsf{A}_{ij} = 0$ *if and only if* $\mathsf{A}_{ji} = 0$.

*Moreover, let* $S \subseteq \{1, \dots, r\}$ *be non-empty* (*with* $r = \operatorname{rank} \mathfrak{g}$) *and let* $\mathsf{A}^S$ *denote the* $|S| \times |S|$ *matrix obtained from* $\mathsf{A}$ *by removing, for every* $i \notin S$, *the* $i$*-th row and column. Then,* $\det \mathsf{A}^S > 0$.

*Proof.* a follows immediately from the definition (4.91). b likewise follows from the definition by combining Propositions 4.15 and 4.16 with Lemma 4.18. c follows from the table in Proposition 4.16 and a. d is also an easy consequence of the definition and Proposition 4.10:

$$\mathsf{A}_{ji} = \frac{2(\alpha_j, \alpha_i)}{\|\alpha_j\|^2} = \frac{\|\alpha_i\|^2}{\|\alpha_j\|^2} \frac{2(\alpha_i, \alpha_j)}{\|\alpha_i\|^2} = \frac{\|\alpha_i\|^2}{\|\alpha_j\|^2} \mathsf{A}_{ij}. \tag{4.99}$$

It remains to establish $\det \mathsf{A}^S > 0$ for all non-empty $S$. Let $\mathsf{R}^S$ be the subspace of $\mathsf{R}$ spanned by the $\alpha_i$ with $i \in S$. The key is to note that the $|S| \times |S|$ matrix $\mathsf{B}^S$ with entries

$$\mathsf{B}^S_{ij} = (\alpha_i, \alpha_j) = \frac{\|\alpha_i\|^2}{2} \mathsf{A}_{ij}, \qquad \text{for all } i, j \in S, \tag{4.100}$$

is the matrix representing $(\cdot,\cdot)$ on $\mathsf{R}^S$, with respect to the basis $\{\alpha_i : i \in S\}$. As $(\cdot,\cdot)$ is positive-definite on $\mathsf{R}$ (Proposition 4.15), it is also positive-definite on $\mathsf{R}^S$ and so we have $\det \mathsf{B}^S > 0$, whence

$$\det \mathsf{A}^S = \frac{2^r}{\|\alpha_1\|^2 \cdots \|\alpha_r\|^2} \det \mathsf{B}^S > 0. \tag{4.101}$$ ■

The classification of (complex finite-dimensional) semisimple Lie algebras now proceeds by first classifying the matrices A that satisfy the conditions listed in Proposition 4.21. Then, one can ask if every such matrix is indeed the Cartan matrix of some semisimple Lie algebra. The hardest condition to analyse is of course the determinant condition. Putting this aside for a moment, the remaining conditions can be conveniently encoded in graphs known as *Dynkin diagrams*.

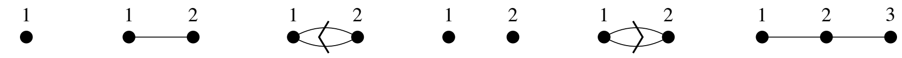
## 4.8. **Dynkin diagrams**

To draw the Dynkin diagram of an $r \times r$ matrix $\mathsf{A}$ satisfying conditions a–d of Proposition 4.21,

- take $r$ vertices labelled from 1 to $r$;
- for $i \neq j$, connect the $i$-th and $j$-th vertices by $\max\{-\mathsf{A}_{ij}, -\mathsf{A}_{ji}\}$ edges;
- if two vertices $i$ and $j$ are connected by more than one edge, draw an arrow on the edges pointing from $j$ to $i$, if $\mathsf{A}_{ij} < -1$, and from $i$ to $j$, if $\mathsf{A}_{ji} < -1$.

The Dynkin diagram of a semisimple Lie algebra is that of its Cartan matrix. Note that because $\|\alpha_i\|^2 \mathsf{A}_{ij} = 2(\alpha_i, \alpha_j) = \|\alpha_j\|^2 \mathsf{A}_{ji}$, the arrows on a Dynkin diagram (assuming it comes from the Cartan matrix of a semisimple Lie algebra) always point from a long simple root to a short one. Note also that a Dynkin diagram has no edges connecting a vertex to itself. However, it might have a loop that involves more than one vertex.

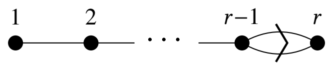
**Example 45.** We list the Cartan matrices and Dynkin diagrams of some low-rank semisimple Lie algebras below.

| $\mathfrak{g}$ | $\mathfrak{sl}(2)$ | $\mathfrak{sl}(3)$ | $\mathfrak{sp}(4)$ | $\mathfrak{so}(4)$ | $\mathfrak{so}(5)$ | $\mathfrak{so}(6)$ |
|---|---|---|---|---|---|---|
| $\mathsf{A}^{\mathfrak{g}}$ | $\begin{pmatrix} 2 \end{pmatrix}$ | $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ | $\begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}$ | $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ | $\begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix}$ | $\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$ |
| $\Gamma^{\mathfrak{g}}$ |  | | | | | |

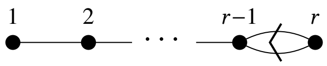
▲

**Example 46.** The Dynkin diagrams of the remaining simple classical Lie algebras (whose Cartan matrices were worked out in Exercise 72) are as follows.

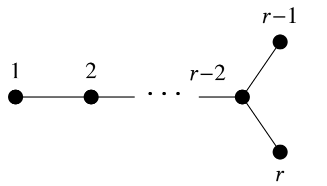

$\Gamma^{\mathfrak{sl}(r+1)}$, $r \geqslant 2$




$\Gamma^{\mathfrak{so}(2r+1)}$, $r \geqslant 3$




$\Gamma^{\mathfrak{sp}(2r)}$, $r \geqslant 3$




$\Gamma^{\mathfrak{so}(2r)}$, $r \geqslant 4$

Beautiful, aren't they? ▲

One of the main advantages of working with the Dynkin diagrams is that they easily account for the reordering of the simple roots that make the Cartan matrix non-unique. Indeed, a Cartan matrix may be obtained from another one by reordering if and only if their Dynkin diagrams (without vertex labels) are isomorphic as graphs. It therefore follows that two semisimple Lie algebras are isomorphic if and only if their Dynkin diagrams are. From the diagrams listed above, it is now easy to deduce that

$$\mathfrak{sp}(4) \simeq \mathfrak{so}(5) \quad \text{and} \quad \mathfrak{sl}(4) \simeq \mathfrak{so}(6). \tag{4.102}$$

A second advantage of Dynkin diagrams relates to the interpretation of whether they are connected or not. We saw in Proposition 4.17 that a simple Lie algebra corresponds to an irreducible root system, *ie.* one that cannot be written as a disjoint union of two non-empty mutually orthogonal subsets. It is clear that if a root system is reducible, so $\Delta = \Delta_1 \cup \Delta_2$ with $\Delta_1, \Delta_2 \neq \varnothing$ and $(\Delta_1, \Delta_2) = 0$, then we have a similar partition of the simple roots: $\Pi = \Pi_1 \cup \Pi_2$ with $\Pi_1, \Pi_2 \neq \varnothing$ and $(\Pi_1, \Pi_2) = 0$.

**Exercise 75.** Prove this statement about $\Pi$. ▼

The algorithm for reconstructing the root system from the simple roots may now be used to show the converse: a partition of simple roots into non-empty orthogonal subsets implies a reducible root system. But, simple roots are orthogonal if and only if $\mathsf{A}_{ij} = 0$, so such a partition allows one to order the simple roots to make the Cartan matrix block-diagonal. We thus have the following sharpening of Proposition 4.17.

**Proposition 4.22.** *A semisimple Lie algebra $\mathfrak{g}$ is simple if and only if its Dynkin diagram $\Gamma^{\mathfrak{g}}$ is connected.*

In particular, the classical Lie algebras $\mathfrak{sl}(r+1)$, $\mathfrak{so}(2r+1)$, $\mathfrak{sp}(2r)$ and $\mathfrak{so}(2r)$, excepting $\mathfrak{so}(2)$ and $\mathfrak{so}(4)$, are simple because their Dynkin diagrams, exhibited in Examples 45 and 46, are connected.

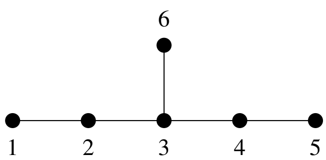
Our aim is to classify the (complex finite-dimensional) simple Lie algebras. The above technology reduces this classification to determining their (connected) Dynkin diagrams, which we know must satisfy the determinant condition of Proposition 4.21. Whilst this condition is a little awkward to analyse, it lends itself to the graph-theoretic world of Dynkin diagrams because it essentially states that every **subgraph** of the Dynkin diagram of a simple Lie algebra must also correspond to a Cartan matrix of positive determinant.

The classification of simple Lie algebras may therefore be attacked by classifying connected Dynkin diagrams whose subgraphs all have positive determinant. This, in turn, may be done iteratively by adding vertices and edges to known positive-determinant diagrams and recording whether the result is a new positive-determinant diagram or something else (which can then never appear as a subdiagram again).

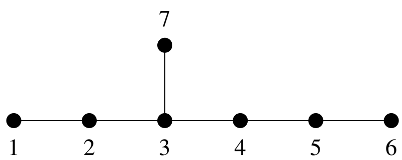
This turns out to be an efficient, though quite combinatorial, route to the classification condition. The determinant condition itself turns out to be so powerful in this respect that one could even drop some of the other constraints of Proposition 4.21 (*cf.* Carter, Ch. 6).

**Theorem 4.23.** *The connected Dynkin diagrams whose corresponding matrices satisfy all the conditions of Proposition 4.21 are exhausted by those given in Examples 45 and 46 (omitting that of $\mathfrak{so}(4)$) and the following five exceptions.*

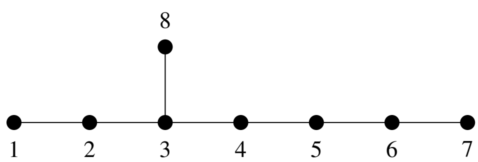
$\Gamma^{\mathfrak{e}_6}$

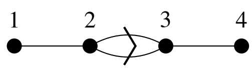



$\Gamma^{\mathfrak{e}_7}$

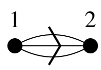



$\Gamma^{\mathfrak{e}_8}$




$\Gamma^{\mathfrak{f}_4}$




$\Gamma^{\mathfrak{g}_2}$

We remark that the ordering of the labels of these exceptional Dynkin diagrams varies considerably between sources. It is therefore common to see people refer to "Bourbaki's ordering", as a standard reference, in papers.

We shall not prove this classification result formally (see Carter, Ch. 6), but instead content ourselves with sketching the basic ideas. First, one shows that the five exceptional Dynkin diagrams do actually satisfy the conditions of Proposition 4.21 (those of Examples 45 and 46 do by construction).

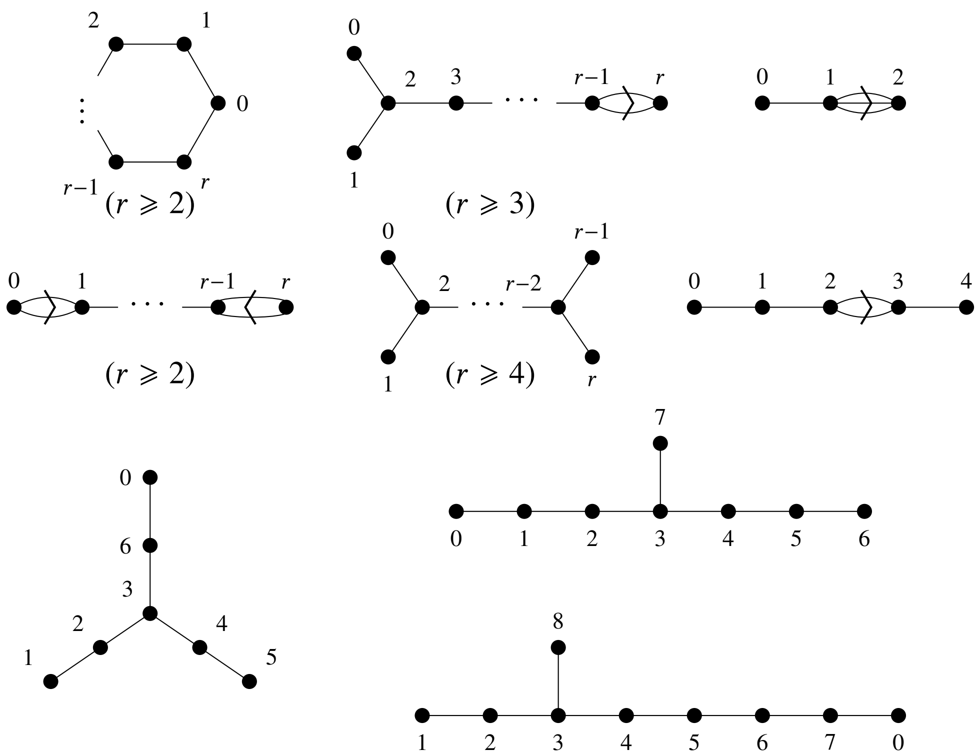
**Exercise 76.**

**(a)** Write down the "Cartan matrices" corresponding to each of the exceptional Dynkin diagrams of Theorem 4.23 and show that each has positive determinant. Explain carefully why computing one determinant each is enough to show that the determinant condition of Proposition 4.21 is satisfied.

**(b)** For completeness, compute the determinants (inductively) of the Cartan matrices corresponding to the Dynkin diagrams of Examples 45 and 46. ▼

Next, one adds a "zeroth" vertex to the Dynkin diagrams of Examples 45 and 46 and Theorem 4.23, connecting it to the original diagram with some new edges so that the determinant of the corresponding extended matrix is 0. These diagrams violate Proposition 4.21 and so cannot be the Dynkin diagram of a simple Lie algebra.

**Exercise 77.** Show that the following diagrams are not the Dynkin diagrams of any simple Lie algebra.




▼

Finally, one uses the fact that the Dynkin diagram of any simple Lie algebra cannot have any of these “forbidden diagrams” as subgraphs to rule out any but those found in Examples 45 and 46 and Theorem 4.23.

This classifies the Dynkin diagrams that satisfy the positive-determinant condition. To complete the classification of the simple Lie algebras, it only remains to show that the exceptional diagrams of Theorem 4.23 are actually the Dynkin diagrams of some simple Lie algebras, now known as *exceptional* Lie algebras. (The fact that such a Lie algebra is unique, if it exists, is Theorem 4.20.) These Lie algebras indeed exist and are denoted, somewhat unimaginatively, by

$$\mathfrak{e}_6, \quad \mathfrak{e}_7, \quad \mathfrak{e}_8, \quad \mathfrak{f}_4 \quad \text{and} \quad \mathfrak{g}_2,$$

as in Theorem 4.23. In each case, the subscript denotes the Lie algebra's rank. They were originally constructed explicitly — indeed, Cartan's doctoral thesis is where these constructions were first found.

Interestingly, there are constructions of each of the exceptional Lie algebras that employ the **octonions** $\mathbb{O}$. This is somewhat in the spirit of orthogonality being associated to $\mathbb{R}$ and unitarity being associated to $\mathbb{C}$, whilst being symplectic is properly regarded as “orthogonality” for the quaternions $\mathbb{H}$. The fact that there are infinite families of simple Lie algebras associated with $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$ (namely $\mathfrak{so}(n)$, $\mathfrak{sl}(n) \simeq \mathfrak{su}(n)^{\mathbb{C}}$ and $\mathfrak{sp}(2n)$) but only finitely many associated with $\mathbb{O}$ may be due to the fact that $\mathbb{O}$ is not associative so a few coincidences need to occur before one can construct an associative Lie group or an associative Lie algebra (in the sense of the Jacobi identity) from it.[1]

Luckily, a general construction was found much later by Serre.

**Theorem 4.24** (Serre)**.** *Let* $\mathsf{A}$ *be the* $r \times r$ *Cartan matrix of a root system* $\Delta$*. Then, the* (*complex*) *Lie algebra* $\mathfrak{g}$ *defined by generators* $E_i$*,* $H_i$ *and* $F_i$*,* $i = 1, \ldots, r$*, and relations*

$$[H_i, H_j] = 0, \quad [H_i, E_j] = \mathsf{A}_{ij} E_j, \quad [H_i, F_j] = -\mathsf{A}_{ij} F_j, \quad [E_i, F_j] = \delta_{ij} H_j, \quad \forall i, j, \tag{4.103a}$$

$$\operatorname{ad}(E_i)^{1-\mathsf{A}_{ij}} E_j = \operatorname{ad}(F_i)^{1-\mathsf{A}_{ij}} F_j = 0, \quad \forall i \neq j, \tag{4.103b}$$

*is semisimple of rank* $r$ *and finite-dimensional. Moreover,* $\mathfrak{g}$ *has* $\operatorname{span}\{H_1, \ldots, H_r\}$ *as a Cartan subalgebra and* $\mathsf{A}$ *as a Cartan matrix.*

The relations (4.103a) are known as the Chevalley relations whilst (4.103b) are the Serre relations. We shall not prove this powerful theorem. However, we can check that these relations are indeed satisfied by all semisimple Lie algebras.

[1]Those who find this fascinating might well look at Adams' *Lectures on exceptional Lie groups* and (*eg.*) https://mathoverflow.net/questions/99736/beautiful-descriptions-of-exceptional-groups , http://math.ucr.edu/home/baez/octonions/ .

**Exercise 78.** Use the root strings of Proposition 4.13 (*cf.* Examples 43 and 44) to prove that the Chevalley-Serre relations (4.103) are satisfied in a semisimple Lie algebra if we let $E_i = e_{\alpha_i}$, $H_i = h_{\alpha_i}$ and $F_i = f_{\alpha_i}$. ▼

This concludes the classification of (complex finite-dimensional) semisimple Lie algebras. They are determined uniquely by their root systems, which are in turn completely determined by their simple roots, which are in turn completely determined by their Cartan matrices or, equivalently, by their Dynkin diagrams. Up to isomorphism, there are four infinite families of simple Lie algebras:

**(a)** $\mathfrak{a}_r \equiv \mathfrak{sl}(r+1)$, for $r \geqslant 1$;
**(b)** $\mathfrak{b}_r \equiv \mathfrak{so}(2r+1)$, for $r \geqslant 1$;
**(c)** $\mathfrak{c}_r \equiv \mathfrak{sp}(2r)$, for $r \geqslant 1$; and
**(d)** $\mathfrak{d}_r \equiv \mathfrak{so}(2r)$, for $r \geqslant 3$.

The *a*–*d* labelling of these classical Lie algebras is traditional (though often capital letters are used instead) and aligns with the names given to the five exceptional simple Lie algebras $\mathfrak{e}_6$, $\mathfrak{e}_7$, $\mathfrak{e}_8$, $\mathfrak{f}_4$ and $\mathfrak{g}_2$. We note again, for completeness, that $\mathfrak{a}_1 \simeq \mathfrak{b}_1 \simeq \mathfrak{c}_1$, $\mathfrak{b}_2 \simeq \mathfrak{c}_2$ and $\mathfrak{a}_3 \simeq \mathfrak{d}_3$; otherwise, the simple Lie algebras listed are all mutually inequivalent.

A (semi)simple Lie algebra whose Dynkin diagram contains no arrows, *ie.*no double or triple bond, is said to be *simply laced*. From the classification result, we see that the simply laced simple Lie algebras are the $\mathfrak{sl}(r+1)$, for $r \geqslant 1$, the $\mathfrak{so}(2r)$, for $r \geqslant 4$, along with $\mathfrak{e}_6$, $\mathfrak{e}_7$ and $\mathfrak{e}_8$. In other words, the simply laced cases are of types $\mathfrak{a}$, $\mathfrak{d}$ or $\mathfrak{e}$. Such "ADE" classifications appear in all sorts of places in mathematics including the platonic solids, the eigenvalues of symmetric matrices, algebraic singularities, finite-type quivers and subfactors.

More interestingly, mathematical physicists have started adding to this list. For example, the modular-invariant partition functions of the Virasoro minimal model conformal field theories also fall into an ADE-type classification. Indeed, quantum mechanics, quantum field theory and string theory have all proven quite adept in recent years at not only generating new examples of ADE, but also of explaining why different ADE classifications are related. But we have no space for this here! We instead conclude by noting that no discussion of the classification of simple Lie algebras would be complete without an investigation of the lowest-rank exceptional Lie algebra.

**Exercise 79.**

**(a)** Given the Dynkin diagram in Theorem 4.23, write down the Cartan matrix of $\mathfrak{g}_2$, the relative lengths of the simple roots and the angle between them.
**(b)** Reconstruct the root system of $\mathfrak{g}_2$ using the algorithm described in Examples 43 and 44. Draw it, making sure to respect the relative lengths of the roots and the angles between them.

**(c)** Show that $\mathfrak{g}_2$ has a **semisimple** subalgebra spanned by the simple coroots and the root vectors $e_\alpha$ with $\alpha$ long.

**(d)** Carefully compute the simple coroots of this subalgebra and show that its Cartan matrix is that of $\mathfrak{sl}(3)$, thereby proving that there is an injective homomorphism $\mathfrak{sl}(3) \hookrightarrow \mathfrak{g}_2$ of Lie algebras.

**(e)** Do the simple coroots and the short root vectors likewise span a subalgebra isomorphic to $\mathfrak{sl}(3)$? ▼

**Exercise 80.** For any simple Lie algebra $\mathfrak{g}$, there is a unique root, called the *highest root* and traditionally denoted by $\theta$, for which $\theta + \alpha_i$ is not a root, for all $i = 1, \dots, \operatorname{rank} \mathfrak{g}$.

**(a)** Determine the root vector $e_\theta$, for each of the simple classical Lie algebras, using the data in Example 41.

**(b)** In each case, fix the proportionality constant between the Killing form and the trace form in the defining representation, *cf.* Exercise 54.

**(c)** Compute $\kappa(e_\theta, e_{-\theta})$ and $\kappa(\theta^\vee, \theta^\vee)$, thereby determining $f_\theta$ and $\|\theta\|^2$.

**(d)** Finish by calculating $b \in \mathbb{C}$ such that $\widetilde{\kappa} = b\kappa$ yields $\|\theta\|^2 = 2$. Compare $\widetilde{\kappa}$ with the trace form in the defining representation. ▼

## 5. Representations of semisimple Lie algebras

Having learned all about Killing forms, roots, Cartan matrices and Dynkin diagrams, it is now time to see how representations of general semisimple Lie algebras work. Our experience with $\mathfrak{sl}(2)$ leads us to expect that things will work out nicely (and they do). However, this is far from trivial. In a representation consisting of $n \times n$ matrices, each independent Lie bracket will give rise to $\sim n^2$ simultaneous equations that have to be satisfied. So, constructing representations is actually not at all an obvious game and classifying them seems to be non-trivial. However, this wouldn't be much of a course if there wasn't a beautiful result waiting to be uncovered...

### 5.1. **Weights and weight modules**

We first introduce a new concept (of tremendous importance): the weights of a representation (or module). This is the generalisation of the concept of roots from the adjoint representation (the Lie algebra itself) to arbitrary representations. For convenience, we assume throughout that $\mathfrak{g}$ denotes a (finite-dimensional complex) semisimple Lie algebra.

So, we define a *weight vector* of a representation $\pi$ on the $\mathfrak{g}$-module $V$ to be a simultaneous eigenvector $v$ of the linear operators $\pi(H)$, for all $H$ in the Cartan subalgebra $\mathfrak{g}_0$ of $\mathfrak{g}$. The *weight* of this eigenvector is then the linear functional $\lambda \in \mathfrak{g}_0^*$ that satisfies

$$\pi(H)v = \lambda(H)v, \qquad ie.\ Hv = \lambda(H)v. \tag{5.1}$$

The set of all weight vectors (extended by 0) of a module $V$, for a given weight, *ie.* a simultaneous eigenspace for the Cartan subalgebra, is called a *weight space* of $V$.

You can see that if $\pi = \mathrm{ad}$, so $V = \mathfrak{g}$, then $Hv = \mathrm{ad}(H)v = [H, v]$ and we recover the definition of the roots of $\mathfrak{g}$ by taking $v$ to be a corresponding root vector. Thus, every root of $\mathfrak{g}$ is a weight of the adjoint module. However, this module has another weight because we may also take $v$ to be a (non-zero) Cartan element so that $[H, v] = 0$. Roots are required to be non-zero, whereas there is no such constraint on weights. Indeed, 0 is always a weight of the adjoint module of a semisimple Lie algebra.

**Exercise 81.** Show that if $V$ and $W$ are isomorphic $\mathfrak{g}$-modules, then they have the same set of weights. Is the converse true? ▼

Every finite-dimensional $\mathfrak{g}$-module $V$ possesses a weight vector. To see this, note that the simple coroot $h_{\alpha_1}$ has a non-zero eigenspace $V_1 \subseteq V$ of eigenvalue $\lambda(h_{\alpha_1})$ (where $\lambda \in \mathfrak{g}_0^*$). If $\operatorname{rank} \mathfrak{g} = 1$, then we are done. Otherwise, observe that $h_{\alpha_2}$ preserves $V_1$ because it commutes with $h_{\alpha_1}$: $v \in V_1$ if and only if $h_{\alpha_1} v = \lambda(h_{\alpha_1})v$, hence $h_{\alpha_1}(h_{\alpha_2} v) = h_{\alpha_2} h_{\alpha_1} v = \lambda(h_{\alpha_1}) h_{\alpha_2} v$ and so $h_{\alpha_2} v \in V_1$. It follows that $h_{\alpha_2}$ has a non-zero eigenspace $V_2 \subseteq V_1$. If $\operatorname{rank} \mathfrak{g} = 2$, then we are done; otherwise, continue until we have a non-zero simultaneous eigenspace of all the simple coroots.

A module that possesses a basis whose elements are all weight vectors is called a *weight module*. We will shortly show that every finite-dimensional irreducible module is a weight module. A later goal is to extend this to all finite-dimensional modules; see Section 5.7.

**Example 47.** Recall from Theorem 3.1 that the finite-dimensional irreducible $\mathfrak{sl}(2)$-modules $\mathcal{L}_\lambda$ are parametrised by a non-negative integer $\lambda$ and that $\mathcal{L}_\lambda$ is spanned by vectors $v_\lambda^{(n)} = f^n v_\lambda$, $n = 0, 1, \dots, \lambda$, that satisfy

$$hv_\lambda^{(n)} = (\lambda - 2n)v_\lambda^{(n)}. \tag{5.2}$$

As $h$ spans a Cartan subalgebra, we see that each $v_\lambda^{(n)}$ is a weight vector corresponding to the weight $\lambda^{(n)}$ defined by $\lambda^{(n)}(h) = \lambda - 2n$. The $\mathfrak{sl}(2)$-modules $\mathcal{L}_\lambda$ are therefore all weight modules and their weight spaces are all 1-dimensional. ▲

**Exercise 82.** If $V$ and $W$ are weight modules, show that $V \oplus W$, $V \otimes W$ and $V^*$ are also weight modules. What are their weights, in terms of those of $V$ and $W$? ▼

**Proposition 5.1.** *A finite-dimensional irreducible* $\mathfrak{g}$*-module* $V$ *is a weight module.*

*Proof.* Fix $i \in \{1, \dots, \operatorname{rank} \mathfrak{g}\}$. As $\dim V < \infty$, the simple coroot $h_{\alpha_i} \in \mathfrak{g}_0$ has an eigenvector when acting upon $V$. Thus, the span $W_i$ of the eigenvectors of $h_{\alpha_i}$ is non-zero. Moreover, $W_i$ is a submodule of $V$ as the simple coroots and the root vectors form a basis of $\mathfrak{g}$:

$$h_{\alpha_i} v = \lambda v \qquad \Rightarrow \qquad \begin{aligned} h_{\alpha_i} h_{\alpha_j} v &= \lambda h_{\alpha_j} v, \\ h_{\alpha_i} e_\alpha v &= \big(\lambda + \alpha(h_{\alpha_i})\big) e_\alpha v. \end{aligned} \tag{5.3}$$

Because $V$ is assumed to be irreducible, we must have $W_i = V$. The simple coroot $h_{\alpha_i}$ therefore acts diagonalisably on $V$.

Since $i$ was arbitrary, this conclusion holds for all simple coroots. As they commute, there is a basis of $V$ that simultaneously diagonalises their action. This basis also simultaneously diagonalises the action of $\mathfrak{g}_0$ because the simple coroots span the Cartan subalgebra. This basis therefore consists of weight vectors, by definition, hence $V$ is a weight module. ■

As the simple coroots $h_{\alpha_i} = \alpha_i^\vee$ form a basis of the Cartan subalgebra $\mathfrak{g}_0$, it is often convenient to equate a weight $\lambda \in \mathfrak{g}_0^*$ with the vector whose $i$-th entry is the eigenvalue $\lambda(\alpha_i^\vee)$, $i = 1, \dots, \operatorname{rank} \mathfrak{g}$. For the $\mathfrak{sl}(2)$-modules $\mathcal{L}_\lambda$, this means identifying $\lambda^{(n)}$ with the 1-vector $(\lambda - 2n)$ (or just $\lambda - 2n$). In general, the numbers

$$\lambda_i = \lambda(\alpha_i^\vee) \tag{5.4}$$

are called the *Dynkin labels* of the weight $\lambda$. The Dynkin labels of the simple root $\alpha_j$ are then $\alpha_j(\alpha_i^\vee) = \mathsf{A}_{ij}$, $i = 1, \dots, \operatorname{rank} \mathfrak{g}$, *ie.* the Dynkin labels of the simple roots are the columns of the Cartan matrix.

Abstractly, identifying weights with their Dynkin labels is equivalent to choosing the basis $\{\omega_i\}$ of $\mathfrak{g}_0^*$ that is *dual* to the simple coroot basis of $\mathfrak{g}_0$:

$$\omega_i(\alpha_j^\vee) = \delta_{ij}, \qquad \text{for all } i, j = 1, \dots, \operatorname{rank} \mathfrak{g}. \tag{5.5}$$

More precisely, (5.4) and (5.5) together imply that

$$\lambda = \sum_{i=1}^{\operatorname{rank} \mathfrak{g}} \lambda_i \omega_i. \tag{5.6}$$

The basis elements $\omega_i$ are called the *fundamental weights* of $\mathfrak{g}$ because of the fundamental role that they will play in the theory of finite-dimensional representations of $\mathfrak{g}$. We remark that the fundamental weights are orthogonal to the simple roots in the following sense:

$$(\omega_i, \alpha_j) = \omega_i(\imath^{-1}(\alpha_j)) = \frac{\|\alpha_j\|^2}{2} \omega_i(\alpha_j^\vee) = \frac{\|\alpha_j\|^2}{2} \delta_{ij}. \tag{5.7}$$

However, we emphasise that this does not mean that $\omega_i$ and $\alpha_i$ are proportional!

**Example 48.** When $\mathfrak{g} = \mathfrak{sl}(2)$, $\dim \mathfrak{g}_0^* = 1$ and so the fundamental weight $\omega_1$ and the simple root $\alpha_1$ must be proportional. Writing $\omega_1 = a\alpha_1$, we see that (5.7) gives

$$\frac{\|\alpha_1\|^2}{2} = (\omega_1, \alpha_1) = a\|\alpha_1\|^2 \quad \Rightarrow \quad a = \frac{1}{2}. \tag{5.8}$$

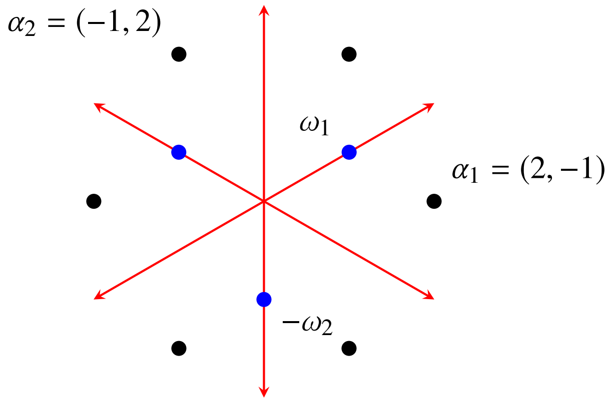
We thus have $\alpha_1 = 2\omega_1$, consistent with our earlier discovery that the Dynkin labels of the simple roots, in this case the sole Dynkin label is $(\alpha_1)_1 = 2$, form the columns of the Cartan matrix, which in this case is just $\begin{pmatrix} 2 \end{pmatrix}$. ▲

**Example 49.** In the defining $\mathfrak{sl}(3)$-module $\mathbb{C}^3$, the simple coroots $\alpha_1^\vee$ and $\alpha_2^\vee$ are (customarily) represented by the matrices $E_{11} - E_{22}$ and $E_{22} - E_{33}$, respectively:

$$\alpha_1^\vee \longmapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \alpha_2^\vee \longmapsto \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{5.9}$$

Being diagonal, it is easy to read off the simultaneous eigenvalues: $(1, 0)$, $(-1, 1)$ and $(0, -1)$. These are therefore the weights of the defining module, expressed in the basis of fundamental weights, *ie.* these are the Dynkin labels of the three weights, $\omega_1$, $\omega_2 - \omega_1$ and $-\omega_2$, of the defining representation.

If we draw these weights on the same diagram that we used for roots, we get the following (beautifully symmetric) picture.




Here, we've drawn the weights of the defining module in blue. The (hopefully by now familiar) root system of $\mathfrak{sl}(3)$ is shown, for comparison, in black. Note that $\alpha_1$ is orthogonal to $\omega_2$ and $\alpha_2$ is orthogonal to $\omega_1$, as dictated by (5.7). ▲

**Exercise 83.** Repeat the above example for the dual of the defining representation of $\mathfrak{sl}(3)$, recalling that the dual was defined in Exercise 29. Repeat again for the defining representations of $\mathfrak{sp}(4)$ and $\mathfrak{so}(4)$, then repeat for the dual representations. ▼

One thing that we can see in the weight diagram of Example 49 is that the difference between any two distinct weights of the defining module of $\mathfrak{sl}(3)$ is a root of $\mathfrak{sl}(3)$. We can understand why this, or rather something like this, must be true for any irreducible weight module: If we start from a weight vector $v$ of weight $\lambda$, then acting with a Cartan element does not change the weight, whereas acting with the root vector $e_\alpha$ shifts the weight to $\lambda + \alpha$. Acting repeatedly with root vectors on $v$ necessarily generates the whole irreducible module — otherwise the vectors generated from $v$ by this action would span a non-zero proper submodule. It follows that all the weights of an irreducible weight module must differ from any one of its weights by integer linear combinations of roots.

We can make this more precise by introducing the *root lattice* $\mathsf{Q}$ which is the free abelian subgroup of $\mathfrak{g}_0^*$ spanned by the simple roots $\alpha_i$. Its elements therefore consist of linear combinations of the simple roots whose coefficients are all integers. Clearly, every root belongs to $\mathsf{Q}$.

Recall that an indecomposable $\mathfrak{g}$-module is one that cannot be written as the direct sum of two submodules.

**Proposition 5.2.** *If $\lambda$ and $\mu$ are weights of an indecomposable weight module, then $\lambda - \mu \in \mathsf{Q}$.*

*Proof.* Suppose that $\lambda - \mu \notin \mathsf{Q}$. Then, there is a non-zero submodule generated by the weight vectors whose weights are equal to $\lambda$ mod $\mathsf{Q}$ and a non-zero submodule generated by the weight vectors whose weights are not equal to $\lambda$ mod $\mathsf{Q}$. As neither submodule is zero, but their intersection must be zero, it follows that the module decomposes as the direct sum of these two submodules. ■

Along with the root lattice $\mathsf{Q}$, we also introduce the *weight lattice* $\mathsf{P}$, which is the free abelian subgroup of $\mathfrak{g}_0^*$ spanned by the fundamental weights $\omega_i$. We will also find it convenient to define $\mathsf{Q}_\geqslant$ and $\mathsf{P}_\geqslant$ to be the **non-negative** integer linear combinations of the $\alpha_i$ and $\omega_i$, respectively. We note that $\mathsf{Q} \subseteq \mathsf{P}$, because the Dynkin labels of the simple roots are the entries of the Cartan matrix (*ie.* they are integers), but that $\mathsf{Q}_\geqslant \nsubseteq \mathsf{P}_\geqslant$. One reason to care about the weight lattice is the following result:

**Proposition 5.3.** *If $\lambda$ is a weight of a finite-dimensional $\mathfrak{g}$-module $V$, then its Dynkin labels are integers, ie. $\lambda \in \mathsf{P}$.*

*Proof.* Fix $i \in \{1, \dots, \operatorname{rank} \mathfrak{g}\}$ and recall from Theorem 4.11 that the simple root $\alpha_i$ gives rise to an inclusion (injective homomorphism of Lie algebras) $\mathfrak{sl}(2) \hookrightarrow \mathfrak{g}$ in which the Cartan element $h \in \mathfrak{sl}(2)$ is mapped to the coroot $\alpha_i^\vee$. By composing this homomorphism with the representation of $\mathfrak{g}$ on $V$, $\mathfrak{sl}(2) \hookrightarrow \mathfrak{g} \to \mathfrak{gl}(V)$, we obtain a finite-dimensional representation of $\mathfrak{sl}(2)$ on $V$.

By Weyl's theorem (Theorem 3.5), $V$ is completely reducible as an $\mathfrak{sl}(2)$-module, decomposing (uniquely) into a direct sum of irreducible $\mathfrak{sl}(2)$-modules. Each irreducible is isomorphic to one of the $\mathcal{L}_\lambda$, by Theorem 3.1, and so the eigenvalues of $h$, hence those of $\alpha_i^\vee$ as well, are integers. But, if $v \in V$ is a weight vector of weight $\lambda$ (with respect to the action of $\mathfrak{g}$), then $\alpha_i^\vee v = \lambda(\alpha_i^\vee) v = \lambda_i v$ shows that the $i$-th Dynkin label of $\lambda$ is an integer. Since $i$ was arbitrary, the proposition is proved. ■

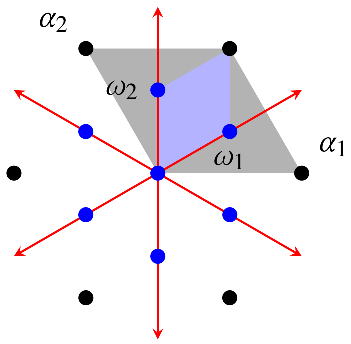
This result allows us to settle another detail relating to the trace forms $\kappa_\pi$, $\pi$ a finite-dimensional $\mathfrak{g}$-representation, introduced in Section 4.1. We recall from Exercise 54 that every trace form is proportional to the Killing form when $\mathfrak{g}$ is simple. We can now say something about when the proportionality constant is zero.

**Exercise 84.** Use Proposition 5.3 to show that if $\mathfrak{g}$ is simple and $\pi$ is irreducible and finite-dimensional, then the trace form $\kappa_\pi$ of (4.20) is degenerate if and only if $\pi$ is trivial. ▼

On a somewhat different note, the fact that $\mathsf{Q}$ is a sublattice of $\mathsf{P}$ leads those of the algebraic persuasion to ask about the quotient lattice $\mathsf{P}/\mathsf{Q}$ (the quotient is as an abelian group). This turns out to be quite interesting! In particular, the number of elements in the quotient serves as a measure of the relative sizes of the root and weight lattices. We can either measure this directly by counting integral points, or we can realise $|\mathsf{P}/\mathsf{Q}|$ as the ratio of the volumes of the parallelepipeds defined by the basis vectors.

**Example 50.** The root lattice of $\mathfrak{sl}(2)$ is spanned by the simple root $\alpha_1 = (2)$ and the weight lattice is spanned by the fundamental weight $\omega_1 = (1)$. We may therefore identify $\mathsf{Q}$ and $\mathsf{P}$ with $2\mathbb{Z}$ and $\mathbb{Z}$, respectively. It's pretty clear that $|\mathsf{P}/\mathsf{Q}| = 2$.

Similarly, the picture below shows that the parallelepiped defined by the simple roots of $\mathfrak{sl}(3)$ has three times the area of that defined by its fundamental weights, *ie.* $|\mathsf{P}/\mathsf{Q}| = 3$.




▲

**Exercise 85.** For $\mathfrak{sp}(4)$, $\mathfrak{so}(4)$ and $\mathfrak{g}_2$, compute the order of $\mathsf{P}/\mathsf{Q}$ by measuring the areas of the simple root and fundamental weight parallelepipeds (with diagrams). Compare your results, as well as those for $\mathfrak{sl}(2)$ and $\mathfrak{sl}(3)$, with the determinants of the corresponding Cartan matrices. Explain any coincidences that you observe. ▼

## 5.2. **Universal enveloping algebras**

In this section, we'll start developing some of the fundamental tools of representation theory that we need, the first of these being the notion of a universal enveloping algebra of a Lie algebra $\mathfrak{g}$. If you have studied group representation theory, then this associative algebra is to $\mathfrak{g}$ as the group algebra is to a group — in particular, the representation theories of $\mathfrak{g}$ and its universal enveloping algebra are identical. In what follows, we shall temporarily suspend our automatic assumption that $\mathfrak{g}$ is semisimple and finite-dimensional.

The universal enveloping algebra of $\mathfrak{g}$ is an associative algebra that is best motivated by thinking about the matrices (or linear operators) that one works with in a representation. Unlike Lie algebra elements, you can multiply matrices and this multiplication makes the space of endomorphisms of a module $V$ into an associative algebra. From an associative algebra, one can always define a Lie algebra by taking the Lie bracket to be the commutator $[A, B] = AB - BA$, but then we lose the ability to multiply elements. The universal enveloping algebra is then an associative algebra, constructed from a given Lie algebra, in which the multiplication is consistent with the Lie bracket being a commutator, just as we require in a representation. This is the motivation — passing to the universal enveloping algebra does not change the representation theory, but we gain a more powerful algebraic structure (associative multiplication).

We want to be able to multiply elements of a Lie algebra $\mathfrak{g}$: How do we do that, abstractly, whilst maintaining all the usual properties that a multiplication should have (distributivity, associativity, *etc.*)? This is what tensor products are for! Regarding $\mathfrak{g}$ as a vector space (for now), we let

$$\mathsf{T}^n(\mathfrak{g}) = \underbrace{\mathfrak{g} \otimes \cdots \otimes \mathfrak{g}}_{n>0 \text{ times}} \quad \text{and} \quad \mathsf{T}^0(\mathfrak{g}) = \mathbb{C}. \tag{5.10}$$

We then form the direct sum

$$\mathsf{T}(\mathfrak{g}) = \bigoplus_{n=0}^{\infty} \mathsf{T}^n(\mathfrak{g}). \tag{5.11}$$

We mention that the elements of an infinite direct sum are defined to be linear combinations of vectors from the direct summands where only **finitely many** of the coefficients are non-zero. Here, each summand $\mathsf{T}^n(\mathfrak{g})$ consists of linear combinations of products of $n$ elements of $\mathfrak{g}$. Technically, the addition of two elements of different degrees $n$ should be denoted by $\oplus$, but we will identify this with the addition $+$ of each $\mathsf{T}^n(\mathfrak{g})$, for notational convenience. We can now make the vector space $\mathsf{T}(\mathfrak{g})$ into an associative algebra with unit $1 \in \mathbb{C} = \mathsf{T}^0(\mathfrak{g})$ upon defining

$$(x_1 \otimes \cdots \otimes x_m)(y_1 \otimes \cdots \otimes y_n) = x_1 \otimes \cdots \otimes x_m \otimes y_1 \otimes \cdots \otimes y_n, \tag{5.12}$$

for all $x_1, \dots, x_m, y_1, \dots, y_n \in \mathfrak{g}$. This algebra $\mathsf{T}(\mathfrak{g})$ is called the *tensor algebra* of $\mathfrak{g}$.

To get the universal enveloping algebra, we not only have to be able to multiply, but we also need to take into account the fact that the Lie bracket is replaced by the matrix commutator in representations. In the abstract setting, this is done by considering the following two-sided ideal of $\mathsf{T}(\mathfrak{g})$:

$$\mathsf{J} = \{U I_{x,y} U' : x, y \in \mathfrak{g} \text{ and } U, U' \in \mathsf{T}(\mathfrak{g})\}, \quad \text{where } I_{x,y} = x \otimes y - y \otimes x - [x, y]. \tag{5.13}$$

Setting this ideal to zero will therefore force the Lie bracket $[x, y]$ to coincide with the (abstract tensor algebra) commutator $x \otimes y - y \otimes x$. We therefore define the *universal enveloping algebra* of $\mathfrak{g}$ to be the quotient

$$\mathsf{U}(\mathfrak{g}) = \frac{\mathsf{T}(\mathfrak{g})}{\mathsf{J}}. \tag{5.14}$$

$\mathsf{U}(\mathfrak{g})$ is thus an associative algebra. Because $\mathsf{J} \cap \mathsf{T}^0(\mathfrak{g})$ is obviously 0, the constants survive the quotienting and so the unit $1 \in \mathsf{T}(\mathfrak{g})$ determines a unit for $\mathsf{U}(\mathfrak{g})$ (which we shall denote by $\mathbb{1} = \overline{1}$).

**Example 51.** When $\mathfrak{g}$ is an abelian Lie algebra, the ideal $\mathsf{J}$ is generated by elements of the form $x \otimes y - y \otimes x$, where $x, y \in \mathfrak{g}$. The universal enveloping algebra $\mathsf{U}(\mathfrak{g})$ may therefore be identified with the algebra of polynomials in the elements of $\mathfrak{g}$. More precisely, if $\{x_i : i \in I\}$ is a basis of $\mathfrak{g}$, then $\mathsf{U}(\mathfrak{g}) \simeq \mathbb{C}[x_i : i \in I]$. ▲

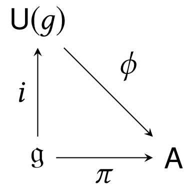
It is, unfortunately, not clear whether $\mathsf{J} \cap \mathsf{T}^1(\mathfrak{g}) = \{0\}$, *ie.* if any non-zero elements of $\mathsf{T}^1(\mathfrak{g}) = \mathfrak{g}$ are set to zero in $\mathsf{U}(\mathfrak{g})$. If there were, then this would mean that the natural composition

$$i \colon \mathfrak{g} = \mathsf{T}^1(\mathfrak{g}) \hookrightarrow \mathsf{T}(\mathfrak{g}) \twoheadrightarrow \mathsf{U}(\mathfrak{g}) \tag{5.15}$$

would not be an inclusion. It seems difficult to imagine how one could obtain an element of $\mathsf{T}^1(\mathfrak{g})$ from the $I_{x,y}$; the proof that one cannot is therefore correspondingly unintuitive. However, the result is very important and is usually obtained as a consequence of an even more important result known as the Poincaré–Birkhoff–Witt theorem. Before coming to this, let us see why $\mathsf{U}(\mathfrak{g})$ is referred to as being **universal**.

**Proposition 5.4.** *Let* $\mathsf{A}$ *be a unital associative algebra* (*over* $\mathbb{C}$) *and let* $\pi \colon \mathfrak{g} \to \mathsf{A}$ *be a homomorphism of Lie algebras* (*taking the Lie bracket of* $\mathsf{A}$ *to be the commutator*). *Then, there is a unique homomorphism* $\phi \colon \mathsf{U}(\mathfrak{g}) \to \mathsf{A}$ *of unital associative algebras such that*

$$\phi \circ i = \pi. \tag{5.16}$$




*In other words, all morphisms from* $\mathfrak{g}$ *to* $\mathsf{A}$ *must factor uniquely through* $\mathsf{U}(\mathfrak{g})$.

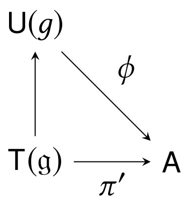
*Proof.* First, extend $\pi$ to a homomorphism $\pi' \colon \mathsf{T}(\mathfrak{g}) \to \mathsf{A}$ of unital associative algebras by setting

$$\pi'(x_1 \otimes \cdots \otimes x_n) = \pi(x_1) \dots \pi(x_n), \quad \text{for all } x_1, \dots, x_n \in \mathfrak{g}, \tag{5.17}$$

and extending by linearity. Since

$$\pi'(x \otimes y - y \otimes x - [x, y]) = \pi(x)\pi(y) - \pi(y)\pi(x) - \pi([x, y]) = 0, \tag{5.18}$$

because $\pi$ is a Lie algebra homomorphism, it follows that $\pi'(\mathsf{J}) = 0$. $\pi'$ therefore induces a homomorphism $\phi \colon \mathsf{U}(\mathfrak{g}) \to \mathsf{A}$ such that




commutes. Restricting the domain of the horizontal and vertical maps to $\mathfrak{g}$ gives the desired commutative diagram.

It remains to prove the uniqueness of $\phi$. So suppose that $\widetilde{\phi} \circ i = \pi$, for some homomorphism $\widetilde{\phi} \colon \mathsf{U}(g) \to \mathsf{A}$ of unital associative algebras. As $\mathfrak{g} = \mathsf{T}^1(\mathfrak{g})$ generates $\mathsf{T}(\mathfrak{g})$, it follows that $i(\mathfrak{g})$ generates $\mathsf{U}(\mathfrak{g})$. Thus, the actions of $\phi$ and $\widetilde{\phi}$ on $\mathsf{U}(\mathfrak{g})$ are completely determined by their restrictions to $i(\mathfrak{g})$. But these restrictions are identical, $\widetilde{\phi} \circ i = \pi = \phi \circ i$, so indeed we have $\phi = \widetilde{\phi}$. ■

We remark that a homomorphism of unital associative algebras must not only respect the algebraic operations, but also send the unit to the unit. A representation of a unital associative algebra is, of course, just a homomorphism of unital associative algebras into $\operatorname{End}(V)$, for some vector space $V$.

**Exercise 86.** Use Proposition 5.4 to prove that the representations of $\mathfrak{g}$ and $\mathsf{U}(\mathfrak{g})$ are in bijection by:

**(a)** Showing that any representation $\pi$ of $\mathfrak{g}$ on $V$ induces a representation $\phi$ of $\mathsf{U}(\mathfrak{g})$ on $V$.
**(b)** Showing that any representation $\phi$ of $\mathsf{U}(\mathfrak{g})$ on $V$ induces a representation $\pi$ of $\mathfrak{g}$ on $V$.
**(c)** Showing that these two constructions are inverse to one another. ▼

We now turn to the all-important Poincaré–Birkhoff–Witt theorem and its somewhat subtle proof.

**Theorem 5.5** (Poincaré–Birkhoff–Witt)**.** *The canonical map from* $\mathfrak{g}$ *into* $\mathsf{U}(\mathfrak{g})$ *is an inclusion. Moreover, if* $\{x_i : i \in I\}$ *is an ordered basis of* $\mathfrak{g}$ (*so the index set* $I$ *admits a total order*), *then the* (*equivalence classes of the*) *monomials*

$$x_{i_1} \otimes x_{i_2} \otimes \cdots \otimes x_{i_n}, \quad \text{for all } n \in \mathbb{Z}_{\geqslant 0} \text{ and } i_1 \leqslant i_2 \leqslant \cdots \leqslant i_n \text{ in } I, \tag{5.19}$$

*form a basis of* $\mathsf{U}(\mathfrak{g})$, *called a* Poincaré–Birkhoff–Witt basis.

*Proof.* Proving that the monomials (5.19) span $\mathsf{U}(\mathfrak{g})$ is not hard — one just repeatedly applies the commutation rules

$$x_i \otimes x_j = x_j \otimes x_i + [x_i, x_j], \quad \text{for all } i, j \in I, \tag{5.20}$$

to rewrite any other monomial as a linear combination of monomials with the desired order. Because $[x_i, x_j] \in \mathfrak{g}$ has (tensor) degree 1, whilst $x_i \otimes x_j$ and $x_j \otimes x_i$ have degree 2, each application results in one term in which the $x_i$ and $x_j$ have been correctly ordered and another term of strictly lower degree which may be dealt with inductively.

The hard part is proving that the monomials (5.19) are linearly independent. The inclusion property then follows easily. To show the linear independence, we shall construct a linear map $\overline{\psi}\colon \mathsf{U}(\mathfrak{g}) \to \mathbb{C}[z_i : i \in I]$ such that the images of the monomials (5.19) are obviously linearly independent. This will be constructed by exhibiting a linear map $\psi\colon \mathsf{T}(\mathfrak{g}) \to \mathbb{C}[z_i : i \in I]$ such that the images of the monomials (5.19) are linearly independent **and** the two-sided ideal $\mathsf{J}$ of $\mathsf{T}(\mathfrak{g})$ is mapped to 0. Because the linear independence of the images implies that of the monomials (5.19), the proof will therefore be complete once the existence of $\psi$ has been established.

We claim that a linear map $\psi$ is defined by the following two requirements:

$$\psi(x_{i_1} \otimes \cdots \otimes x_{i_n}) = z_{i_1} \cdots z_{i_n}, \tag{5.21a}$$

for all $n \in \mathbb{Z}_{\geqslant 0}$ and $i_1 \leqslant \cdots \leqslant i_n$ in $I$, and

$$\begin{aligned} \psi(x_{i_1} \otimes \cdots \otimes x_{i_m} \otimes x_{i_{m+1}} \otimes \cdots \otimes x_{i_n}) &= \psi(x_{i_1} \otimes \cdots \otimes x_{i_{m+1}} \otimes x_{i_m} \otimes \cdots \otimes x_{i_n}) \\ &\quad + \psi(x_{i_1} \otimes \cdots \otimes [x_{i_m}, x_{i_{m+1}}] \otimes \cdots \otimes x_{i_n}), \end{aligned} \tag{5.21b}$$

for all $n \in \mathbb{Z}_{\geqslant 2}$ and $1 \leqslant m < n$. Clearly, (5.21a) defines $\psi$ on the correctly ordered monomials (5.19) and (5.21b) is designed to define $\psi$ inductively on the remaining monomials by swapping $x_{i_m}$ and $x_{i_{m+1}}$ in they are incorrectly ordered. Moreover, (5.21b) also guarantees that $\psi$ annihilates the ideal J.

It therefore only remains to check that the linear map $\psi$ is well-defined: one can reduce an incorrectly ordered monomial to a correctly ordered one in many ways, all of which must result in the same value of $\psi$. We check that $\psi$ is well-defined using a double induction argument. Note first that for $n \leqslant 2$, the action of $\psi$ is well-defined because there is at most one way to reorder a monomial. Moreover, if we define the index of a monomial $x_{i_1} \otimes \cdots \otimes x_{i_n}$ to be the number of $m$, $1 \leqslant m < n$, such that $i_m > i_{m+1}$, then the action of $\psi$ on every index-0 monomial is well-defined, by (5.21a).

These are the base cases — the inductive step requires us to assume that the action of $\psi$ on monomials of degree less than $n$, and on monomials of degree $n$ and index less than $k$, is well-defined. We then need to prove that the action on monomials of degree $n$ and index $k$ is well-defined. This is the purpose of Exercise 87 below (mwouhaha!). ■

**Exercise 87.** Complete the proof of Theorem 5.5 by showing inductively that if $m$ and $m'$, where $1 \leqslant m < m' < n$, satisfy $i_m > i_{m+1}$ and $i_{m'} > i_{m'+1}$, then applying (5.21b) twice to

$$T = \psi(x_{i_1} \otimes \cdots \otimes x_{i_m} \otimes x_{i_{m+1}} \otimes \cdots \otimes x_{i_{m'}} \otimes x_{i_{m'+1}} \otimes \cdots \otimes x_{i_n}) \tag{5.22}$$

results in a well-defined action of $\psi$, **no matter which pair is swapped first**. You should consider the case $m' = m + 1$ separately. ▼

It follows that the universal enveloping algebra of any (non-zero) Lie algebra is infinite-dimensional. We remark that in all practical circumstances, we can (and will!) omit the tensor product symbols when working in $\mathsf{U}(\mathfrak{g})$ and replace the Lie bracket by the commutator, as one does in a representation.

**Example 52.** If we order the standard basis of $\mathfrak{sl}(2)$ as $\{f, h, e\}$, then the corresponding Poincaré–Birkhoff–Witt basis of $\mathsf{U}(\mathfrak{g})$ is

$$\{f^i h^j e^k : i, j, k \in \mathbb{Z}_{\geqslant 0}\}. \tag{5.23}$$

Of course, different orderings of $\{f, h, e\}$ lead to different Poincaré–Birkhoff–Witt bases.

Similarly, taking $\{f_\theta, f_{\alpha_2}, f_{\alpha_1}, h_{\alpha_1}, h_{\alpha_2}, e_{\alpha_1}, e_{\alpha_2}, e_\theta\}$ as an ordered basis of $\mathfrak{sl}(3)$ (we recall that $\theta = \alpha_1 + \alpha_2$ is the highest root of $\mathfrak{sl}(3)$) results in the Poincaré–Birkhoff–Witt basis

$$\{f_\theta^i f_{\alpha_2}^j f_{\alpha_1}^k h_{\alpha_1}^\ell h_{\alpha_2}^m e_{\alpha_1}^n e_{\alpha_2}^p e_\theta^q : i, j, k, \ell, m, n, p, q \in \mathbb{Z}_{\geqslant 0}\}. \tag{5.24}$$ ▲

**Exercise 88.** Show that if $\mathfrak{g}$ has a triangular decomposition $\mathfrak{g} = \mathfrak{g}_- \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_+$, *cf.* Equation (4.22), then its universal enveloping algebra decomposes as

$$\mathsf{U}(\mathfrak{g}) = \mathsf{U}(\mathfrak{g}_-) \otimes \mathsf{U}(\mathfrak{g}_0) \otimes \mathsf{U}(\mathfrak{g}_+). \tag{5.25}$$ ▼

**Exercise 89.** Prove that $\mathsf{U}(\mathfrak{g})$ has no zero-divisors, *ie.* that if $U, U' \in \mathsf{U}(\mathfrak{g})$ are both non-zero, then $UU' \neq 0$. ▼

## 5.3. **Highest-weight modules**

We now resume assuming that the Lie algebra $\mathfrak{g}$ is semisimple and finite-dimensional, combining the weight module theory that we have seen in Section 5.1 with the Poincaré–Birkhoff–Witt theorem (Theorem 5.5).

An incredibly important concept is that of a highest weight. Because a finite-dimensional $\mathfrak{g}$-module may only have a finite number of weights, there must exist maximal weights (for any sensible notion of "maximal"). With respect to a given set $\Pi$ of simple roots $\alpha_i$, $i = 1, \dots, \operatorname{rank} \mathfrak{g}$, we will take maximal to mean the following: A weight $\lambda$ of a $\mathfrak{g}$-module $V$ is said to be a *highest weight* of $V$ if there is a corresponding weight vector $v_\lambda \in V$, called a *highest-weight vector*, that is annihilated by the action of the simple root vectors $e_{\alpha_i}$. It therefore follows that every finite-dimensional $\mathfrak{g}$-module has a highest weight.

Since the positive root vectors $e_\alpha$, $\alpha \in \Delta_+$, may all be expressed in terms of (nested) Lie brackets of the $e_{\alpha_i}$ (this follows from Proposition 4.13 and the algorithm of Section 4.7 for constructing $\Delta_+$ from $\Pi$), it follows that highest-weight vectors are characterised by

$$H v_\lambda = \lambda(H) v_\lambda \quad \text{and} \quad e_\alpha v_\lambda = 0, \qquad \text{for all } H \in \mathfrak{g}_0 \text{ and } \alpha \in \Delta_+. \tag{5.26}$$

We remark that a sufficient condition for $\lambda$ to be a highest weight, for a given $\mathfrak{g}$-module $V$, is that $\lambda + \alpha_i$ is not a weight of the module for any $i = 1, \dots, \operatorname{rank} \mathfrak{g}$. However, this need not be necessary if $V$ is reducible.

**Exercise 90.** Strengthen the proof of Proposition 5.3 to conclude that every highest weight of a finite-dimensional weight module belongs to $\mathsf{P}_{\geqslant}$. ▼

We remark that weights $\lambda$ with non-negative integer Dynkin labels, *ie.* $\lambda \in \mathsf{P}_{\geqslant}$, are said to be *dominant integral*.

Recall that we have a particularly nice basis of $\mathfrak{g}$ consisting of the root vectors $e_\alpha$ and $f_\alpha$, $\alpha \in \Delta_+$, and the simple coroot vectors $\alpha_i^\vee = h_{\alpha_i}$, $i = 1, \dots, \operatorname{rank} \mathfrak{g}$. The Poincaré–Birkhoff–Witt theorem (Theorem 5.5) has some important things to say when we order this basis appropriately.

**Proposition 5.6.** *An irreducible module with a highest weight has a unique highest weight and the corresponding highest-weight vector is unique up to scalar multiples.*

*Proof.* Let $V$ be irreducible and let $\lambda$ and $\mu$ be highest weights of $V$ with respective highest-weight vectors $v$ and $w$. By irreducibility, the submodule generated by $v \in V$ is $V$ itself, hence $w \in V$ may be written as a linear combination of Poincaré–Birkhoff–Witt basis elements acting on $v$. We choose to order these basis elements so that, for each $\alpha \in \Delta_+$, the $f_\alpha$ always appear to the left of the $h_{\alpha_i}$, which in turn always appear to the left of the $e_\alpha$.

The highest weight property of $v$ then implies that any basis element involving a non-zero power of an $e_\alpha$ will annihilate $v$ and any basis element involving a non-zero power of an $h_{\alpha_i}$ will return a multiple of $v$. It follows that $w$ can be obtained from $v$ by acting with a linear combination of Poincaré–Birkhoff–Witt basis elements which only involve the $f_\alpha$. Thus, the weight difference $\lambda - \mu$ must be a **non-negative integer** linear combination of the simple roots: $\lambda - \mu \in \mathsf{Q}_{\geqslant}$.

Now exchange the roles of $v$ and $w$ in this argument. Then, we conclude that $\mu - \lambda \in \mathsf{Q}_{\geqslant}$ and so arrive at $\lambda - \mu \in \mathsf{Q}_{\geqslant} \cap (-\mathsf{Q}_{\geqslant}) = 0$, *ie.* $\lambda = \mu$. But, this implies that $v$ and $w$ may be obtained from one another by acting with a linear combination of Poincaré–Birkhoff–Witt basis elements involving only the $h_{\alpha_i}$. Since $v$ and $w$ are weight vectors, they are thereby proportional. ∎

Proposition 5.1 tells us that all irreducible finite-dimensional modules are weight modules and Exercise 90 tells us that the highest weights of such modules are dominant integral (belong to $\mathsf{P}_{\geqslant}$). As all finite-dimensional modules must have a highest weight, we now conclude that every irreducible finite-dimensional module has **exactly** one highest weight. We call any module that is generated by a single highest-weight vector a *highest-weight module*.

**Corollary 5.7.** *Every finite-dimensional irreducible* $\mathfrak{g}$*-module is a highest-weight module whose highest weight is dominant integral.*

This innocuous corollary is in fact one direction of the classification result for finite-dimensional irreducible $\mathfrak{g}$-modules. It turns out that they are, up to isomorphism, in bijection with the set $\mathsf{P}_{\geqslant}$ of dominant integral weights. However, it will take quite some effort still before we can establish this fundamental fact.

For now, we can leverage this knowledge to algorithmically determine *all* the weights of a finite-dimensional, irreducible module, given only its highest weight. The key here, as it was with analysing the roots of semisimple Lie algebras, is our knowledge of $\mathfrak{sl}(2)$ representation theory.

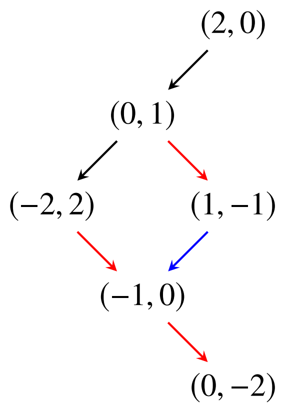
An algorithm for computing the weights of an irreducible, finite-dimensional $\mathfrak{g}$-module is then:

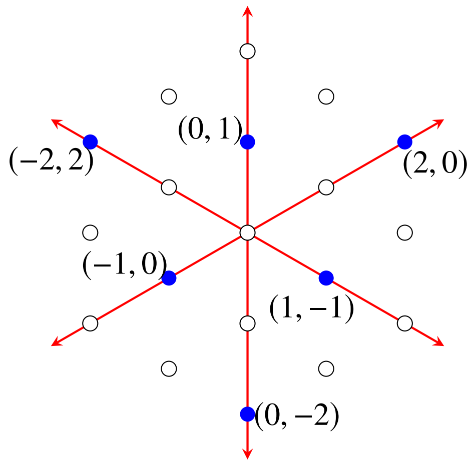
**(a)** Start with the highest weight $\lambda = (\lambda_1, \dots, \lambda_r)$, expressed in terms of its Dynkin labels (recall (5.4) for the definition).
**(b)** Given any weight $\mu = (\mu_1, \dots, \mu_r)$ already found, more may be found if any of the $\mu_i$ are positive: $\mu_i > 0$ implies that $\mu - k\alpha_i$ is a weight, for all $k = 0, 1, \dots, \mu_i$.
**(c)** Repeat the previous step until all the new weights found have non-positive Dynkin labels.

This algorithm must terminate as there are only a finite number of weights to find. The fact that $\mu - k\alpha_i$ is a weight follows from $\mathfrak{sl}(2)$ representation theory: a positive Dynkin label $\mu_i$ implies that we may apply $f_{\alpha_i}$ $\mu_i$ times to get new weight vectors. It is clear that all of the weights generated by this algorithm must be weights of the $\mathfrak{g}$-module. To see that it generates all the weights, we note that every $f_\alpha$ may be written as a nested Lie bracket of the $f_{\alpha_i}$, so it is enough to consider the action of the latter on weight vectors. Moreover, acting exhaustively on the highest-weight vector generates a non-zero submodule, hence it must be the entire module by irreducibility.

It is important to note, however, that this algorithm only produces the weights. It does not tell us their *multiplicities*, meaning the dimensions of the corresponding weight spaces. It therefore cannot be used to compute the dimension of the $\mathfrak{g}$-module. In particular, it doesn't even tell us which dominant integral weights $\lambda$ actually lead to finite-dimensional modules!

**Example 53.** Let $\mathfrak{g} = \mathfrak{sl}(3)$ and consider the irreducible module $V$ of highest weight $\lambda = 2\omega_1 = (2, 0)$. Recalling that the Dynkin labels of the simple roots are the columns of the Cartan matrix, we have $\alpha_1 = (2, -1)$ and $\alpha_2 = (-1, 2)$. As $\lambda_1 = 2$, we know that $\lambda - \alpha_1 = (0, 1)$ and $\lambda - 2\alpha_1 = (-2, 2)$ are both weights of $V$. As the second Dynkin label of both these weights is positive, it follows that $\lambda - \alpha_1 - \alpha_2 = (1, -1)$ is a weight of $V$ as are $\lambda - 2\alpha_1 - \alpha_2 = (-1, 0)$ and $\lambda - 2\alpha_1 - 2\alpha_2 = (0, -2)$. These last two weights have no positive Dynkin labels, they generate nothing new. Moreover, $(1, -1)$ only generates $(-1, 0)$ which we already have. These are therefore all the weights. We draw them as they were determined and in blue in a (once again beautifully symmetric) weight diagram.




The arrows pointing to the left in the first picture represent subtracting $\alpha_1$, the arrows pointing right are for subtracting $\alpha_2$. They are colour-coded according to the order they appear in the algorithm (black, then red, then blue).

As you did Exercise 83, you will notice that the weight diagram on the right contains that of the dual of the defining module. This does not mean that the latter module is a submodule, or even a quotient module, of the irreducible module with highest weight $(2, 0)$ — irreducibility tells us that it cannot be! ▲

**Exercise 91.** Draw the weight diagrams for the irreducible $\mathfrak{g}$-module of highest weight $\omega_2 = (0, 1)$, for $\mathfrak{g} = \mathfrak{sp}(4)$ and $\mathfrak{g}_2$. Find the weights of the fundamental $\mathfrak{so}(7)$-modules, *ie.* those whose highest weights are $\omega_1, \omega_2$ and $\omega_3$, and (**optional!**) draw the weight diagrams if you are a connoisseur of the third dimension. (Make sure that you use the same ordering for the simple roots as in Examples 45 and 46 and Theorem 4.23.) ▼

**Exercise 92.** Show that the tensor square of any $\mathfrak{g}$-module $V$ decomposes into $\mathfrak{g}$-modules as $V \otimes V \simeq S^2V \oplus \wedge^2 V$, where

$$\begin{aligned} S^2V &= \operatorname{span}\{v \otimes w + w \otimes v : v, w \in V\}, \\ \text{and} \quad \wedge^2 V &= \operatorname{span}\{v \otimes w - w \otimes v : v, w \in V\}. \end{aligned} \tag{5.27}$$

When $V$ is the defining module of $\mathfrak{sl}(3)$, show that $S^2V$ has dimension 6 and highest weight $(2, 0)$. Can we therefore conclude that the module in Example 53 is (isomorphic to) $S^2V$? If so, why? If not, why not? ▼

## 5.4. **Verma modules**

We have developed enough theory now to realise that highest-weight modules are the key to the finite-dimensional representation theory of a given semisimple Lie algebra $\mathfrak{g}$. However, we have not yet managed to prove any general existence theorems. In particular, we would like to answer the following questions:

- Given $\lambda \in \mathfrak{g}_0^*$, is there a $\mathfrak{g}$-module with $\lambda$ as a highest weight?

- Given $\lambda \in \mathsf{P}_{\geqslant}$, is there a finite-dimensional $\mathfrak{g}$-module with $\lambda$ as a highest weight?
- In both cases, can we take the $\mathfrak{g}$-module to be irreducible?
- Assuming existence, when are these $\mathfrak{g}$-modules unique (up to isomorphism)?

In this section, we will answer the first of these questions by explicitly constructing a highest-weight module for each highest weight $\lambda \in \mathfrak{g}_0^*$. Much of what follows will also hold for arbitrary Lie algebras with a triangular decomposition. If you're interested in this level of generality, see the excellent book by Moody and Pianzola.

Recall that $\mathfrak{g}$ has a basis consisting of the simple coroots and the root vectors. The Poincaré–Birkhoff–Witt theorem (Theorem 5.5) then shows that there is a basis of the universal enveloping algebra $\mathsf{U}(\mathfrak{g})$ in which each basis element has the form $U = U_- U_0 U_+$, where $U_0$ is a product of simple coroots and $U_+$ ($U_-$) is a product of positive (negative) root vectors. Each of these products may be empty, of course, the meaning of which is that we may have any of $U_0$ or $U_\pm$ being $\mathbb{1}$.

If one of these basis elements $U$ acts on a highest-weight vector $v$ of weight $\lambda$, then the result will be zero unless $U_+ = \mathbb{1}$, because $e_\alpha v = 0$ for all $\alpha \in \Delta_+$. If $U_+ = \mathbb{1}$, then the $U_0$-part of the basis element acts as a multiple of the identity because $h_{\alpha_i} v = \lambda_i v$. It follows that $Uv$ will always be proportional to $U_- v$, hence that the set of all $U_- v$, where $U_-$ runs over a basis of the universal enveloping algebra $\mathsf{U}(\mathfrak{g}_-)$, will span the same space that the $Uv$ do. We will show that this space is a $\mathfrak{g}$-module and, moreover, that we can construct such a $\mathfrak{g}$-module where this spanning set is a basis.

We pause to remark that a highest-weight vector $v$ is also known as a *vacuum*, or (sometimes) a *ground state*, in physics because it often has the property of having the minimal energy (or is extremal for some other physical quantity like spin, *cf.* Section 3.3). The fact that $v$ is defined to be annihilated by the positive root vectors leads to the latter being termed *annihilation operators*. The negative root vectors are then the *creation operators* because they usually act non-trivially on $v$ and so create new weight vectors. If the vacuum has minimal energy, then these new vectors have higher energies and are therefore referred to as *excited states*. The Cartan elements, on the other hand, are often called *zero modes* or, in quantum physics, the *quantum observables* because their eigenvalues are these energies and/or some other physically measurable quantities.

In any case, to show that $\mathsf{U}(\mathfrak{g}_-)v$ is a $\mathfrak{g}$-module, we must first address the subtlety that arises because to construct this $\mathfrak{g}$-module, we start from a highest-weight vector $v$ which, by definition, must belong to some $\mathfrak{g}$-module. We can get around this in two distinct ways, though of course they are equivalent at the end of the day. The first is concrete and "top-down", though it takes a bit of setting up; the second is abstract and "bottom-up", but requires a fair whack of mathematical maturity.

**Exercise 93.** Show that $\mathsf{U}(\mathfrak{g})$ becomes a $\mathfrak{g}$-module when the action is defined to be by left-multiplication: $x \cdot U = xU$, for all $x \in \mathfrak{g}$ and $U \in \mathsf{U}(\mathfrak{g})$. Now show that right-multiplication by any $V \in \mathsf{U}(\mathfrak{g})$ defines a $\mathfrak{g}$-module endomorphism $\phi_V \colon \mathsf{U}(\mathfrak{g}) \to \mathsf{U}(\mathfrak{g})$, *ie.* define $\phi_V$ by $\phi_V(U) = UV$. ▼

It follows that the images $\operatorname{im}\phi_V$ are submodules (left ideals) of $\mathsf{U}(\mathfrak{g})$, for any $V \in \mathsf{U}(\mathfrak{g})$. Obviously, $\operatorname{im}\phi_0 = 0$ and $\operatorname{im}\phi_{\mathbb{1}} = \mathsf{U}(\mathfrak{g})$.

**Proposition 5.8.** *Given any $\lambda \in \mathfrak{g}_0^*$, the quotient*

$$\mathcal{V}_\lambda = \frac{\mathsf{U}(\mathfrak{g})}{\mathsf{I}_\lambda}, \quad \text{where } \mathsf{I}_\lambda = \sum_{\alpha\in\Delta_+} \operatorname{im}\phi_{e_\alpha} + \sum_{i=1}^{\operatorname{rank}\mathfrak{g}} \operatorname{im}\phi_{h_{\alpha_i}-\lambda_i\mathbb{1}}, \tag{5.28}$$

*is a highest-weight module over $\mathfrak{g}$ with highest weight $\lambda$.*

*Proof.* Consider the equivalence class $\overline{\mathbb{1}} \in \mathcal{V}_\lambda$ of the unit $\mathbb{1} \in \mathsf{U}(\mathfrak{g})$. The Poincaré–Birkhoff–Witt theorem ensures that $\mathbb{1}$ is not in $\operatorname{im}\phi_V$, whenever $V$ is not a multiple of $\mathbb{1}$, hence $\mathbb{1}$ is not in the sum of any such images. In particular, $\overline{\mathbb{1}} \neq 0$ in $\mathcal{V}_\lambda$ and explicit computation shows that $\overline{\mathbb{1}}$ is the desired highest-weight vector of $\mathcal{V}_\lambda$:

$$\begin{aligned} e_\alpha\overline{\mathbb{1}} &= \overline{e_\alpha} = 0, && \text{for all } \alpha \in \Delta_+, && \text{as } e_\alpha = \phi_{e_\alpha}(\mathbb{1}) \in \operatorname{im}\phi_{e_\alpha}, \\ h_{\alpha_i}\overline{\mathbb{1}} &= \overline{h_{\alpha_i}} = \overline{\lambda_i\mathbb{1}} = \lambda_i\overline{\mathbb{1}}, && \text{for all } i = 1,\dots,\operatorname{rank}\mathfrak{g}, && \text{as } h_{\alpha_i} - \lambda_i\mathbb{1} \in \operatorname{im}\phi_{h_{\alpha_i}-\lambda_i\mathbb{1}}. \end{aligned} \tag{5.29}$$

As $\overline{\mathbb{1}}$ obviously generates $\mathcal{V}_\lambda$, this completes the proof. ■

The highest-weight module $\mathcal{V}_\lambda$ constructed in this proposition is called the *Verma module* of highest weight $\lambda$. It answers our first existence question in the affirmative. The reason why we take the (sum of the) images of the $\phi_{e_\alpha}$ and $\phi_{h_{\alpha_i}-\lambda_i\mathbb{1}}$ is that these images are spanned by elements of $\mathsf{U}(\mathfrak{g})$ of the form $Ue_\alpha$ and $U(h_{\alpha_i} - \lambda_i\mathbb{1})$, respectively, both of which annihilate a highest-weight vector.

The second, more abstract, construction requires the notion of an induced module. This concept takes a module over a subalgebra and lifts it to a module over the full algebra in a canonical way. The subalgebra that we will use is the universal enveloping algebra $\mathsf{U}(\mathfrak{b}) \subset \mathsf{U}(\mathfrak{g})$, where $\mathfrak{b} = \mathfrak{g}_0 \oplus \mathfrak{g}_+$ is the *Borel subalgebra* associated to the triangular decomposition $\mathfrak{g} = \mathfrak{g}_- \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_+$.

We start with the one-dimensional $\mathfrak{g}_0$-module $\mathbb{C}_\lambda = \operatorname{span}\{v\}$, where $v$ is a weight vector of weight $\lambda \in \mathfrak{g}_0^*$: $h_{\alpha_i}v = \lambda_i v$ for all $i = 1,\dots,\operatorname{rank}\mathfrak{g}$. This is extended to a $\mathfrak{b}$-module, hence a $\mathsf{U}(\mathfrak{b})$-module (*cf.* Exercise 86), by requiring that the elements of $\mathfrak{g}_+$ act as zero: $e_\alpha v = 0$, for all $\alpha \in \Delta_+$. This indeed gives $\mathbb{C}_\lambda$ a $\mathfrak{b}$-module structure because $[\mathfrak{g}_0, \mathfrak{g}_+] \subseteq \mathfrak{g}_+$.

Now comes the induced module construction. We lift this one-dimensional $\mathsf{U}(\mathfrak{b})$-module to a $\mathsf{U}(\mathfrak{g})$-module, hence a $\mathfrak{g}$-module, $\mathcal{V}'_\lambda$ by defining the latter as the tensor product

$$\mathcal{V}'_\lambda = \mathsf{U}(\mathfrak{g}) \otimes_{\mathsf{U}(\mathfrak{b})} \mathbb{C}_\lambda. \tag{5.30}$$

The symbol $\otimes_{\mathsf{U}(\mathfrak{b})}$ indicates that it is "permeable" to elements of $\mathsf{U}(\mathfrak{b})$, rather than just to scalars in $\mathbb{C}$ (*cf.* the complexifications introduced in Section 2.4):

$$UV \otimes_{\mathsf{U}(\mathfrak{b})} v = U \otimes_{\mathsf{U}(\mathfrak{b})} Vv, \quad \text{for all } U \in \mathsf{U}(\mathfrak{g}) \text{ and } V \in \mathsf{U}(\mathfrak{b}). \tag{5.31}$$

As a vector space, $\mathcal{V}'_\lambda$ may be identified with a quotient of the usual tensor product over $\mathbb{C}$:

$$\mathcal{V}'_\lambda \simeq \frac{\mathsf{U}(\mathfrak{g}) \otimes_{\mathbb{C}} \mathbb{C}_\lambda}{\operatorname{span}\{UV \otimes_{\mathbb{C}} v - U \otimes_{\mathbb{C}} Vv : U \in \mathsf{U}(\mathfrak{g}) \text{ and } V \in \mathsf{U}(\mathfrak{b})\}}. \tag{5.32}$$

The $\mathfrak{g}$- and $\mathsf{U}(\mathfrak{g})$-module structures of $\mathcal{V}'_\lambda$ are given by left-multiplication on the first factor: $U \cdot (U' \otimes_{\mathsf{U}(\mathfrak{b})} v) = (UU') \otimes_{\mathsf{U}(\mathfrak{b})} v$.

For experts, what's going on here is that $\mathbb{C}_\lambda$ is a **left** $\mathsf{U}(\mathfrak{b})$-module whilst we may regard $\mathsf{U}(\mathfrak{g})$ as a **right** $\mathsf{U}(\mathfrak{b})$-module under $U \cdot V = UV$, where $V \in \mathsf{U}(\mathfrak{b})$ acts on $U \in \mathsf{U}(\mathfrak{g})$ by right-multiplication. The tensor product $\mathsf{U}(\mathfrak{g}) \otimes_{\mathsf{U}(\mathfrak{b})} \mathbb{C}_\lambda$ is then a vector space. However, as $\mathsf{U}(\mathfrak{g})$ is also a **left** $\mathsf{U}(\mathfrak{g})$-module and this left action commutes with the right $\mathsf{U}(\mathfrak{b})$-action (so it is a $(\mathsf{U}(\mathfrak{g}), \mathsf{U}(\mathfrak{b}))$-**bimodule**), it follows that this tensor product vector space is also naturally a **left** $\mathsf{U}(\mathfrak{g})$-module. Phew!

**Example 54.** To make the induced module construction a little less abstract, consider $\mathfrak{g} = \mathfrak{sl}(2)$ with the triangular decomposition $\mathfrak{sl}(2)_+ = \operatorname{span}\{e\}$, $\mathfrak{sl}(2)_0 = \operatorname{span}\{h\}$ and $\mathfrak{sl}(2)_- = \operatorname{span}\{f\}$. Start from the **trivial** $\mathfrak{sl}(2)_0$-module $\mathbb{C}_0 = \operatorname{span}\{v\}$ on which $h$ acts as 0. Extend to a module over $\mathfrak{b} = \operatorname{span}\{e, h\}$ by letting $e$ also act as 0. Now induce to form the $\mathfrak{sl}(2)$-module $\mathcal{V}'_0 = \mathsf{U}(\mathfrak{sl}(2)) \otimes_{\mathsf{U}(\mathfrak{b})} \mathbb{C}_0$. Since $\mathsf{U}(\mathfrak{sl}(2))$ has a Poincaré–Birkhoff–Witt basis consisting of the monomials $f^i h^j e^k$, where $i, j, k \in \mathbb{Z}_{\geqslant 0}$, the module $\mathcal{V}'_0$ is spanned by elements of the form $f^i h^j e^k \otimes_{\mathsf{U}(\mathfrak{b})} v = f^i \otimes_{\mathsf{U}(\mathfrak{b})} h^j e^k v$. As these elements vanish if $j$ or $k$ is positive, we have

$$\mathcal{V}'_0 = \operatorname{span}\{f^i \otimes_{\mathsf{U}(\mathfrak{b})} v : i \in \mathbb{Z}_{\geqslant 0}\} \tag{5.33}$$

(as a vector space). The $\mathfrak{sl}(2)$-action is by left-multiplication so

$$\begin{aligned} f \cdot f^i \otimes_{\mathsf{U}(\mathfrak{b})} v &= f^{i+1} \otimes_{\mathsf{U}(\mathfrak{b})} v, \\ h \cdot f^i \otimes_{\mathsf{U}(\mathfrak{b})} v &= f^i \otimes_{\mathsf{U}(\mathfrak{b})} hv + [h, f^i] \otimes_{\mathsf{U}(\mathfrak{b})} v = -2f^i \otimes_{\mathsf{U}(\mathfrak{b})} v, \\ e \cdot f^i \otimes_{\mathsf{U}(\mathfrak{b})} v &= f^i \otimes_{\mathsf{U}(\mathfrak{b})} ev + [e, f^i] \otimes_{\mathsf{U}(\mathfrak{b})} v = if^{i-1} \otimes_{\mathsf{U}(\mathfrak{b})} \big(h - (i-1)\mathbb{1}\big)v \\ &= -i(i-1)f^{i-1} \otimes_{\mathsf{U}(\mathfrak{b})} v, \end{aligned} \tag{5.34}$$

by Exercise 35. Note that $\mathbb{1} \otimes_{\mathsf{U}(\mathfrak{b})} v$ is a highest-weight vector that generates $\mathcal{V}'_0$ and so $\mathcal{V}'_0$ is a highest-weight $\mathfrak{sl}(2)$-module. However, it is clearly **not** finite-dimensional because the $f^i \otimes_{\mathsf{U}(\mathfrak{b})} v$, with $i \in \mathbb{Z}_{\geqslant 0}$, are linearly independent (so they form a basis of $\mathcal{V}'_0$). ▲

**Exercise 94.** Check if the $\mathfrak{sl}(2)$-module $\mathcal{V}'_0$ constructed in the previous example is isomorphic to the $\mathfrak{sl}(2)$-module $\mathbb{C}[z]$ studied in Exercises 39 and 42. ▼

**Exercise 95.** Write down a basis for a general Verma module $\mathcal{V}'_\lambda$ over $\mathfrak{sl}(3)$, $\mathfrak{sp}(4)$ and $\mathfrak{g}_2$, drawing the weight diagram (schematically) in each case. ▼

**Exercise 96.** Show, in general, that $\mathbb{1} \otimes_{\mathsf{U}(\mathfrak{b})} v$ is always a non-zero highest-weight vector of $\mathcal{V}'_\lambda$ of weight $\lambda$ that generates $\mathcal{V}'_\lambda$. In other words, show that $\mathcal{V}'_\lambda$ is a highest-weight module. Use the Poincaré–Birkhoff–Witt theorem to explain why $\mathcal{V}'_\lambda \simeq \mathsf{U}(\mathfrak{g}_-)$, as vector spaces, and thereby write down a basis of $\mathcal{V}'_\lambda$. ▼

**Exercise 97.** Show that every weight of $\mathcal{V}'_\lambda$ comes with finite multiplicity, *ie.* the dimension of each weight space is finite. Explain why the multiplicity of the highest-weight vector $\mathbb{1} \otimes_{\mathsf{U}(\mathfrak{b})} v$ is 1. ▼

The concrete and abstract answers to the existence of highest-weight modules of course are the same. We shall therefore also refer to $\mathcal{V}'_\lambda$ as a Verma module.

**Lemma 5.9.** *Given any $\lambda \in \mathfrak{g}_0^*$, we have the vector space decomposition*

$$\mathsf{U}(\mathfrak{g}) = \mathsf{U}(\mathfrak{g}_-) \oplus \mathsf{I}_\lambda, \quad \mathsf{I}_\lambda = \sum_{\alpha\in\Delta_+} \operatorname{im}\phi_{e_\alpha} + \sum_{i=1}^{\operatorname{rank}\mathfrak{g}} \operatorname{im}\phi_{h_{\alpha_i}-\lambda_i\mathbb{1}}. \tag{5.35}$$

*Proof.* We first note that $\mathsf{U}(\mathfrak{g}_+) = \mathbb{C}\mathbb{1} \oplus \sum_{\alpha\in\Delta_+} \mathsf{U}(\mathfrak{g}_+)e_\alpha$, because a Poincaré–Birkhoff–Witt basis monomial either contains a positive root vector or is $\mathbb{1}$, and similarly that $\mathsf{U}(\mathfrak{g}_0) = \mathbb{C}\mathbb{1} \oplus \sum_{i=1}^{\operatorname{rank}\mathfrak{g}} \mathsf{U}(\mathfrak{g}_0)(h_{\alpha_i} - \lambda_i\mathbb{1})$. Exercise 88 then gives (dropping tensor product symbols for brevity)

$$\begin{aligned}\mathsf{U}(\mathfrak{g}) &= \mathsf{U}(\mathfrak{g}_-)\mathsf{U}(\mathfrak{g}_0)\Big[\mathbb{C}\mathbb{1} \oplus \sum_{\alpha\in\Delta_+}\mathsf{U}(\mathfrak{g}_+)e_\alpha\Big] = \mathsf{U}(\mathfrak{g}_-)\mathsf{U}(\mathfrak{g}_0) \oplus \sum_{\alpha\in\Delta_+}\mathsf{U}(\mathfrak{g})e_\alpha \\ &= \mathsf{U}(\mathfrak{g}_-)\Big[\mathbb{C}\mathbb{1} \oplus \sum_{i=1}^{\operatorname{rank}\mathfrak{g}}\mathsf{U}(\mathfrak{g}_0)(h_{\alpha_i} - \lambda_i\mathbb{1})\Big] \oplus \sum_{\alpha\in\Delta_+}\mathsf{U}(\mathfrak{g})e_\alpha \\ &= \mathsf{U}(\mathfrak{g}_-) \oplus \sum_{\alpha\in\Delta_+}\mathsf{U}(\mathfrak{g})e_\alpha \oplus \sum_{i=1}^{\operatorname{rank}\mathfrak{g}}\mathsf{U}(\mathfrak{g}_-)\mathsf{U}(\mathfrak{g}_0)(h_{\alpha_i} - \lambda_i\mathbb{1}).\end{aligned} \tag{5.36}$$

This is almost what we want because $\operatorname{im}\phi_U = \mathsf{U}(\mathfrak{g})U$, for all $U \in \mathsf{U}(\mathfrak{g})$. We therefore need to replace $\mathsf{U}(\mathfrak{g}_-)\mathsf{U}(\mathfrak{g}_0)$ in the last direct summand above by $\mathsf{U}(\mathfrak{g})$. Obviously,

$$\sum_{i=1}^{\operatorname{rank}\mathfrak{g}}\mathsf{U}(\mathfrak{g})(h_{\alpha_i} - \lambda_i\mathbb{1}) \supseteq \sum_{i=1}^{\operatorname{rank}\mathfrak{g}}\mathsf{U}(\mathfrak{g}_-)\mathsf{U}(\mathfrak{g}_0)(h_{\alpha_i} - \lambda_i\mathbb{1}). \tag{5.37}$$

One the other hand, we have

$$\sum_{i=1}^{\operatorname{rank}\mathfrak{g}}\mathsf{U}(\mathfrak{g})(h_{\alpha_i} - \lambda_i\mathbb{1}) = \sum_{i=1}^{\operatorname{rank}\mathfrak{g}}\mathsf{U}(\mathfrak{g}_-)\mathsf{U}(\mathfrak{g}_0)\Big[\mathbb{C}\mathbb{1} \oplus \sum_{\alpha\in\Delta_+}\mathsf{U}(\mathfrak{g}_+)e_\alpha\Big](h_{\alpha_i} - \lambda_i\mathbb{1}) \tag{5.38}$$

$$= \sum_{i=1}^{\operatorname{rank}\mathfrak{g}} \mathsf{U}(\mathfrak{g}_-)\mathsf{U}(\mathfrak{g}_0)(h_{\alpha_i} - \lambda_i \mathbb{1}) \oplus \sum_{\alpha\in\Delta_+} \mathsf{U}(\mathfrak{g}) \sum_{i=1}^{\operatorname{rank}\mathfrak{g}} \bigl(h_{\alpha_i} - \bigl(\lambda_i + 2\alpha(h_{\alpha_i})\bigr)\mathbb{1}\bigr)e_\alpha$$

$$\subseteq \sum_{\alpha\in\Delta_+} \mathsf{U}(\mathfrak{g})e_\alpha \oplus \sum_{i=1}^{\operatorname{rank}\mathfrak{g}} \mathsf{U}(\mathfrak{g}_-)\mathsf{U}(\mathfrak{g}_0)(h_{\alpha_i} - \lambda_i \mathbb{1}).$$

It therefore follows that

$$\sum_{\alpha\in\Delta_+} \mathsf{U}(\mathfrak{g})e_\alpha + \sum_{i=1}^{\operatorname{rank}\mathfrak{g}} \mathsf{U}(\mathfrak{g})(h_{\alpha_i} - \lambda_i \mathbb{1}) = \sum_{\alpha\in\Delta_+} \mathsf{U}(\mathfrak{g})e_\alpha \oplus \sum_{i=1}^{\operatorname{rank}\mathfrak{g}} \mathsf{U}(\mathfrak{g}_-)\mathsf{U}(\mathfrak{g}_0)(h_{\alpha_i} - \lambda_i \mathbb{1}), \tag{5.39}$$

completing the proof. ∎

**Proposition 5.10.** *The* $\mathfrak{g}$*-modules* $\mathcal{V}_\lambda$ *and* $\mathcal{V}'_\lambda$ *are isomorphic, for all* $\lambda \in \mathfrak{g}_0^*$.

*Proof.* Define a map $\Phi\colon \mathsf{U}(\mathfrak{g}) \to \mathcal{V}'_\lambda = \mathsf{U}(\mathfrak{g}) \otimes_{\mathsf{U}(\mathfrak{b})} \mathbb{C}_\lambda$ by $\Phi(U) = U \otimes_{\mathsf{U}(\mathfrak{b})} v$, where $v$ is some fixed spanning element of $\mathbb{C}_\lambda$ with $h_{\alpha_i} v = \lambda_i v$ and $e_\alpha v = 0$ for $\alpha \in \Delta_+$. This map is obviously linear and surjective. Moreover, it is also a $\mathfrak{g}$-module homomorphism:

$$x\Phi(U) = x(U \otimes_{\mathsf{U}(\mathfrak{b})} v) = (xU) \otimes_{\mathsf{U}(\mathfrak{b})} v = \Phi(xU), \quad \text{for all } x \in \mathfrak{g} \text{ and } U \in \mathsf{U}(\mathfrak{g}). \tag{5.40}$$

Obviously, $\ker\Phi \supseteq \mathsf{I}_\lambda = \sum_{\alpha\in\Delta_+} \operatorname{im}\phi_{e_\alpha} + \sum_{i=1}^{\operatorname{rank}\mathfrak{g}} \operatorname{im}\phi_{h_{\alpha_i}-\lambda_i\mathbb{1}}$, because both the $e_\alpha$ and the $h_{\alpha_i} - \lambda_i\mathbb{1}$ belong to $\mathsf{U}(\mathfrak{b})$ and annihilate $v$. However, if this was strict, then some non-zero element $U \in \mathsf{U}(\mathfrak{g}_-)$ would belong to $\ker\Phi$, by Lemma 5.9. But, $\Phi(U) = U \otimes_{\mathsf{U}(\mathfrak{b})} v \neq 0$ because $\Phi|_{\mathsf{U}(\mathfrak{g}_-)}$ maps a Poincaré–Birkhoff–Witt basis of $\mathsf{U}(\mathfrak{g}_-)$ to a basis of $\mathcal{V}'_\lambda$. Thus, $\ker\Phi = \mathsf{I}_\lambda$ and so the first isomorphism theorem for modules (Proposition 2.4) gives

$$\mathcal{V}_\lambda = \frac{\mathsf{U}(\mathfrak{g})}{\ker\Phi} \simeq \operatorname{im}\Phi = \mathcal{V}'_\lambda. \tag{5.41}$$ ∎

Exercise 96 makes it clear that, no matter how one chooses to construct them, Verma modules are infinite-dimensional. But, we said before that our goal was to classify the finite-dimensional irreducible $\mathfrak{g}$-modules, for $\mathfrak{g}$ semisimple, so devoting time to Verma modules deserves some justification. The point of these modules are that they have the **universal** property of being maximal among highest-weight modules in the sense that they are generated from a highest-weight vector that satisfies only the requirements of being a highest-weight vector.

**Proposition 5.11.** *If* $W$ *is a highest-weight* $\mathfrak{g}$*-module that is generated by a highest-weight vector* $w$ *of weight* $\lambda$*, then there exists a surjective* $\mathfrak{g}$*-module homomorphism* $\psi\colon \mathcal{V}_\lambda \to W$*. In other words,* $W$ *is isomorphic to the quotient* $\mathcal{V}_\lambda/\ker\psi$.

*Proof.* Consider the (obviously linear) map $\psi'\colon \mathsf{U}(\mathfrak{g}) \to W$ defined by $\psi'(U) = Uw$. This is a surjective $\mathfrak{g}$-module homomorphism, since $x\psi'(U) = xUw = \psi'(xU)$ and $W$ is generated by $w$. Moreover, $\psi'$ annihilates $\mathsf{I}_\lambda = \sum_{\alpha\in\Delta_+} \operatorname{im}\phi_{e_\alpha} + \sum_i \operatorname{im}\phi_{h_{\alpha_i}-\lambda_i\mathbb{1}}$ because

$w$ is a highest-weight vector. Thus, $\psi'$ descends to a well-defined surjective $\mathfrak{g}$-module homomorphism from $\mathcal{V}_\lambda = \mathsf{U}(\mathfrak{g})/\mathsf{I}_\lambda$ to $W$. This is $\psi$. ■

Let $v_\lambda$ denote the generating highest-weight vector of $\mathcal{V}_\lambda$, *ie.* $v_\lambda = \overline{\mathbb{1}}$ in the notation of the proof of Proposition 5.8 and $v_\lambda = \mathbb{1} \otimes_{\mathsf{U}(\mathfrak{b})} v$ in that of Exercise 96. As a generator, $v_\lambda$ cannot belong to any **proper** submodule of $\mathcal{V}_\lambda$. The (frequently non-direct) sum of **all** the proper submodules of $\mathcal{V}_\lambda$ is therefore still proper and so is the (unique) maximal proper submodule of $\mathcal{V}_\lambda$. We denote this maximal proper submodule by $\mathcal{J}_\lambda$. By Exercise 33, $\mathcal{V}_\lambda$ has a unique irreducible quotient $\mathcal{L}_\lambda = \mathcal{V}_\lambda/\mathcal{J}_\lambda$. Moreover, the class $\overline{v_\lambda} \in \mathcal{L}_\lambda$ is non-zero, hence is a highest-weight vector of weight $\lambda$. Since every irreducible highest-weight module has a unique highest weight (Proposition 5.6), we obtain the following:

**Corollary 5.12.** *Any irreducible highest-weight module $\mathcal{L}$ is isomorphic to the quotient $\mathcal{L}_\lambda$ of $\mathcal{V}_\lambda$, where $\lambda$ is the unique highest weight of $\mathcal{L}$.*

We emphasise that this not only gives us a uniform construction of an irreducible highest-weight module, for any highest weight $\lambda \in \mathfrak{g}_0^*$, it also proves that such an irreducible highest-weight module is *unique*, up to isomorphism.

Finally, we raise the (very real) possibility that a Verma module $\mathcal{V}_\lambda$ may have more than one highest weight, *ie.* that there may exist a highest-weight vector $w_\mu \in \mathcal{V}_\lambda$, of weight $\mu \neq \lambda$, in addition to the generating highest-weight vector $v_\lambda$. If such a $w_\mu$ exists, then $\mathcal{V}_\lambda$ must be reducible by Proposition 5.6. These additional highest-weight vectors are called *singular vectors*. (Of course, singular vectors can exist in modules that aren't Verma!)

**Exercise 98.** Show that $f^{\lambda+1}v_\lambda$ is a singular vector in the Verma module $\mathcal{V}_\lambda$ of $\mathfrak{sl}(2)$ whenever $\lambda \in \mathbb{Z}_{\geqslant 0}$. (Here, we are identifying $\lambda \in \mathfrak{sl}(2)_0^*$ with its Dynkin label $\lambda_1 \in \mathbb{C}$.) Identify the (isomorphism classes of the) submodule $W$ generated by this singular vector and the quotient $\mathcal{V}_\lambda/W$. ▼

**Exercise 99.** Use the Poincaré–Birkhoff–Witt Theorem to prove that a singular vector $w_\mu \in \mathcal{V}_\lambda$ of weight $\mu$ always generates a submodule of $\mathcal{V}_\lambda$ that is isomorphic to $\mathcal{V}_\mu$. Conclude that every non-zero $\mathfrak{g}$-module homomorphism between Verma modules is injective. ▼

**Proposition 5.13.** *Let $v_\lambda$ be the generating highest-weight vector of the Verma module $\mathcal{V}_\lambda$. Then, the vector $w_j = f_{\alpha_j}^{\lambda_j+1} v_\lambda$ is singular in $\mathcal{V}_\lambda$ whenever the $j$-th Dynkin label $\lambda_j$ is a non-negative integer.*

*Proof.* Suppose that $\lambda_j \in \mathbb{Z}_{\geqslant 0}$. As $w_j$ is a weight vector (of weight $\lambda - (\lambda_j + 1)\alpha_j$), we only need to show that $e_\alpha w_j = 0$ for all $\alpha \in \Delta_+$. However, we note that the positive root vectors are generated under the Lie bracket by the simple root vectors $e_{\alpha_i}$, so it will be enough to show that $e_{\alpha_i} w_j = 0$, for all $i = 1, \dots, \operatorname{rank} \mathfrak{g}$.

When $i \neq j$, we have $[e_{\alpha_i}, f_{\alpha_j}] = 0$ because $\alpha_i - \alpha_j \notin \Delta$ (the difference of two simple roots is never a root). Thus,

$$e_{\alpha_i} w_j = e_{\alpha_i} f_{\alpha_j}^{\lambda_j+1} v_\lambda = f_{\alpha_j}^{\lambda_j+1} e_{\alpha_i} v_\lambda = 0. \tag{5.42}$$

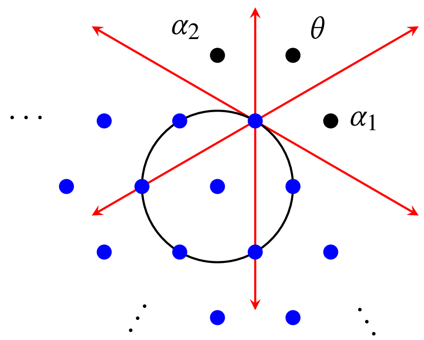
However, when $i = j$, $e_{\alpha_j}$, $f_{\alpha_j}$ and $h_{\alpha_j}$ span an $\mathfrak{sl}(2)$-subalgebra so that $e_{\alpha_j} w_j = 0$ follows from Exercise 98 and the fact that $v_\lambda$ is a highest-weight vector of weight $\lambda_j$ with respect to this subalgebra. ■

**Example 55.** We illustrate this proposition with the Verma module $\mathcal{V}_0$ of $\mathfrak{sl}(3)$. As both Dynkin labels of the highest weight are 0, Proposition 5.13 gives two singular vectors: $w_1 = f_{\alpha_1} v_0$ and $w_2 = f_{\alpha_2} v_0$. Moreover, the weight of $w_1$ is $-\alpha_1 = (-2, 1)$ and that of $w_2$ is $-\alpha_2 = (1, -2)$. As $w_1$ and $w_2$ generate submodules isomorphic to Verma modules, by Exercise 99, Proposition 5.13 gives two more singular vectors $w_3 = f_{\alpha_2}^2 w_1 = f_{\alpha_2}^2 f_{\alpha_1} v_0$ and $w_4 = f_{\alpha_1}^2 w_2 = f_{\alpha_1}^2 f_{\alpha_2} v_0$. Moreover, $w_3$ and $w_4$ have weights $-\alpha_1 - 2\alpha_2 = (0, -3)$ and $-2\alpha_1 - \alpha_2 = (-3, 0)$, respectively, hence Proposition 5.13 gives two more singular vectors $w_5 = f_{\alpha_1} w_3 = f_{\alpha_1} f_{\alpha_2}^2 f_{\alpha_1} v_0$ and $w_6 = f_{\alpha_2} w_4 = f_{\alpha_2} f_{\alpha_1}^2 f_{\alpha_2} v_0$. They both have weight $-2\alpha_1 - 2\alpha_2 = (-2, -2)$, so Proposition 5.13 generates no further singular vectors.

In fact, the last two singular vectors coincide: taking $f_{\alpha_1} = E_{21}$, $f_{\alpha_2} = E_{32}$ and $f_\theta = E_{31}$ gives $[f_{\alpha_2}, f_{\alpha_1}] = f_\theta$ and $[f_\theta, f_{\alpha_1}] = [f_\theta, f_{\alpha_1}] = 0$, hence

$$f_{\alpha_1} f_{\alpha_2}^2 f_{\alpha_1} = f_{\alpha_1} f_{\alpha_2} f_{\alpha_1} f_{\alpha_2} + f_{\alpha_1} f_{\alpha_2} f_\theta = f_{\alpha_1} f_{\alpha_2} f_{\alpha_1} f_{\alpha_2} + f_\theta f_{\alpha_1} f_{\alpha_2} = f_{\alpha_2} f_{\alpha_1}^2 f_{\alpha_2}. \tag{5.43}$$

One can check explicitly that there are no other singular vectors lurking in the weights spaces considered here. In fact, there are no other singular vectors in $\mathcal{V}_0$: we shall see in Corollary 5.21 below that the weight $\mu$ of any such singular vector must satisfy $\|\mu + \theta\|^2 = \|\theta\|^2$. *ie.*, $\mu$ lies on the circle (in $\mathbb{R}$) centred at $-\theta$ that passes through 0.




The only weights of $\mathcal{V}_0$ (blue in the above diagram) lying on this circle are those in which we have already found a singular vector. ▲

**Exercise 100.** Use Proposition 5.13 to find 8 singular vectors (including the highest-weight vector) of the $\mathfrak{sp}(4)$-module $\mathcal{V}_{\alpha_1+\alpha_2}$. Use Corollary 5.21 below to argue that the corresponding 8 weights exhaust the possible singular vector weights. ▼

**Exercise 101.** The singular vectors of Proposition 5.13 need not exhaust the singular vectors of a Verma module. Show that the $\mathfrak{sl}(3)$-module $\mathcal{V}_{-(\alpha_1+\alpha_2)/2}$ has a singular vector of weight $-\frac{3}{2}(\alpha_1+\alpha_2)$ that is not covered by Proposition 5.13. ▼

A singular vector of weight $\mu$ generates a submodule of $\mathcal{V}_\lambda$ isomorphic to $\mathcal{V}_\mu$, by Exercise 99. This Verma submodule may well have singular vectors generating Verma submodules of $\mathcal{V}_\mu$, which will of course be submodules of $\mathcal{V}_\lambda$, and so on. The question of how all of these submodules are embedded into one another is interesting and important, but far too intricate to answer here — the main keyword for those who are interested is *Kazhdan–Lusztig polynomials*. Instead, we content ourselves with a simple observation.

**Proposition 5.14.** *Every non-zero submodule of a Verma module possesses a highest-weight vector.*

*Proof.* Let $W$ be a non-zero submodule and $w \in W$ a non-zero element. As Verma modules are weight modules, $w$ may be written as a **finite** linear combination of weight vectors $w_\mu$: $w = \sum_{\mu\in\Gamma} c_\mu w_\mu$ for some $c_\mu \in \mathbb{C}$. Here, $\mu$ is the weight of $w_\mu$ and the set $\Gamma = \{\mu : c_\mu \neq 0\}$ is finite. Choose an arbitrary weight $\nu \in \Gamma$. Then, for each $\mu \in \Gamma \setminus \{\nu\}$, there exists $H_\mu \in \mathfrak{g}_0$ such that $\mu(H_\mu) \neq \nu(H_\mu)$. We now consider the **finite** product

$$U = \prod_{\mu\in\Gamma\setminus\{\nu\}} (H_\mu - \mu(H_\mu)\mathbb{1}) \in \mathsf{U}(\mathfrak{g}_0). \tag{5.44}$$

Each of the weight vectors $w_\mu$, with $\mu \neq \nu$, is annihilated by $U$; however, $w_\nu$ is not. Thus, $Uw \in W$ is non-zero and proportional to $w_\nu$. This shows that $W$ contains a weight vector.

But, if $w \in W$ is a weight vector, then its weight has the form $\lambda - \sum_i m_i\alpha_i$, where $\lambda$ is the (generating) highest weight of the Verma module and the $m_i$ are non-negative integers. Call the sum of the $m_i$ the *depth* of $w$. Consider the action of the simple root vectors $e_{\alpha_i}$ on $w$. If the result is 0 for all $i$, then $w$ is a highest-weight vector of $W$ and we are done. If the result is non-zero for some $i$, then we obtain a new weight vector in $W$ whose depth is one less than that of $w$. Act on this new weight vector with the $e_{\alpha_i}$ and repeat until we arrive at a highest-weight vector of $W$; this is guaranteed because the depth decreases by one with each iteration and is bounded below by 0. ■

This proposition says that every non-zero submodule of a Verma module has singular vectors. However, it **doesn't** say that every non-zero submodule of a Verma module is

**generated** by its singular vectors — that statement is false. Moreover, non-zero submodules that are not generated by their singular vectors only appear when rank $\mathfrak{g} \geqslant 3$, showing that one can't always trust the intuition gleaned from small-rank examples.

## 5.5. **The Weyl group and finite-dimensionality**

We've just seen how to construct a Verma module with an arbitrary highest weight $\lambda \in \mathfrak{g}_0^*$ We've also shown that every irreducible highest-weight module, hence every irreducible finite-dimensional module, can be constructed as a quotient of a Verma module. To answer the question of whether all dominant integral $\lambda \in \mathsf{P}_{\geqslant}$ yield finite-dimensional irreducible quotients or not, we need another basic tool of Lie theory: the Weyl group.

Recall from Proposition 4.13 that $\alpha, \beta \in \Delta$ implies that $\beta - \beta(\alpha^\vee)\alpha \in \Delta$. It follows that the *Weyl reflections* $\mathsf{w}_\alpha$, defined on $\mathfrak{g}_0^*$ for each $\alpha \in \Delta_+$ by

$$\mathsf{w}_\alpha(\lambda) = \lambda - \lambda(\alpha^\vee)\alpha, \tag{5.45}$$

send roots to roots, *ie.* they are symmetries of the root system. The name arises because in the rational vector space $\mathsf{R}$ (or its real completion), $\mathsf{w}_\alpha$ may be identified geometrically with the reflection about the codimension-1 hyperplane orthogonal to $\alpha$:

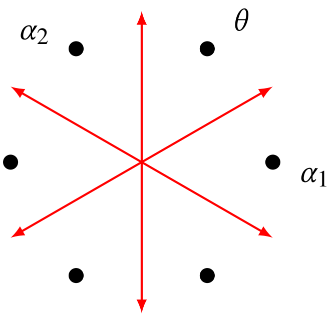
$$\mathsf{w}_\alpha(\alpha) = -\alpha \quad \text{and} \quad \mathsf{w}_\alpha(\lambda) = \lambda \quad \text{if } (\lambda, \alpha) = 0. \tag{5.46}$$

We define the *Weyl group* $\mathsf{W}$ to be the group generated by the *simple Weyl reflections* $\mathsf{w}_{\alpha_i}$. (This turns out to coincide with the group generated by all the Weyl reflections, though we will not need this fact!) Since the generators permute the roots, it follows that the Weyl group is a subgroup of the symmetric group on $\Delta$; in particular, $\mathsf{W}$ is a finite group.

**Exercise 102.** Check that the Weyl reflections define orthogonal linear maps on $\mathsf{R}$, with respect to the inner product $(\cdot,\cdot)$, that square to the identity. (They are thus automorphisms of the root system.) Show that the Weyl reflections preserve the weight lattice $\mathsf{P}$. ▼

**Example 56.** The Weyl group of $\mathfrak{sl}(2)$ is generated by a single Weyl reflection $\mathsf{w}_\alpha$ that maps $\alpha$ to $-\alpha$. The Weyl group is therefore cyclic of order 2. ▲

**Example 57.** The Weyl group of $\mathfrak{sl}(3)$ is generated by the two simple Weyl reflections $\mathsf{w}_{\alpha_1}$ and $\mathsf{w}_{\alpha_2}$. Since $\mathsf{w}_{\alpha_1}\mathsf{w}_{\alpha_2}$ maps $\alpha_1$ to $\alpha_2$ and $\alpha_2$ to $-\theta$, where $\theta = \alpha_1 + \alpha_2$, we may identify $\mathsf{w}_{\alpha_1}\mathsf{w}_{\alpha_2}$ with an anticlockwise rotation by 120°. Similarly, $\mathsf{w}_{\alpha_2}\mathsf{w}_{\alpha_1}$ corresponds to a clockwise rotation by 120°. Moreover, both $\mathsf{w}_{\alpha_1}\mathsf{w}_{\alpha_2}\mathsf{w}_{\alpha_1}$ and $\mathsf{w}_{\alpha_2}\mathsf{w}_{\alpha_1}\mathsf{w}_{\alpha_2}$ send $\alpha_1$ to $-\alpha_2$ and $\alpha_2$ to $-\alpha_1$, hence they may be identified with $\mathsf{w}_\theta$. It follows that the Weyl group has order 6 (it is isomorphic to the symmetric group $\mathsf{S}_3$).




Notice that $\mathsf{W}$ is not the full symmetry group of the root system — rotations by 60° are missing for example. The full symmetry group is isomorphic to the dihedral group $\mathsf{D}_6$ (of order 12). ▲

We remark that all elements of the Weyl group of $\mathfrak{g}$ may be identified as *inner automorphisms*, meaning that their action is equivalent to conjugating by an element of a Lie group whose Lie algebra is $\mathfrak{g}$. The root system symmetries that are not in the Weyl group are therefore often referred to as *outer automorphisms*. It turns out that the quotient of the group of all symmetries of the root system of $\mathfrak{g}$ by the inner automorphisms $\mathsf{W}$ is isomorphic to the group of symmetries of the Dynkin diagram of $\mathfrak{g}$!

To illustrate this, consider the case $\mathfrak{g} = \mathfrak{sl}(3)$ as in Example 57. The quotient $\mathsf{D}_6/\mathsf{S}_3$ is clearly isomorphic to $\mathbb{Z}_2$ which is indeed the group of symmetries of the Dynkin diagram $\Gamma^{\mathfrak{sl}(3)}$. Moreover, the non-trivial Dynkin symmetry swaps the vertices 1 and 2, whilst swapping $\alpha_1$ and $\alpha_2$ in the root system corresponds to a reflection about the hyperplane through $\theta$. This is, of course, not a Weyl reflection.

It turns out that the Weyl groups of the classical Lie algebras (and $\mathfrak{g}_2$) are fairly familiar, whilst those of the other exceptional Lie algebras are not (though they are known explicitly). We present the familiar cases (and the orders of all Weyl groups) in the following table.

| $\mathfrak{g}$ | $\mathfrak{sl}(r+1)$ | $\mathfrak{so}(2r+1)$ | $\mathfrak{sp}(2r)$ | $\mathfrak{so}(2r)$ | $\mathfrak{g}_2$ |
|---|---|---|---|---|---|
| $\mathsf{W}$ | $\mathsf{S}_{r+1}$ | $\mathsf{S}_r \ltimes \mathbb{Z}_2^r$ | $\mathsf{S}_r \ltimes \mathbb{Z}_2^r$ | $\mathsf{S}_r \ltimes \mathbb{Z}_2^{r-1}$ | $\mathsf{D}_6$ |
| $\lvert\mathsf{W}\rvert$ | $(r+1)!$ | $2^r r!$ | $2^r r!$ | $2^{r-1} r!$ | 12 |

For completeness, the orders of the Weyl groups of $\mathfrak{f}_4$, $\mathfrak{e}_6$, $\mathfrak{e}_7$ and $\mathfrak{e}_8$ are 1152, 51840, 2903040 and 696729600, respectively.

Don't worry if you don't know what a "semidirect product" $\ltimes$ is — we won't be using these identifications in what follows. To do so, we'd also need to know how these abstract groups act on the root system and this isn't always obvious. It's a beautiful story of course and you can read about it at length in (*eg.*) the books of Fulton and Harris or Humphreys.

**Exercise 103.**

- Use the root systems of $\mathfrak{so}(4)$, $\mathfrak{sp}(4)$, and $\mathfrak{g}_2$ to verify their Weyl groups. In each case, determine the minimal $m \in \mathbb{Z}_{>0}$ such that $(\mathsf{w}_{\alpha_1}\mathsf{w}_{\alpha_2})^m$ is the identity.

- Compute the root system of $\mathfrak{sp}(6)$ and determine the action of the simple Weyl reflections on it. Again, compute the minimal $m_{ij} \in \mathbb{Z}_{>0}$ such that $(\mathsf{w}_{\alpha_i}\mathsf{w}_{\alpha_j})^{m_{ij}}$ is the identity. ▼

As you can see, the Weyl group explains the beautiful symmetries we have observed in the root diagrams. We have also observed symmetries in the weight diagrams of the finite-dimensional irreducible modules (though not in those of the infinite-dimensional Verma modules). In fact, these symmetries are also manifestations of the Weyl group: $\mathsf{W}$ permutes the weights of any given finite-dimensional module. This is the key to proving the finite-dimensionality of the irreducible highest-weight modules with dominant integral highest weights and, as always, the argument reduces to $\mathfrak{sl}(2)$-theory.

Recall that for $\mathfrak{sl}(2)$, the simple root is $\alpha = (2)$, the simple coroot is $\alpha^\vee = h$ and a weight $\lambda = (\lambda)$ is mapped by the single Weyl reflection to $\lambda - \lambda(H)\alpha = -\lambda$. Since every finite-dimensional $\mathfrak{sl}(2)$-module is a direct sum of irreducibles (Theorem 3.5) and every finite-dimensional irreducible $\mathfrak{sl}(2)$-module has weights $\lambda, \lambda - 2, \dots, -\lambda + 2, -\lambda$ (Theorem 3.1 and Example 47), it follows that the weights of any finite-dimensional $\mathfrak{sl}(2)$-module are permuted by the action of the Weyl group.

**Lemma 5.15.** *When $\lambda \in \mathsf{P}_{\geqslant}$, the weights of the irreducible highest weight $\mathfrak{g}$-module $\mathcal{L}_\lambda$ are permuted by* $\mathsf{W}$.

*Proof.* Let $v$ be the highest-weight vector of $\mathcal{L}_\lambda$ and $v'$ the highest-weight vector of its Verma cover $\mathcal{V}_\lambda$, guaranteed to exist by Corollary 5.12. Since $\lambda$ is dominant integral, the vectors $f_{\alpha_i}^{\lambda_i+1} v'$ are singular in $\mathcal{V}_\lambda$ for each $i$, by Proposition 5.13, hence they generate non-zero proper submodules. These submodules are set to zero in constructing the irreducible quotient $\mathcal{L}_\lambda$, so $f_{\alpha_i}^{\lambda_i+1} v = 0$ in $\mathcal{L}_\lambda$, for all $i = 1, \dots, \operatorname{rank} \mathfrak{g}$.

It follows that for each $i$, the subspace of $\mathcal{L}_\lambda$ spanned by the $f_{\alpha_i}^n v$, with $n = 0, 1, \dots, \lambda_i$, is a finite-dimensional module for the $\mathfrak{sl}(2)$-subalgebra spanned by $e_{\alpha_i}$, $h_{\alpha_i}$ and $f_{\alpha_i}$. We will denote this subalgebra by $\mathfrak{sl}(2)_i$ for convenience. We conclude that for each $i$, the set of finite-dimensional $\mathfrak{sl}(2)_i$-submodules of $\mathcal{L}_\lambda$ is not empty.

Fix $i$ and let $M_i$ be the sum of all the finite-dimensional $\mathfrak{sl}(2)_i$-submodules of $\mathcal{L}_\lambda$. We have just shown that $M_i$ is non-zero, so if we can show that it is a $\mathfrak{g}$-module, then it will follow that $M_i = \mathcal{L}_\lambda$, by irreducibility. Consider therefore an arbitrary finite-dimensional $\mathfrak{sl}(2)_i$-submodule $W \subseteq \mathcal{L}_\lambda$. Then, the subspace

$$W' = \operatorname{span}\{yw : y \in \mathfrak{g} \text{ and } w \in W\} \tag{5.47}$$

is a $\mathfrak{sl}(2)_i$-submodule of $\mathcal{L}_\lambda$ because $x \in \mathfrak{sl}(2)_i$ gives $x(yw) = y(xw) + [x, y]w \in W'$. Moreover, the dimension of $W'$ is finite, being bounded above by $\dim \mathfrak{g} \dim W$. Thus, acting with $y \in \mathfrak{g}$ maps each finite-dimensional $\mathfrak{sl}(2)_i$-submodule $W$ into another one: $W'$. $M_i$ is therefore a $\mathfrak{g}$-module, hence it coincides with $\mathcal{L}_\lambda$ for each $i = 1, \dots, \operatorname{rank} \mathfrak{g}$.

In other words, $\mathcal{L}_\lambda$ is, for each $i$, a sum of finite-dimensional $\mathfrak{sl}(2)_i$-submodules. Since the weights of such a submodule are permuted by the simple Weyl reflection $\mathsf{w}_{\alpha_i}$, we conclude that the weights of $\mathcal{L}_\lambda$ are likewise permuted by each $\mathsf{w}_{\alpha_i}$. But, as this is true for each $i$, and as $\mathsf{W}$ is generated by the $\mathsf{w}_{\alpha_i}$, it follows that the weights of $\mathcal{L}_\lambda$ are permuted by the Weyl group of $\mathfrak{g}$. ■

**Corollary 5.16.** *The weights of any finite-dimensional* $\mathfrak{g}$*-module are permuted by* $\mathsf{W}$.

This permutation result will allow us to complete the classification of finite-dimensional irreducible $\mathfrak{g}$-modules. The following theorem is the generalisation of Theorem 3.1 from $\mathfrak{sl}(2)$ to general (finite-dimensional complex) semisimple Lie algebras $\mathfrak{g}$.

**Theorem 5.17.** *The irreducible* $\mathfrak{g}$*-module* $\mathcal{L}_\lambda$ *is finite-dimensional if and only if* $\lambda \in \mathsf{P}_\geqslant$.

*Proof.* We've already shown that finite-dimensionality implies the dominant integral condition (Corollary 5.7). We therefore assume that $\lambda \in \mathsf{P}_\geqslant$.

Since any weight $\mu$ of $\mathcal{L}_\lambda$ differs from $\lambda$ by an element of the root lattice $\mathsf{Q} \subseteq \mathsf{P}$ (Proposition 5.2), it is integral: $\mu \in \mathsf{P}$. The set of weights of $\mathcal{L}_\lambda$ is also permuted by $\mathsf{W}$ (Lemma 5.15), so it decomposes into orbits $\mathsf{W}(\mu) = \{\mathsf{w}(\mu) : \mathsf{w} \in \mathsf{W}\} \subset \mathsf{P}$ under the Weyl group action. Moreover, these orbits are finite sets, as $\mathsf{W}$ is finite, and so have maximal elements (for any reasonable definition of "maximal"). We say that $\nu \in \mathsf{W}(\mu)$ is *maximal* if for all $\nu' \in \mathsf{W}(\mu)$, $\nu' - \nu$ is not a non-negative integer linear combination of simple roots: $\nu' - \nu \notin \mathsf{Q}_\geqslant$.

Now, if $\nu$ is not dominant integral, *ie.* $\nu \notin \mathsf{P}_\geqslant$, then $\nu_i < 0$ for some $i$, so

$$\mathsf{w}_{\alpha_i}(\nu) - \nu = -\nu(\alpha_i^\vee)\alpha_i = |\nu_i|\alpha_i \in \mathsf{Q}_\geqslant \tag{5.48}$$

and thus $\nu$ is not maximal. We conclude that maximal weights in a Weyl group orbit are dominant integral; every Weyl group orbit therefore includes a dominant integral weight.

How many dominant integral weights can $\mathcal{L}_\lambda$ have? Well, $\mu \in \mathsf{P}_\geqslant$ implies that $\lambda+\mu \in \mathsf{P}_\geqslant$, so $\lambda_i + \mu_i \in \mathbb{Z}_{\geqslant 0}$ for all $i$, and $\lambda - \mu \in \mathsf{Q}_\geqslant$, so $\lambda - \mu = \sum_j m_j \alpha_j$ for some $m_j \in \mathbb{Z}_{\geqslant 0}$. Thus,

$$(\lambda + \mu, \lambda - \mu) = \sum_{i,j=1}^{\operatorname{rank}\mathfrak{g}} (\lambda_i + \mu_i) m_j (\omega_i, \alpha_j) = \frac{1}{2} \sum_{i=1}^{\operatorname{rank}\mathfrak{g}} (\lambda_i + \mu_i) m_i \|\alpha_i\|^2 \geqslant 0. \tag{5.49}$$

However, $(\lambda + \mu, \lambda - \mu) = \|\lambda\|^2 - \|\mu\|^2$, hence we must have $\|\mu\|^2 \leqslant \|\lambda\|^2$. Now, the disc $\|\mu\|^2 \leqslant \|\lambda\|^2$ is compact and $\mathsf{P}_\geqslant$ is discrete, so their intersection is finite. But, if there are only finitely many dominant integral weights in $\mathcal{L}_\lambda$, then there can only be finitely many weights in $\mathcal{L}_\lambda$ because every Weyl group orbit is finite and contains at least one dominant integral weight. Finally, the dimension of every weight space of $\mathcal{L}_\lambda$ is bounded above the dimension of the corresponding weight space of its Verma cover $\mathcal{V}_\lambda$, which is finite by Exercise 97. We therefore conclude that $\mathcal{L}_\lambda$ is a finite-dimensional $\mathfrak{g}$-module. ■

**Corollary 5.18.** *The irreducible finite-dimensional $\mathfrak{g}$-modules are* classified*, up to isomorphism, by the set $\mathsf{P}_\geqslant$ of dominant integral weights.*

We close this section with a bit more fun that one can have with the Weyl group. The *Weyl vector* $\rho$ is an interesting weight that arises in many places (*eg*. in Proposition 5.20 below). It's defined as half the sum of the positive roots:

$$\rho = \frac{1}{2} \sum_{\alpha \in \Delta_+} \alpha. \tag{5.50}$$

With this definition, it looks unlikely to be an element of the root lattice $\mathsf{Q}$ or even the weight lattice $\mathsf{P}$. Nevertheless, we can compute the Dynkin labels of $\rho$ in various cases:

- $\mathfrak{sl}(2)$ has $\Delta_+ = \{(2)\}$, hence $\rho = (1)$.
- $\mathfrak{sl}(3)$ has $\Delta_+ = \{(2,-1),(-1,2),(1,1)\}$, hence $\rho = (1,1)$.
- $\mathfrak{sp}(4)$ has $\Delta_+ = \{(2,-1),(-2,2),(0,1),(2,0)\}$, hence $\rho = (1,1)$.
- $\mathfrak{g}_2$ has $\Delta_+ = \{(2,-3),(-1,2),(1,-1),(0,1),(-1,3),(1,0)\}$, hence $\rho = (1,1)$.
- $\mathfrak{so}(7)$ has

$$\begin{aligned} \Delta_+ = \{&(2,-1,0),(-1,2,-2),(0,-1,2),(1,1,-2), \\ &(-1,1,0),(1,0,0),(-1,0,2),(1,-1,2),(0,1,0)\}, \end{aligned} \tag{5.51}$$

  hence $\rho = (1,1,1)$.

It sure looks like $\rho$ is in fact the element $\sum_i \omega_i$ of $\mathsf{P}_\geqslant$. This can obviously be checked for all the classical Lie algebras by exhaustive and explicit computation, but it's much nicer to have an elegant proof.

**Exercise 104.** In this exercise, we show that

$$\frac{1}{2} \sum_{\alpha \in \Delta_+} \alpha = \rho = \sum_{i=1}^{\operatorname{rank} \mathfrak{g}} \omega_i. \tag{5.52}$$

**(a)** Explain why the only point fixed by all of the simple Weyl reflections $\mathsf{w}_{\alpha_i}$ is the origin.
**(b)** Compute $\mathsf{w}_{\alpha_i}(\alpha_j)$ for each $i, j$ and conclude that $\mathsf{w}_{\alpha_i}$ permutes the set $\Delta_+ \setminus \{\alpha_i\}$.
**(c)** Compute the action of each $\mathsf{w}_{\alpha_i}$ on $\rho = \frac{1}{2} \sum_{\alpha \in \Delta_+} \alpha$ and $\rho' = \sum_i \omega_i$.
**(d)** Conclude that $\rho = \rho'$. ▼

Of course, there's always more than one way to skin a cat. We shall encounter a second way to demonstrate (5.52) shortly (see Exercise 111 below).

## 5.6. The quadratic Casimir

In Section 3.2, we introduced the quadratic Casimir of $\mathfrak{sl}(2)$ as a tool to help prove Weyl's complete reducibility theorem (Theorem 3.5) for finite-dimensional $\mathfrak{sl}(2)$-modules. Recall that it was defined as a linear operator acting on an arbitrary $\mathfrak{sl}(2)$-module. Now that we have more sophisticated tools up our sleeves, we can redefine the quadratic Casimir as an element of its true home: the universal enveloping algebra of $\mathfrak{sl}(2)$. This also provides us with an opportunity to define the quadratic Casimir of a general semisimple Lie algebra $\mathfrak{g}$ as an element of $\mathsf{U}(\mathfrak{g})$, in preparation for proving Weyl's theorem for finite-dimensional $\mathfrak{g}$-modules (Theorem 5.23 below).

Recall from Exercise 15 that the centre $\mathfrak{z}(\mathfrak{g})$ of a Lie algebra $\mathfrak{g}$ is an abelian ideal, hence is 0 for all semisimple $\mathfrak{g}$. In a way, this is a pity because non-trivial central elements are extremely useful in representation theory. However, Exercise 86 shows that the representation theories of $\mathfrak{g}$ and $\mathsf{U}(\mathfrak{g})$ are identical and nothing to date has said anything about central elements of $\mathsf{U}(\mathfrak{g})$ (except multiples of the unit $\mathbb{1}$ which act predictably).

Let $\mathsf{Z}(\mathfrak{g})$ denote the centre of $\mathsf{U}(\mathfrak{g})$. This consists of the $Z \in \mathsf{U}(\mathfrak{g})$ such that $ZU = UZ$, for all $U \in \mathsf{U}(\mathfrak{g})$. As $\mathsf{U}(\mathfrak{g})$ is generated by $\mathfrak{g}$ (embedded into $\mathsf{U}(\mathfrak{g})$ by the canonical map of Theorem 5.5), it suffices to check that $Zx = xZ$ for all $x \in \mathfrak{g}$. Note that $\mathsf{Z}(\mathfrak{g})$ is an abelian (unital associative) subalgebra of $\mathsf{U}(\mathfrak{g})$.

Each $Z \in \mathsf{Z}(\mathfrak{g})$ then defines a $\mathfrak{g}$-module endomorphism on any $\mathfrak{g}$-module $V$:

$$xZv = Zxv, \quad \text{for all } x \in \mathfrak{g} \text{ and } v \in V. \tag{5.53}$$

(Here, we have identified $\mathfrak{g}$-modules with $\mathsf{U}(\mathfrak{g})$-modules for simplicity.) If $V$ happens to be irreducible and finite-dimensional, then each $Z \in \mathsf{Z}(\mathfrak{g})$ acts on $V$ as some multiple $\chi_V(Z)$ of the identity operator, by Schur's lemma (Lemma 3.3). Since $Z$ always acts linearly, it follows that $\chi_V$ is a linear functional on $\mathsf{Z}(\mathfrak{g})$, *ie.* $\chi_V \in \mathsf{Z}(\mathfrak{g})^*$. $\chi_V$ is called the *central character* of the $\mathfrak{g}$-module $V$.

If two $\mathfrak{g}$-modules are isomorphic, then they have the same central characters. The central character is thus an invariant of irreducible finite-dimensional $\mathfrak{g}$-modules and so may be used to distinguish them. It is often important to extend this to more general classes of $\mathfrak{g}$-modules. To this end, let us say that an arbitrary $\mathfrak{g}$-module *admits a central character* if every $Z \in \mathsf{Z}(\mathfrak{g})$ acts on it as a multiple of the identity.

**Exercise 105.** Prove if $V$ and $W$ are isomorphic $\mathfrak{g}$-modules that admit central characters, then their central characters coincide as elements of $\mathsf{Z}(\mathfrak{g})^*$. ▼

We remark that each $Z \in \mathsf{Z}(\mathfrak{g})$ has a single eigenvalue when acting on any finite-dimensional indecomposable $\mathfrak{g}$-module, by Lemma 3.2, but this does not mean that the module admits a central character because some of the $Z$ might act non-diagonalisably.

A $\mathfrak{g}$-module, not necessarily finite-dimensional or irreducible, is said to be *cyclic* if it is generated by a single element (*cf.* a cyclic group), called a *cyclic vector*. All irreducible modules are cyclic, with any non-zero vector qualifying as a cyclic vector. On the other hand, Verma modules (and highest-weight modules in general) are also examples of cyclic modules with the (generating) highest-weight vector as cyclic vector.

**Exercise 106.** Generalise Schur's lemma (Lemma 3.3) by proving the following result:

Let $V$ be a cyclic $\mathfrak{g}$-module whose cyclic vector is an eigenvector of a $\mathfrak{g}$-module endomorphism $Q$. Then, $Q$ acts as a multiple of the identity on $V$. ▼

A convenient fact for representation theory is that it isn't too difficult to construct a quadratic (*ie.* of tensor degree 2) element of $\mathsf{Z}(\mathfrak{g})$, if one has an non-degenerate invariant bilinear form $\kappa$. Let $\{x_i\}$ denote a basis of $\mathfrak{g}$ and let $\{y_j\}$ denote the dual basis with respect to $\kappa$:

(5.54) $$\kappa(x_i, y_j) = \delta_{ij}.$$

We now define the *quadratic Casimir* of $\mathfrak{g}$ to be the element

(5.55) $$Q = \sum_{i=1}^{\dim \mathfrak{g}} x_i y_i \in \mathsf{U}(\mathfrak{g}).$$

Recall that when $\mathfrak{g}$ is simple, every non-degenerate invariant bilinear form is proportional to the Killing form (Exercise 54). This is reflected in the (obvious) fact that $Q$ may be rescaled.

**Exercise 107.** Prove that $Q$ does not depend upon the choice of basis $\{x_i\}$. ▼

We remark that it is often convenient to rewrite the quadratic Casimir in the somewhat more explicit form

(5.56) $$Q = \sum_{i,j} \kappa^{-1}(x_i, x_j) x_i x_j,$$

where $\kappa^{-1}(\cdot, \cdot)$ denotes the bilinear form whose representing matrix (with respect to an arbitrary basis of $\mathfrak{g}$) is the inverse of that of $\kappa$.

**Example 58.** For $\mathfrak{g} = \mathfrak{sl}(2)$, we may choose the standard basis $\{e, h, f\}$ and take $\kappa$ to be the Killing form. We computed the matrix representative of $\kappa$, with respect to this basis, in Example 31, so the inverse is easily determined:

(5.57) $$\kappa = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 0 \end{pmatrix} \qquad \Rightarrow \qquad \kappa^{-1} = \begin{pmatrix} 0 & 0 & \frac{1}{4} \\ 0 & \frac{1}{8} & 0 \\ \frac{1}{4} & 0 & 0 \end{pmatrix}.$$

The quadratic Casimir of $\mathfrak{sl}(2)$ is therefore

(5.58) $$Q = \frac{1}{8} h^2 + \frac{1}{4} ef + \frac{1}{4} fe = \frac{1}{4} \left( \frac{1}{2} h^2 + ef + fe \right) \in \mathsf{U}(\mathfrak{sl}(2)).$$

Note that if we replace the Killing form by the trace form in the defining representation $\pi$ (see Example 22), then

$$\kappa_\pi(h,h) = \operatorname{tr}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^2 = 2 \tag{5.59}$$

shows that $\kappa_\pi = \frac{1}{4}\kappa$, hence $\kappa_\pi^{-1} = 4\kappa^{-1}$ and so $Q$ would be multiplied by 4. In this way, we recover the definition of the quadratic Casimir of $\mathfrak{sl}(2)$ used in Section 3.2 (but as an element of $\mathsf{U}(\mathfrak{sl}(2))$, rather than as a linear operator acting on an arbitrary $\mathfrak{sl}(2)$-module). ▲

**Exercise 108.** Compute the quadratic Casimirs of $\mathfrak{sl}(3)$ and $\mathfrak{sp}(4)$ explicitly, in both cases taking $\kappa$ to be the trace form in the defining representation. ▼

**Proposition 5.19.** *Let $\mathfrak{g}$ be a Lie algebra with a non-degenerate invariant bilinear form $\kappa$. Then, the quadratic Casimir is central*: $Q \in \mathsf{Z}(\mathfrak{g})$.

*Proof.* Choose dual bases $\{x_i\}$ and $\{y_i\}$ of $\mathfrak{g}$, with respect to $\kappa$. Given $z \in \mathfrak{g}$, let $(a_{ij})$ and $(b_{ij})$ denote the matrix representatives of $\operatorname{ad}(z)$ with respect to the $x_i$ and $y_i$, respectively:

$$[z,x_i] = \sum_j a_{ij}x_j \quad \text{and} \quad [z,y_i] = \sum_j b_{ij}y_j, \quad \text{for some } a_{ij}, b_{ij} \in \mathbb{C}. \tag{5.60}$$

The entries of these matrices are related by the invariance of $\kappa$:

$$b_{ij} = \sum_k b_{ik}\kappa(x_j,y_k) = \kappa(x_j,[z,y_i]) = -\kappa([z,x_j],y_i) = -\sum_k a_{jk}\kappa(x_k,y_i) = -a_{ji}. \tag{5.61}$$

With this relation, we now compute that for any $z \in \mathfrak{g}$,

$$\begin{aligned}[z,Q] &= \sum_i [z,x_iy_i] = \sum_i ([z,x_i]y_i + x_i[z,y_i]) = \sum_{i,j}(a_{ij}x_jy_i + b_{ij}x_iy_j) \\ &= \sum_{i,j}(a_{ji}+b_{ij})x_iy_j = 0.\end{aligned} \tag{5.62}$$

■

**Exercise 109.** Prove that the eigenvalue of $Q$ on the adjoint module of a simple Lie algebra is 1. Be careful not to fall into the trap of thinking that $\operatorname{ad}(Q)x = [Q,x] = 0$. [*Hint: what happens if you take the trace of* $\operatorname{ad}(Q)$*?*] ▼

It is interesting to note that when $\operatorname{rank}\mathfrak{g} > 1$, *ie.* when $\mathfrak{g} \not\cong \mathfrak{sl}(2)$, there are other algebraically independent elements in $\mathsf{Z}(\mathfrak{g})$ (meaning elements that are not polynomials in $\mathbb{1}$ and $Q$). For example, $\mathfrak{sl}(3)$ has a cubic Casimir that plays an occasional role in quantum chromodynamics calculations. In general, $\mathsf{Z}(\mathfrak{g})$ is a polynomial algebra in $\operatorname{rank}\mathfrak{g}$ algebraically independent generators, known as the *Casimir elements* of $\mathfrak{g}$. Their degrees $d_i$ (as polynomials) define the *exponents* $e_i$ of $\mathfrak{g}$ by $d_i = e_i + 1$. These exponents appear mysteriously in such places as heights of roots, eigenvalues of Dynkin diagram adjacency matrices and even the cardinality of the Weyl group.

Be that as it may, one of the first things we did in Section 3.2 with the quadratic Casimir of $\mathfrak{sl}(2)$ was to explicitly determine its eigenvalues on the finite-dimensional irreducible $\mathfrak{sl}(2)$-modules $\mathcal{L}_\lambda$, $\lambda \in \mathbb{Z}_{\geqslant 0}$. We did this by acting on an $h$-eigenvector (a weight vector) which was annihilated by $e$ (so a highest-weight vector), see Equation (3.18). Our next task is to generalise this to the quadratic Casimir of an arbitrary semisimple Lie algebra $\mathfrak{g}$ and evaluate its action on an arbitrary highest-weight vector. For this, we recall the Weyl vector $\rho \in \mathfrak{g}_0^*$, defined in (5.50) as half the sum of the positive roots.

**Proposition 5.20.** *The eigenvalue of the quadratic Casimir $Q$ on a highest-weight vector of weight $\lambda$ is $(\lambda, \lambda + 2\rho)$.*

*Proof.* We work in the basis of simple coroots $\alpha_i^\vee$, $i = 1, \dots, \operatorname{rank} \mathfrak{g}$, and root vectors $e_\alpha$ and $f_\alpha$, $\alpha \in \Delta_+$. As the Killing form is block-diagonal in this basis (Lemma 4.7), finding the dual basis is straightforward. First, note that the dual basis of $\{e_\alpha, f_\alpha\}$ is

$$\left\{ \frac{\lVert\alpha\rVert^2}{2} f_\alpha, \frac{\lVert\alpha\rVert^2}{2} e_\alpha \right\}, \tag{5.63}$$

by Equation (4.53). Moreover, (5.5) gives

$$\kappa(\iota^{-1}(\omega_i), \alpha_j^\vee) = \omega_i(\alpha_j^\vee) = \delta_{ij}, \tag{5.64}$$

hence the basis of $\mathfrak{g}_0$ dual to the simple coroots $\alpha_i^\vee$ is given by the $\iota^{-1}(\omega_i)$, $i = 1, \dots, \operatorname{rank} \mathfrak{g}$. The dual basis of $\mathfrak{g}$ is thus given by the $\iota^{-1}(\omega_i)$ and the elements of (5.63), for $\alpha \in \Delta_+$.

Let the calculation begin! We substitute these bases into the definition (5.55) of $Q$:

$$\begin{aligned} Q &= \sum_i \alpha_i^\vee \iota^{-1}(\omega_i) + \sum_{\alpha \in \Delta_+} \frac{\lVert\alpha\rVert^2}{2} (e_\alpha f_\alpha + f_\alpha e_\alpha) \\ &= \sum_i \alpha_i^\vee \iota^{-1}(\omega_i) + \sum_{\alpha \in \Delta_+} \frac{\lVert\alpha\rVert^2}{2} \alpha^\vee + \sum_{\alpha \in \Delta_+} \lVert\alpha\rVert^2 f_\alpha e_\alpha \\ &= \sum_i \alpha_i^\vee \iota^{-1}(\omega_i) + \sum_{\alpha \in \Delta_+} \iota^{-1}(\alpha) + \sum_{\alpha \in \Delta_+} \lVert\alpha\rVert^2 f_\alpha e_\alpha \\ &= \sum_i \alpha_i^\vee \iota^{-1}(\omega_i) + 2\iota^{-1}(\rho) + \sum_{\alpha \in \Delta_+} \lVert\alpha\rVert^2 f_\alpha e_\alpha. \end{aligned} \tag{5.65}$$

Evaluating $Q$ on a highest-weight vector $v$ of weight $\lambda$, we note that the third term gives 0 and we are left with

$$\begin{aligned} Qv &= \left[ \sum_i \lambda(\alpha_i^\vee) \lambda\bigl(\iota^{-1}(\omega_i)\bigr) + 2\lambda\bigl(\iota^{-1}(\rho)\bigr) \right] v = \left[ \sum_i \lambda_i (\lambda, \omega_i) + 2(\lambda, \rho) \right] v \\ &= \left[ (\lambda, \sum_i \lambda_i \omega_i) + (\lambda, 2\rho) \right] v = (\lambda, \lambda + 2\rho) v. \end{aligned} \tag{5.66}$$

■

**Corollary 5.21.** *If $V$ is an indecomposable $\mathfrak{g}$-module with two singular vectors (highest-weight vectors) of weights $\lambda$ and $\mu$, then $\|\lambda + \rho\|^2 = \|\mu + \rho\|^2$.*

**Exercise 110.** Combine this result with Proposition 5.14 to conclude that the Verma module $\mathcal{V}_{-\rho}$ is irreducible. ▼

**Exercise 111.** Recall from Proposition 5.13 that $\mathcal{V}_\lambda$ has a Verma submodule isomorphic to $\mathcal{V}_{\lambda-(\lambda_j+1)\alpha_j}$ whenever $\lambda_j \in \mathbb{Z}_{\geqslant 0}$. Combine this with Exercise 106 and Proposition 5.20 to conclude that the Dynkin labels of the Weyl vector $\rho$ are all 1. ▼

**Exercise 112.**

**(a)** Show that the irreducible highest-weight module $\mathcal{L}_0$ is 1-dimensional and trivial.
**(b)** Prove that the quadratic Casimir $Q$ acts as zero on a finite-dimensional irreducible $\mathfrak{g}$-module $V$ if and only if $V$ is trivial. ▼

Of course, $Q$ would also act as 0 on any irreducible $\mathfrak{g}$-module with $\lambda = -2\rho$ as a highest weight, but such a module is not finite-dimensional (*cf.* Exercise 90).

We recall that a simple Lie algebra possesses a unique highest root, traditionally denoted by $\theta$. This was mentioned in Exercise 80, but now follows easily from Example 24 and Proposition 5.6. There is a standard convention for simple Lie algebras, mentioned several times already, which normalises the highest root so that $\|\theta\|^2 = 2$. As we have discussed, this may be achieved by rescaling the Killing form (*cf.* Exercise 80).

From now on, the bilinear forms $\kappa(\cdot,\cdot)$ and $(\cdot,\cdot)$, as well as the isomorphism $\iota\colon \mathfrak{g}_0 \to \mathfrak{g}_0^*$ will be **rescaled** so that $\|\theta\|^2 = 2$. When we wish to refer to the original Killing form and the form it induces on $\mathfrak{g}_0^*$, we shall use the notation $\kappa_{\text{ad}}(\cdot,\cdot)$ and $(\cdot,\cdot)_{\text{ad}}$ instead. Similarly, the original isomorphism shall now be denoted by $\iota_{\text{ad}}$.

**Corollary 5.22.** *The convention $\|\theta\|^2 = 2$ corresponds to rescaling the Killing form $\kappa_{\text{ad}}$ of $\mathfrak{g}$ to*

$$\kappa = \frac{1}{2\mathsf{h}^\vee}\kappa_{\text{ad}}, \tag{5.67}$$

*where $\mathsf{h}^\vee = 1 + (\theta,\rho)$ is called the* dual Coxeter number *of $\mathfrak{g}$.*

*Proof.* We start by noting that setting $\kappa = \zeta\kappa_{\text{ad}}$, for some non-zero $\zeta \in \mathbb{C}$, implies that $\iota = \zeta\iota_{\text{ad}}$ and so $(\cdot,\cdot) = \zeta^{-1}(\cdot,\cdot)_{\text{ad}}$, see Equation (4.41) and Exercise 65. Imposing $(\theta,\theta) = 2$ then requires that

$$2 = \zeta^{-1}(\theta,\theta)_{\text{ad}} = \zeta^{-1}(1 - 2(\theta,\rho)_{\text{ad}}) = \zeta^{-1} - 2(\theta,\rho) \quad \Rightarrow \quad \zeta^{-1} = 2(1 + (\theta,\rho)), \tag{5.68}$$

because Exercise 109 and Proposition 5.20 give $(\theta,\theta)_{\text{ad}} + 2(\theta,\rho)_{\text{ad}} = 1$. ■

Given $\|\theta\|^2 = 2$, as above, we can explicitly compute the value of the dual Coxeter number of each simple Lie algebra by writing $\theta = \sum_i a_i \alpha_i$ or $\theta^\vee = \sum_i a_i^\vee \alpha_i^\vee$ and noting that

$$\mathsf{h}^\vee = 1 + \sum_{i=1}^{\operatorname{rank}\mathfrak{g}} a_i^\vee = 1 + \sum_{i=1}^{\operatorname{rank}\mathfrak{g}} \frac{\|\alpha_i\|^2}{2} a_i. \tag{5.69a}$$

The $a_i$ and $a_i^\vee$ are called the *marks* and *comarks* of $\mathfrak{g}$, respectively. These numbers play a passing role in the theory of semisimple Lie algebras, but really come into their own when one generalises to the closely related (infinite-dimensional) affine Kac–Moody algebras that arise in mathematical physics (conformal field theory, string theory, *etc.*).

The name "dual Coxeter number" suggests that there is also a Coxeter number $\mathsf{h}$ associated to each simple Lie algebra $\mathfrak{g}$. There is, of course, and it is defined by

$$\mathsf{h} = 1 + \sum_i a_i = 1 + \sum_{i=1}^{\operatorname{rank}\mathfrak{g}} \frac{2}{\|\alpha_i\|^2} a_i^\vee, \tag{5.69b}$$

*ie.* it is one plus the height of $\theta$, *cf.* (4.94). The Coxeter number also turns out to be one plus the largest exponent of $\mathfrak{g}$ and is equal to the number of roots divided by the rank. Unlike the dual Coxeter number, the Coxeter number is secretly an invariant of the Weyl group of $\mathfrak{g}$ (the fact that Weyl groups are examples of Coxeter groups gives a hint about the naming).

We tabulate the Coxeter and dual Coxeter numbers of the simple Lie algebras for convenience.

| $\mathfrak{g}$ | $\mathfrak{sl}(r+1)$ | $\mathfrak{so}(2r+1)$ | $\mathfrak{sp}(2r)$ | $\mathfrak{so}(2r)$ | $\mathfrak{e}_6$ | $\mathfrak{e}_7$ | $\mathfrak{e}_8$ | $\mathfrak{f}_4$ | $\mathfrak{g}_2$ |
|---|---|---|---|---|---|---|---|---|---|
| $\mathsf{h}$ | $r+1$ | $2r$ | $2r$ | $2r-2$ | 12 | 18 | 30 | 12 | 6 |
| $\mathsf{h}^\vee$ | $r+1$ | $2r-1$ | $r+1$ | $2r-2$ | 12 | 18 | 30 | 9 | 4 |

Of course, they coincide if $\mathfrak{g}$ is simply laced (*ie.* its Dynkin diagram has no arrows). You might wish to compare the quoted values of $\mathsf{h}^\vee$ with your answers for Exercise 80.

We conclude by noting a convenient consequence of setting $\|\theta\|^2 = 2$: the ratio $\frac{2}{\|\alpha\|^2}$ that appears in the definition (4.54) of the coroot $\alpha^\vee$ is now constrained to lie in the set $\{1, 2, 3\}$, by Proposition 4.16. Indeed, if $\mathfrak{g}$ is simply laced, then all roots have the same length and so this ratio is always 1.

To see an example of this computational convenience, suppose that we had need to explicitly compute $(\lambda, \mu)$, for $\lambda, \mu \in \mathfrak{g}_0$, directly from the Dynkin labels of $\lambda$ and $\mu$. If we let $\vec{\lambda}$ and $\vec{\mu}$ denote the column vectors consisting of the Dynkin labels of $\lambda$ and $\mu$, respectively, then we would have

$$(\lambda, \mu) = \sum_{i,j=1}^{\operatorname{rank}\mathfrak{g}} \lambda_i (\omega_i, \omega_j) \mu_j = \vec{\lambda}^{\mathsf{T}} \mathsf{G} \vec{\mu}, \tag{5.70}$$

where $\mathsf{G}$ is the matrix whose $(i, j)$-th entry is $(\omega_i, \omega_j)$. It is thus the *Gram matrix* of the bilinear form $(\cdot, \cdot)$ in the basis of fundamental weights. Because this basis is dual to the basis of $\mathfrak{g}_0$ consisting of the simple coroots, $\mathsf{G}$ is the inverse of the Gram matrix $\mathring{\mathsf{A}}$ of $\kappa(\cdot,\cdot)$ in this simple coroot basis. However, $\mathring{\mathsf{A}}$ is just a symmetrisation of the Cartan matrix $\mathsf{A}$ obtained by multiplying each column by an appropriate integer:

$$\mathring{\mathsf{A}}_{ij} = \kappa(\alpha_i^\vee, \alpha_j^\vee) = \frac{2}{\|\alpha_j\|^2}\alpha_j(\alpha_i^\vee) = \frac{2}{\|\alpha_j\|^2}\mathsf{A}_{ij}. \tag{5.71}$$

Having small integers for the ratios $\frac{2}{\|\alpha\|^2}$ is certainly nice for such explicit computations.

**Example 59.** Because $\alpha_1$ is a short root of $\mathfrak{sp}(4)$, whilst $\alpha_2$ is long, the symmetrisation of the Cartan matrix is obtained by multiplying the first column by $\frac{2}{\|\alpha_1\|^2} = 2$:

$$\mathsf{A} = \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix} \quad \Rightarrow \quad \mathring{\mathsf{A}} = \begin{pmatrix} 4 & -2 \\ -2 & 2 \end{pmatrix} \quad \Rightarrow \quad \mathsf{G} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}. \tag{5.72}$$

It is thus easy to compute, for example, that

$$\begin{aligned}
\|\alpha_1\|^2 &= \vec{\alpha}_1{}^{\mathsf{T}}\mathsf{G}\vec{\alpha}_1 = \begin{pmatrix} 2 & -1 \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \end{pmatrix} = 1, \\
(\alpha_1,\alpha_2) &= \vec{\alpha}_1{}^{\mathsf{T}}\mathsf{G}\vec{\alpha}_2 = \begin{pmatrix} 2 & -1 \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} -2 \\ 2 \end{pmatrix} = -1, \\
\|\alpha_2\|^2 &= \vec{\alpha}_2{}^{\mathsf{T}}\mathsf{G}\vec{\alpha}_2 = \begin{pmatrix} -2 & 2 \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} -2 \\ 2 \end{pmatrix} = 2,
\end{aligned} \tag{5.73}$$

hence that the angle between the two simple roots is 135°, as in Example 37. ▲

**Exercise 113.** Compute the Gram matrix $\mathsf{G}$ for each of the classical simple Lie algebras. ▼

## 5.7. Weyl's theorem: complete reducibility

It's finally time to generalise Theorem 3.5 from $\mathfrak{sl}(2)$ to a general semisimple Lie algebra $\mathfrak{g}$ and thereby demonstrate Weyl's theorem: that all finite-dimensional $\mathfrak{g}$-modules are completely reducible. The analogous statement for finite groups, or rather the corresponding complex group algebras, is called Maschke's Theorem and it is proved using an averaging trick. Weyl's original proof copied this by setting up the same averaging trick, but on the (connected, simply-connected) compact Lie group whose Lie algebra is $\mathfrak{g}$. We don't have the luxury of introducing integration on compact Lie groups, nor to discuss averaging with respect to their invariant (Haar) measures. There is a purely algebraic proof due to Brauer. However, it is rather involved, and we are somewhat pressed for time, so we'll omit it.

To hint at the difficulties involved, recall from Section 3.2 that in proving Weyl's theorem for $\mathfrak{sl}(2)$, we exploited the fact that the eigenvalue of the quadratic Casimir completely distinguished the finite-dimensional irreducible $\mathfrak{sl}(2)$-modules, up to isomorphism. Unfortunately, the same is not true for general $\mathfrak{g}$, so one has to work much harder.

**Exercise 114.** Show that the eigenvalue of the quadratic Casimir on the defining $\mathfrak{sl}(3)$-module coincides with that on the dual of the defining module. Explain why this defining module and its dual cannot be isomorphic. ▼

In fact, it turns out that finite-dimensional irreducible $\mathfrak{g}$-modules are completely distinguished by their central characters, suggesting yet another algebraic path to generalise the proof of Weyl's theorem we gave for $\mathfrak{sl}(2)$. However, this requires delving much more deeply into the beautiful world of semisimple Lie algebras than we have time for (keyword: Harish-Chandra homomorphism). We shall therefore have to content ourselves with just stating Weyl's theorem.

**Theorem 5.23** (Weyl)**.** *Every finite-dimensional module $V$ of a (finite-dimensional complex) semisimple Lie algebra $\mathfrak{g}$ is completely reducible.*

It is perhaps a little easy to overlook the importance of Weyl's Theorem in the theoretical development that we are following. So let us emphasise that the complete reducibility of the finite-dimensional $\mathfrak{g}$-modules is the best news that we could possibly have hoped for, both on theoretical and practical grounds. When the representations of a given algebraic structure are not completely reducible, then life becomes significantly more difficult very quickly. In fact, it can be shown (in a rigorous sense) that one will usually not be able to classify the modules. And without a solid formalism for working with representations, one may well be left with inefficient "brute-force" methods to tackle practical calculations.

To give a sense for how Weyl's theorem can simplify calculations, we present some examples in the next section that detail how one can decompose the tensor product of irreducible $\mathfrak{g}$-modules into irreducible $\mathfrak{g}$-modules. The method presented is not particularly efficient — much better methods are available, particularly when $\mathfrak{g}$ is classical (names include Clebsch–Gordan, Littlewood–Richardson and Schur–Weyl) — but it is illustrative. And would be much much harder if we didn't know that the result was completely reducible.

We conclude this section with a simple theoretical application of Weyl's theorem to the relationship between irreducibility and finite-dimensionality.

**Lemma 5.24.** *Every finite-dimensional highest-weight module of highest weight $\lambda$ is isomorphic to the irreducible $\mathcal{L}_\lambda$.*

*Proof.* Every highest-weight module with highest weight $\lambda$ is isomorphic to a (non-zero) quotient of the Verma module $\mathcal{V}_\lambda$, by Proposition 5.11. Being finite-dimensional, this quotient is a direct sum of irreducible highest-weight modules, by Weyl's theorem. As $\lambda$ has multiplicity 1 in $\mathcal{V}_\lambda$ (Exercise 97), it has multiplicity 1 in the quotient as well. It follows that only one of the irreducibles in the direct sum decomposition may have $\lambda$ as a weight. However, the quotient is clearly generated by the image of the highest-weight vector $v \in \mathcal{V}_\lambda$ of weight $\lambda$, so the direct sum can only be this irreducible module, *ie.* the quotient is irreducible. As it is a quotient of $\mathcal{V}_\lambda$, it is therefore $\mathcal{L}_\lambda$, by Corollary 5.12. ■

This simple observation has a rather neat application to the identification of the maximal proper submodule $\mathcal{J}_\lambda$ of $\mathcal{V}_\lambda$, *cf.* Section 5.4, when $\lambda \in \mathsf{P}_\geqslant$. For this, we need two additional ingredients. First, we need the following identity (the proof is an exercise of course).

**Exercise 115.** Consider an associative algebra in which we define linear endomorphisms $\mathrm{ad}(U)$ by $\mathrm{ad}(U)V = [U, V] = UV - VU$ (note that this does not make ad into an associative algebra homomorphism). Prove the following identity for all positive integers $n$:

$$[U^n, V] = \sum_{k=1}^{n} \binom{n}{k} \mathrm{ad}(U)^k V \cdot U^{n-k}. \tag{5.74}$$ ▼

The second ingredient is to remark that as $\mathfrak{g}$ is finite-dimensional, each $\mathrm{ad}(e_\alpha)$, $\alpha \in \Delta$, defines a *nilpotent* endomorphism of the adjoint module. There are only finitely many roots, so for $\alpha, \beta \in \Delta$, we must have $\beta + n\alpha \notin \Delta$ for all $n$ sufficiently large (*ie.* root strings are finite), hence $\mathrm{ad}(e_\alpha)^n e_\beta = 0$ for such $n$ (*cf.* Serre's theorem — Theorem 4.24 — for the simple root case). As there are only finitely many roots, we can therefore choose $N$ so that

$$\mathrm{ad}(e_\alpha)^N e_\beta = 0, \quad \text{for } all\ \alpha, \beta \in \Delta. \tag{5.75}$$

**Theorem 5.25.** *For $\lambda \in \mathsf{P}_\geqslant$, $\mathcal{J}_\lambda$ is generated by the singular vectors $f_{\alpha_i}^{\lambda_i+1} v$, where $v \in \mathcal{V}_\lambda$ is the highest-weight vector of weight $\lambda$ and $i = 1, \ldots, \operatorname{rank} \mathfrak{g}$. In other words,*

$$\mathcal{L}_\lambda \simeq \frac{\mathcal{V}_\lambda}{\sum_{i=1}^{\operatorname{rank} \mathfrak{g}} \mathcal{V}_{\lambda-(\lambda_i+1)\alpha_i}}. \tag{5.76}$$

*Proof.* The aim is to show that the (obviously non-zero) quotient on the right-hand side of (5.76) is finite-dimensional. Then, it is isomorphic to $\mathcal{L}_\lambda$, by Lemma 5.24.

To demonstrate this finite-dimensionality, we show that the quotient is a sum of finite-dimensional $\mathfrak{sl}(2)_i$-modules, for each $i = 1, \ldots, \operatorname{rank} \mathfrak{g}$, as in the proof of Lemma 5.15. Its weights are thus permuted by $\mathsf{W}$ and so the finite-dimensionality follows as in the proof of Theorem 5.17. Here, $\mathfrak{sl}(2)_i$ denotes the subalgebra of $\mathfrak{g}$ spanned by $e_{\alpha_i}$, $h_{\alpha_i}$ and $f_{\alpha_i}$. To demonstrate that the $\mathfrak{sl}(2)_i$-modules are finite-dimensional, we will show that for each vector $w$ in a Poincaré–Birkhoff–Witt basis of $\mathcal{V}_\lambda$, there exists $n$ such that $f_{\alpha_i}^n \overline{w}$ is zero in the quotient. Since the corresponding statement with $f_{\alpha_i}$ replaced by $e_{\alpha_i}$ is always true (because the quotient is a highest-weight module), this will show that each basis vector generates a finite-dimensional $\mathfrak{sl}(2)_i$-module.

Consider therefore a Poincaré–Birkhoff–Witt basis vector $w = f_{\beta_m} \cdots f_{\beta_1} v \in \mathcal{V}_\lambda$, where $\beta_1, \ldots, \beta_m \in \Delta_+$. We show that $\overline{w}$ is annihilated by a sufficiently high power of $f_{\alpha_i}$ by induction on $m$. If $m = 0$, then $\overline{w} = \overline{v}$ is annihilated by the $f_{\alpha_i}^{\lambda_i+1}$, by definition of the quotient module. So, we may assume that there exists $n$ such that $f_{\alpha_i}^n$ annihilates $f_{\beta_\ell} \cdots f_{\beta_1} \overline{v}$, for all $\ell < m$. It follows then that

$$f_{\alpha_i}^{n+N} \overline{w} = f_{\alpha_i}^{n+N} f_{\beta_m} \cdots f_{\beta_1} \overline{v} = \sum_{j=1}^{m} f_{\beta_m} \cdots [f_{\alpha_i}^{n+N}, f_{\beta_j}] \cdots f_{\beta_1} \overline{v} \tag{5.77}$$

$$= \sum_{j=1}^{m} f_{\beta_m} \cdots \sum_{k=1}^{n+N} \binom{n+N}{k} \operatorname{ad}(f_{\alpha_i})^k f_{\beta_j} \cdot f_{\alpha_i}^{n+N-k} f_{\beta_{j-1}} \cdots f_{\beta_1} \overline{v}, \tag{5.78}$$

by (5.74). However, (5.75) now truncates the second sum so that

$$f_{\alpha_i}^{n+N} \overline{w} = \sum_{j=1}^{m} f_{\beta_m} \cdots \sum_{k=1}^{N-1} \binom{n+N}{k} \operatorname{ad}(f_{\alpha_i})^k f_{\beta_j} \cdot f_{\alpha_i}^{n+N-k} f_{\beta_{j-1}} \cdots f_{\beta_1} \overline{v}. \tag{5.79}$$

As $n + N - k > n$ and $j - 1 < m$, $f_{\alpha_i}^{n+N-k} f_{\beta_{j-1}} \cdots f_{\beta_1} \overline{v} =$ by the induction hypothesis. This completes the induction and thus also the proof. ■

## 5.8. **An application to quantum field theory**

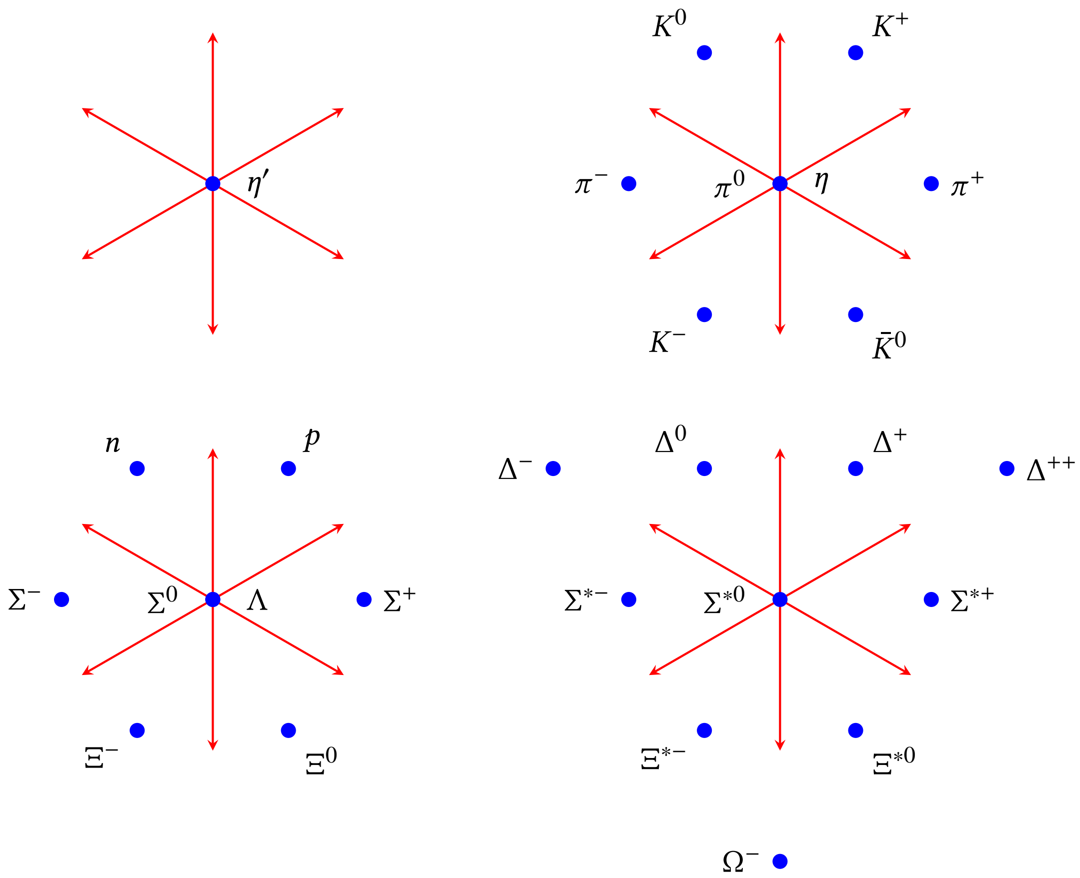
One *raison d'être* of quantum field theory is to explain the fundamental features of particle physics. In the first half of the twentieth century, physicists were quite happy with their particles. They had convincingly demonstrated that atoms were composed of electrons, protons and neutrons and, aside from predictions of such beasties as neutrinos (by Pauli) and pions (by Yukawa) and the inconvenient discovery of the muon (Rabi: "Who ordered that?"), that seemed to be that.

However, the late forties and fifties saw an unprecedented increase in funding for fundamental physics, probably as a result of the somewhat uncomfortable role it played in the second world war. This led to the discovery, first in cosmic rays and then later in cyclotrons and other accelerators, of three different pions, four completely new particles called kaons, as well as the eta meson and the lambda baryon, among others.

Even while physicists were failing to come up with cool names for all these unexpected new particles, efforts were underway to measure their properties and bring some semblance of order to them. As it turned out, a significant part of this order derived from the finite-dimensional representation theory of the complex Lie algebra $\mathfrak{sl}(3)$ (which is the same as the finite-dimensional representation theory of the real Lie algebra $\mathfrak{su}(3)$ and the compact real Lie group $\mathsf{SU}(3)$).

Physicists had long been accustomed to distinguishing elementary particles by their *isospins*, a term introduced by Heisenberg to quantify (in terms of symmetry) the difference between protons and neutrons. The name "isospin" was chosen to reflect its mathematical formulation in terms of finite-dimensional $\mathfrak{sl}(2)$-modules. For example, protons and neutrons together span a 2-dimensional fundamental $\mathfrak{sl}(2)$-module with the proton representing the quantum state of isospin $+\frac{1}{2}$ and the neutron being its isospin $-\frac{1}{2}$ partner (recall that physicists define quantum spin as half the $h$-eigenvalue). To understand the rapidly growing "particle zoo" being discovered, physicists then added a second scalar quantity to distinguish particles which they imaginatively called the *hypercharge*.

Two scalar eigenvalues means, when interpreted in the language of semisimple Lie theory, a 2-dimensional Cartan subalgebra. Indeed, Gell-Mann and Ne'eman independently realised in the early sixties that the particle zoo could be naturally described in terms of certain representations of $\mathfrak{sl}(3)$. In particular, they found that the operators whose eigenvalues gave the isospin and hypercharge could be assigned to the (orthogonal) Cartan elements $\frac{1}{2}h_{\alpha_1}$ and $\frac{1}{3}(h_{\alpha_1} + 2h_{\alpha_2}) = \iota^{-1}(\omega_2)$, respectively. We indicate where the isospins and hypercharges of these fundamental particles fall on our familiar pictures of the (rational or real) inner product space $\mathsf{R} \subset \mathfrak{sl}(3)^*_0$.



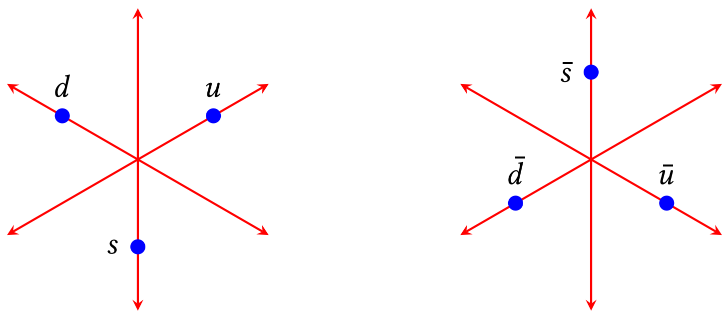

Here, $p$ and $n$ denote the proton and the neutron, respectively. You can guess the unimaginative names of the other particles or look them up. Note that there are also distinct antiparticles for the particles in the second row that lead to $\mathfrak{sl}(3)$-modules that are the duals of those indicated. Note also that the fanciful notion of hypercharge is actually somewhat redundant as one can replace it by the electric charge of the particle, given by the eigenvalue of $\iota^{-1}(\omega_1) = \frac{1}{3}(2h_{\alpha_1} + h_{\alpha_2})$, which is not orthogonal to isospin.

This observed (but by no means understood) $\mathfrak{sl}(3)$-symmetry is called *flavour symmetry* by physicists to distinguish it from *colour symmetry* (which is also associated with $\mathfrak{sl}(3)$).

Because the masses of the particles in the diagrams can be quite different, it is an "approximate" or "broken" symmetry of nature, at least at the energy scales that our experiments can access. Nevertheless, this flavour symmetry allowed Gell-Mann to predict the existence of the $\Omega^-$ particle, which at the time had not been discovered, and its mass. Its subsequent discovery was the confirmation of the role of $\mathfrak{sl}(3)$ in particle physics, earning him the 1969 Nobel prize in physics. Some refer to this classification scheme, which Gell-Mann dubbed the "eightfold way" after a Buddhist doctrine (remember this was the sixties), as the periodic table of particle physics.

Aside from the obvious question of why nature chose $\mathfrak{sl}(3)$ for this duty, a natural question to ask is why do we observe these particular representations and not some completely different ones. Gell-Mann had an answer for that as well (also independently proposed by Zweig at the same time). The proposed answer was that these fundamental particles are not so fundamental after all, but are actually composites constructed out of new (and unobserved) fundamental particles called *quarks* and *antiquarks*.

This proposal turned out to be mathematically very satisfying because the quarks were assigned to the first fundamental $\mathfrak{sl}(3)$-module $\mathcal{L}_{\omega_1}$, also known as the defining module, whilst the antiquarks were assigned to its dual, the second fundamental $\mathfrak{sl}(3)$-module $\mathcal{L}_{\omega_2}$. Quarks and antiquarks thus come in three flavours, called *up* ($u$), *down* ($d$) and *strange* ($s$), that are interchanged by the (approximate) $\mathfrak{sl}(3)$ flavour symmetry. Their different masses are (partly) responsible for the different masses observed for the particles.




One may therefore form composite particles, like protons and neutrons, by combining either a quark with an antiquark, by combining three quarks or by combining three antiquarks. As quarks are fermions, the first combination gives bosons (these are called *mesons*) whilst the second gives fermions (called *baryons*). These empirical rules led to the postulate of colour symmetry wherein each quark/antiquark is also assigned a "colour"/"anticolour" such that only colourless combinations may be combined. This is the basis of quantum chromodynamics.

Finally, we come to the mathematics. The representation-theoretic interpretation of "combining quarks and antiquarks" to get mesons is taking the **tensor product** of the fundamental $\mathfrak{sl}(3)$-modules $\mathcal{L}_{\omega_1}$ and $\mathcal{L}_{\omega_2}$. Since both have dimension 3, being the defining module and its dual, respectively, this tensor product is 9-dimensional. By Weyl's theorem, it decomposes as a direct sum of irreducibles and, by Corollary 5.7, the highest weight of every irreducible summand is dominant integral.

If you did Exercise 82, you would know that the weights of $\mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_2}$ are obtained by adding a weight of $\mathcal{L}_{\omega_1}$ to a weight of $\mathcal{L}_{\omega_2}$. The weights of $\mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_2}$ are thus as follows.

| $+$ | $(0,1)$ | $(1,-1)$ | $(-1,0)$ |
|---|---|---|---|
| $(1,0)$ | $(1,1)$ | $(2,-1)$ | $(0,0)$ |
| $(-1,1)$ | $(-1,2)$ | $(0,0)$ | $(-2,1)$ |
| $(0,-1)$ | $(0,0)$ | $(1,-2)$ | $(-1,-1)$ |

Since the simple roots are $\alpha_1 = (2,-1)$ and $\alpha_2 = (-1,2)$, we see that $\theta = (1,1)$ is a highest weight of $\mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_2}$ (because $(1,1)+(2,-1) = (3,0)$ and $(1,1)+(-1,2) = (0,3)$ are not weights). Therefore $\mathcal{L}_\theta$ is a direct summand of $\mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_2}$.

The only other dominant integral weight of $\mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_2}$ is 0, which appears with multiplicity 3. However, this does not mean that we have 3 copies of $\mathcal{L}_0$. Indeed, $\mathcal{L}_\theta$ is the adjoint module, hence $\dim(\mathcal{L}_\rho \oplus 3\mathcal{L}_0) = 8+3\times 1 = 11 > \dim \mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_2}$. This calculation makes it clear that there is only room for one copy of $\mathcal{L}_0$ and this is consistent with the fact that $\mathcal{L}_\rho$ already has 0 as a weight with multiplicity equal to rank $\mathfrak{sl}(3) = 2$. We therefore conclude that

$$\mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_2} \simeq \mathcal{L}_\theta \oplus \mathcal{L}_0. \tag{5.80}$$

You will have noticed that the two irreducibles on the right-hand side are precisely those appearing in the two top-row particle diagrams above.

**Exercise 116.** Similarly decompose $\mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_1}$ into irreducibles by:

- **(a)** Showing that $2\omega_1$ is a highest weight of $\mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_1}$.
- **(b)** Using a Poincaré–Birkhoff–Witt basis argument to prove that the multiplicity of $\omega_2$ in the Verma module $\mathcal{V}_{2\omega_1}$ is 1.
- **(c)** Explaining why the multiplicity of $\omega_2$ in $\mathcal{L}_{2\omega_1}$ is therefore also 1.
- **(d)** Explaining why this implies that $\mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_1} \simeq \mathcal{L}_{2\omega_1} \oplus \mathcal{L}_{\omega_2}$.

Follow similar steps to show that $\mathcal{L}_{\omega_1} \otimes \mathcal{L}_{2\omega_1} \simeq \mathcal{L}_{3\omega_1} \oplus \mathcal{L}_\theta$. ▼

To obtain the representation-theoretic content of the baryons, we therefore need to decompose the triple tensor product $\mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_1}$. Given the results of Equation (5.80) and Exercise 116, and recalling that tensor products distribute over direct sums, we get

$$\mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_1} \simeq \mathcal{L}_{\omega_1} \otimes \left(\mathcal{L}_{2\omega_1} \oplus \mathcal{L}_{\omega_2}\right) \simeq \mathcal{L}_{3\omega_1} \oplus 2\mathcal{L}_\theta \oplus \mathcal{L}_0. \tag{5.81}$$

However, we only observe baryons in the 10-dimensional module $\mathcal{L}_{3\omega_1}$ and one of the 8-dimensional $\mathcal{L}_\theta$ modules. The reason for this ultimately boils down to Pauli's exclusion principle: quarks are fermions so the wavefunction of the baryons should be fully antisymmetric under exchanging quarks. But this antisymmetry gets contributions from $\mathfrak{sl}(3)$

colour, $\mathfrak{sl}(3)$ flavour and $\mathfrak{sl}(2)$ spin factors, the final result being that the flavour $\mathcal{L}_0$ and one of the $\mathcal{L}_\theta$ modules fail the exclusion principle.

Of course, physicists have also discovered many more particles that do not fit into this simple $\mathfrak{sl}(3)$ picture. To accommodate these, we now believe that there are in fact six quarks rather than three. The newbies $c$ (*charm*), $b$ (*bottom*) and $t$ (*top*) have been experimentally discovered. However, extending $\mathfrak{sl}(3)$ to $\mathfrak{sl}(6)$ (or something similar) to accommodate them does not seem to be useful. Even the experimental data involving only $u$, $d$, $s$ and $c$, corresponding to a putative $\mathfrak{sl}(4)$, suggests that this symmetry is badly broken at accessible energy scales (if it even exists).

In a different direction, physicists have recently discovered exotic particles (or resonances) described as *tetraquarks*, being composed of two quarks and two antiquarks, and even pentaquarks, which are composed of four quarks and one antiquark. In all cases confirmed to date, one of the quarks has been a $c$ or a $b$. Ignoring this, one could try to explore the representation-theoretic context of such exotics by analysing the tensor products

$$\mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_2} \otimes \mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_2} \quad \text{and} \quad \mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_1} \otimes \mathcal{L}_{\omega_2}. \tag{5.82}$$

However, without any experimental guidance, it isn't clear what this would achieve...

**Exercise 117.** Decompose the tensor product of the $\mathfrak{sp}(4)$-module $\mathcal{L}_{\omega_1}$ with itself into irreducibles, carefully explaining your reasoning at each step. ▼

# 6. Afterword...

So there are many things that one would like to have time to include in this first glimpse into the wonderful world of Lie algebras and their representations. The most serious omission here is that while you know how to determine the weights of a finite-dimensional module over a semisimple Lie algebra, in principle, we have not discussed how to determine their multiplicities, *ie.* the dimensions of the weight spaces. Another important discussion is the fact that complex Lie algebras are just the most accessible part of the wonderful world of Lie theory and we shall take some time to describe some of the generalisations and directions that mathematicians have followed, generally in pursuit of physicists, in their quest to understand symmetry.

## 6.1. **Multiplicities, characters and dimensions**

Let us denote by $\text{mult}_\lambda(\mu)$ the multiplicity of the weight $\mu$ in the irreducible highest-weight module $\mathcal{L}_\lambda$. Then, $\text{mult}_\lambda(\lambda) = 1$ (Exercise 97) and these multiplicities may be computed recursively using the following identity due to Freudenthal:

$$\left(\|\lambda+\rho\|^2 - \|\mu+\rho\|^2\right) \text{mult}_\lambda(\mu) = 2 \sum_{\alpha\in\Delta_+} \sum_{j=1}^{\infty} (\mu + j\alpha, \alpha)\, \text{mult}_\lambda(\mu + j\alpha). \tag{6.1}$$

Note that the sum over $j$ on the right-hand side is effectively finite because $\mu + j\alpha$ cannot be a weight of $\mathcal{L}_\lambda$, *ie.* $\text{mult}_{\mu+j\alpha}(\lambda) = 0$, for all sufficiently large $j$. The norms on the left-hand side and (especially) the shifts by $\rho$ suggest that Freudenthal's identity has something to do with the quadratic Casimir $Q$. Indeed, it is proven by explicitly computing the trace of $Q$ on the weight space of $\mathcal{L}_\lambda$ of weight $\mu$ and noting that $Q$ acts on this weight space as multiplication by $(\lambda, \lambda + 2\rho)$.

Freudenthal's recursion relation is rather tedious to apply by hand, but is not too hard to implement on a computer. In principle, one could use it to compute all the multiplicities of a finite-dimensional irreducible module $\mathcal{L}_\lambda$, *ie.* one with $\lambda \in \mathsf{P}_\geqslant$. Summing these multiplicities would then give us the dimension of $\mathcal{L}_\lambda$, a quantity that has been hitherto out of reach. However, this is not particularly satisfying... wouldn't it be nicer if there was a closed-form expression for this dimension?

There is. But rather than just state it immediately, let's use this as an opportunity to introduce another way of encoding multiplicities which turns out to be theoretically (and practically!) extremely important. This encoding is called the *character* of the module and it applies to quite general weight modules (as long as the multiplicities are finite). It is in fact nothing more than a generating function for the multiplicities, as we shall see.

Define *formal exponentials* $\mathsf{e}^\lambda$, for $\lambda \in \mathfrak{g}_0^*$, so that the usual property $\mathsf{e}^\lambda \mathsf{e}^\mu = \mathsf{e}^{\lambda+\mu}$ is satisfied. The formal exponential $\mathsf{e}^0$ is clearly the multiplicative unit of this abelian group, itself isomorphic to $\mathfrak{g}_0^*$, and we shall denote it by $\mathbb{1}$. The character of a weight module $V$

is then the following generating function $\mathrm{ch}[V]$ in the group algebra $\mathbb{C}\mathfrak{g}_0^*$:

$$\mathrm{ch}[V] = \sum_{\mu \in \mathfrak{g}_0^*} \mathrm{mult}_V(\mu)\, \mathsf{e}^{\mu}. \tag{6.2}$$

Here, $\mathrm{mult}_V(\mu)$ is the multiplicity of the weight $\mu$ in $V$. Note that even though the sum is over a set $\mathfrak{g}_0^*$ whose cardinality is uncountably infinite, the multiplicities are almost always zero for the modules that we have been concerned with, so the sum is effectively finite or countable.

It is common, especially in the physics literature, to write $\mu = \sum_{i=1}^{r} \mu_i \omega_i$, with $r = \operatorname{rank} \mathfrak{g}$, so that we may express the character in the alternative form

$$\mathrm{ch}[V] = \sum_{\lambda \in \mathfrak{g}_0^*} \mathrm{mult}_V(\mu)\, \mathsf{z}_1^{\mu_1} \cdots \mathsf{z}_r^{\mu_r}, \quad \text{where } \mathsf{z}_i = \mathsf{e}^{\omega_i} \in \mathfrak{g}_0^*. \tag{6.3}$$

One often writes the left-hand side as $\mathrm{ch}[V](\mathsf{z})$ to emphasise this change of viewpoint. It certainly makes the "generating function" nature of the character more manifest. Now observe that $\mu_i$ is the eigenvalue of the simple coroot $h_{\alpha_i}$ on the weight space of $V$ of weight $\mu$, whilst $\mathrm{mult}_V(\mu)$ is its algebraic multiplicity (in the weight space of weight $\mu$ anyway). We therefore have

$$\mathrm{ch}[V](\mathsf{z}) = \mathrm{tr}_V\, \mathsf{z}_1^{h_{\alpha_1}} \cdots \mathsf{z}_r^{h_{\alpha_r}}. \tag{6.4}$$

Physicists will recognise this form of the character as a mathematical abstraction of a **partition function** ($Z = \mathrm{tr}\, \mathsf{e}^{-\beta H}$ so $\mathsf{z} \sim \mathsf{e}^{-\beta}$) and thus realise why characters are so important in quantum physics and statistical mechanics.

From the point of view of mathematics, characters not only encode the multiplicities of a representation, they also do so in a way that behaves beautifully with respect to direct sums, tensor products and duals. Indeed, the analysis of Exercise 29 may be slightly generalised to show that

$$\mathrm{ch}[V \oplus W] = \mathrm{ch}[V] + \mathrm{ch}[W] \quad \text{and} \quad \mathrm{ch}[V \otimes W] = \mathrm{ch}[V]\,\mathrm{ch}[W], \tag{6.5}$$

whilst the character of $V^*$ is obtained from that of $V$ by inverting all the formal exponentials ($\mathsf{e}^{\mu} \longrightarrow \mathsf{e}^{-\mu}$). Equivalently, $\mathrm{ch}[V^*]$ is $\mathrm{ch}[V]$ with each $z_i$ replaced by $z_i^{-1}$:

$$\mathrm{ch}[V^*]\mathsf{z} = \mathrm{ch}[V]\mathsf{z}^{-1}. \tag{6.6}$$

Recall that weights are linear functionals on the Cartan subalgebra. For our purposes, it is important to realise that formal exponentials, and hence characters, may be regarded as (non-linear) functionals on the Cartan subalgebra. Equivalently, we may view formal exponentials and characters as functionals on the dual space, *ie.* acting on weights:

$$\mathsf{e}^{\lambda} \colon \mu \longmapsto \mathsf{e}^{(\lambda,\mu)} \in \mathbb{C}, \quad \text{for all } \mu \in \mathfrak{g}_0^*. \tag{6.7}$$

We call this action *evaluation* for hopefully obvious reasons. Note that evaluating the character of a module at 0, *ie.* replacing each $\mathsf{z}_i$ with 1, results in the dimension of the module (assuming that this is finite):

$$\operatorname{ch}[V]\big|_0 = \operatorname{ch}[V](1) = \sum_{\lambda\in\mathfrak{g}_0^*} \operatorname{mult}_V(\mu)\,\mathsf{e}^{(\mu,0)} = \sum_{\lambda\in\mathfrak{g}_0^*} \operatorname{mult}_V(\mu) = \dim V. \tag{6.8}$$

One can check that the character of the irreducible $\mathfrak{sl}(2)$-module $\mathcal{L}_\lambda$ is given by

$$\begin{aligned}\operatorname{ch}[\mathcal{L}_\lambda] &= \mathsf{e}^{\lambda\omega} + \mathsf{e}^{(\lambda-2)\omega} + \cdots + \mathsf{e}^{-(\lambda-2)\omega} + \mathsf{e}^{-\lambda\omega} \\ &= \mathsf{z}^{\lambda} + \mathsf{z}^{\lambda-2} + \cdots + \mathsf{z}^{-(\lambda-2)} + \mathsf{z}^{-\lambda} = \frac{\mathsf{z}^{\lambda} - \mathsf{z}^{-\lambda-2}}{\mathbb{1} - \mathsf{z}^{-2}} = \frac{\mathsf{z}^{\lambda+1} - \mathsf{z}^{-\lambda-1}}{\mathsf{z} - \mathsf{z}^{-1}}.\end{aligned} \tag{6.9}$$

Here, we have written the character as a Laurent polynomial in $\mathsf{z} = \mathsf{e}^{\omega}$ and then as a rational function for brevity. Evaluating at 0, *ie.*setting $\mathsf{z}$ to 1, indeed gives the dimension as $\lambda + 1$, either by summing $\lambda + 1$ copies of 1 or by applying l'Hôpital's rule. One can similarly determine the character of the Verma $\mathfrak{sl}(2)$-module $\mathcal{V}_\lambda$:

$$\operatorname{ch}[\mathcal{V}_\lambda] = \mathsf{z}^{\lambda} + \mathsf{z}^{\lambda-2} + \mathsf{z}^{\lambda-4} + \cdots = \frac{\mathsf{z}^{\lambda}}{\mathbb{1} - \mathsf{z}^{-2}} = \frac{\mathsf{z}^{\lambda+1}}{\mathsf{z} - \mathsf{z}^{-1}}. \tag{6.10}$$

This time, the character must be treated as a power series in $\mathsf{z}^{-2}$. Note that the limit as $\mathsf{z} \to 1$ diverges (even from below) in accord with the fact that $\dim \mathcal{V}_\lambda = \infty$.

This Verma module computation generalises easily to general (semisimple) $\mathfrak{g}$ with the result being that

$$\operatorname{ch}[\mathcal{V}_\lambda] = \frac{\mathsf{e}^{\lambda}}{\prod_{\alpha\in\Delta_+}(1 - \mathsf{e}^{-\alpha})} = \frac{\mathsf{e}^{\lambda+\rho}}{\prod_{\alpha\in\Delta_+}\left(\mathsf{e}^{\alpha/2} - \mathsf{e}^{-\alpha/2}\right)}, \tag{6.11}$$

where we have used (5.50). This must be expanded in **negative** powers of the formal exponentials $\mathsf{e}^{\alpha}$ with $\alpha \in \Delta_+$. The generalisation of the computation for irreducible highest-weight modules is decidedly non-trivial in general. However, for $\lambda \in \mathsf{P}_{\geqslant}$ (*ie.* the case where $\dim \mathcal{L}_\lambda < \infty$), Weyl proved the following beautiful character formula:

$$\operatorname{ch}[\mathcal{L}_\lambda] = \frac{\sum_{w\in\mathsf{W}} \det w\, \mathsf{e}^{w(\lambda+\rho)}}{\prod_{\alpha\in\Delta_+}\left(\mathsf{e}^{\alpha/2} - \mathsf{e}^{-\alpha/2}\right)}. \tag{6.12}$$

Here, $\mathsf{W}$ is the Weyl group and $\det w \in \{\pm 1\}$ is the determinant of $w \in \mathsf{W}$ as a linear transformation on the real/rational inner product space $\mathsf{R}$.

One could wax lyrical about Weyl's character formula for a long long time. Here, we only have time to ruminate on three of its immediate consequences. The first is the specialisation corresponding to taking $\lambda = 0$. As $\mathcal{L}_0$ is the one-dimensional trivial module, its character is just $\mathsf{e}^0 = \mathbb{1}$. We therefore have Weyl's *denominator identity*:

$$\prod_{\alpha\in\Delta_+}\left(\mathsf{e}^{\alpha/2} - \mathsf{e}^{-\alpha/2}\right) = \sum_{w\in\mathsf{W}} \det w\, \mathsf{e}^{w(\rho)}. \tag{6.13}$$

Substituting back, we arrive at an alternative form for Weyl's character formula:

$$\mathrm{ch}[\mathcal{L}_\lambda] = \frac{\sum_{w \in \mathsf{W}} \det w \, \mathsf{e}^{w(\lambda+\rho)}}{\sum_{w \in \mathsf{W}} \det w \, \mathsf{e}^{w(\rho)}}. \tag{6.14}$$

The second consequence is our desired dimension formula. However, its derivation is not merely evaluating the character formula at 0. Weyl's formula is indeterminate at 0 so l'Hôpital's rule is required. More precisely, one evaluates (6.12) at $t\rho$ and then takes the limit as $t \to 0$. The result is remarkably simple:

$$\dim \mathcal{L}_\lambda = \prod_{\alpha \in \Delta_+} \frac{(\lambda + \rho, \alpha)}{(\rho, \alpha)}. \tag{6.15}$$

With the Gram matrices of Section 5.6 or the positive roots as linear combinations of simple roots, as in Section 4.7, it is now straightforward to compute the dimension of $\mathcal{L}_\lambda$ in terms of the Dynkin labels $\lambda_i$, *eg*.

$$\begin{aligned} \text{for } \mathfrak{sl}(3): \quad \dim \mathcal{L}_\lambda &= \tfrac{1}{2}(\lambda_1+1)(\lambda_2+1)(\lambda_1+\lambda_2+2) \\ \text{and for } \mathfrak{sp}(4): \quad \dim \mathcal{L}_\lambda &= \tfrac{1}{6}(\lambda_1+1)(\lambda_2+1)(\lambda_1+\lambda_2+2)(\lambda_1+2\lambda_2+3). \end{aligned} \tag{6.16}$$

Our third consequence is called (I kid you not) the *strange formula* of Freudenthal and de Vries. It determines the norm of the Weyl vector as follows:

$$\frac{\|\rho\|^2}{2\mathsf{h}^\vee} = \frac{\dim \mathfrak{g}}{24}. \tag{6.17}$$

This might seem like a bit of a curiosity and in some ways it is. However, it turns out to be crucial to an important link, predicted by physics, between Lie theory and number theory!

To be slightly more specific, conformal field theory requires that the characters of appropriate representations of certain infinite-dimensional generalisations of semisimple Lie algebras (called Kac–Moody algebras — see below) are *vector-valued modular forms*. This means that these characters span a representation of the modular group $\mathsf{SL}(2;\mathbb{Z})$ and can be expressed in terms of functions that number theorists have been studying for hundreds of years. This input from physics has, in so many ways, completely revitalised this corner of number theory.

To see where the strange formula comes from, recall that the dimension formula (6.15) may be derived by evaluating (carefully) the character formula at 0. This is equivalent to taking a Taylor series about 0 and computing the constant term. If, however, one computes the linear term, then (with a bit of hard work) one arrives at the strange formula. Note that the norm in this formula is scaled so that $\|\theta\|^2 = 2$ as one might expect from the explicit factor of $2\mathsf{h}^\vee$.

### 6.2. **But wait! There's more...**

I'd like to close with the admonishment that we really haven't begun to scratch the surface of Lie theory. There are many many directions remaining to explore. First, of course, there is still a lot that we've omitted about the theory of complex semisimple Lie algebras. Even granting this, one should ask what the classification in the complex setting means for real semisimple Lie algebras (the answer lies in a generalisation of Exercise 24), let alone what happens in other characteristics. And then there's the corresponding Lie groups which we only touched on briefly. With Lie groups, one adds geometry, topology and analysis to the algebraic palette that we have studied here. You should not be surprised to learn that the resulting picture is beautiful and important: for example, understanding Lie groups has driven many advances in algebraic and differential geometry and topology.

In a somewhat different direction, the last thirty years or so have seen many Lie-theoretic advances at the behest of mathematical physicists. More than a few have even been rewarded with Fields medals. One such advance is the mathematical characterisation of the symmetries that physicists uncovered in certain "integrable" systems. This characterisation was dubbed a *quantum group* by Drinfel'd (and was discovered independently by Jimbo). An important class of quantum groups is obtained by "deforming" the universal enveloping algebra of a complex semisimple Lie algebra. This involves introducing a (frequently complex) parameter $q$ into the defining relations (4.103) in a consistent fashion. Interestingly, this doesn't change the representation theory unless $q$ happens to be a root of unity, in which case things change a lot. More interestingly, physical applications of quantum groups seem to require $q$ to be a root of unity...

Another physical advance in Lie theory is the consideration of what are now called *Lie superalgebras*. These are just like Lie algebras except that elements come with a parity and the defining relations involving two odd elements acquire an additional sign. In particular, the Lie bracket of two odd elements is no longer modelled on the commutator, but rather on the anticommutator

$$\{x, y\} = xy + yx. \tag{6.18}$$

The original interpretation in physics is that even elements correspond to bosonic degrees of freedom, whilst odd elements are fermionic. A particular special case is *supersymmetry* in which one has a special type of bijection between the even and odd elements. For a long time, physicists were convinced that nature was secretly supersymmetric. However, nature keeps her secrets well and supersymmetry isn't looking all that great for physicists.

On the other hand, supersymmetry has been awesome for mathematicians. The mathematical theory of simple Lie superalgebras is now fairly well-developed, *eg.* there is a classification due to Kac that is analogous to that of Section 4.8. However, the theory has some nasty surprises. For example, simple Lie superalgebras have root systems, but the roots can have zero (or even negative) norm and they can even occur with multiplicity 2. On the other hand, the representation theory is much harder because Weyl's theorem fails for almost all simple Lie superalgebras: finite-dimensional modules need not be completely reducible.

Yet another physical advance relates to quantum field theory and, more specifically, string theories in which space(time) is a (suitable, *eg*. reductive) Lie group. Strings, being essentially circles, have an infinite number of vibrational modes (*cf.* Fourier series), each of which corresponds to a creation or annihilation operator in an infinite-dimensional Lie algebra. Completing this with zero-modes results in a so-called *loop algebra*. By a quirk of quantum mechanics (that's mathematically explained in terms of Lie algebra cohomology), quantised strings add a "central extension" to the loop algebra and one arrives at what mathematicians now call an *untwisted affine Kac–Moody algebra*.

These are infinite-dimensional Lie algebras with roots and coroots, Cartan matrices and Dynkin diagrams. They have been classified and behave fairly well — no roots have negative norm, though a few have zero norm and the Cartan matrices always have zero determinant. Their (non-trivial) irreducible highest-weight modules are all infinite-dimensional, but there are some which behave rather like the finite-dimensional modules of a semisimple Lie algebra. Luckily, these modules are precisely the ones that arise in the string theory's quantum state space (at least when the semisimple part of the Lie group is compact). It's probably fair to say that affine Kac–Moody algebras are one of the most important advances in modern Lie theory for both physicists and mathematicians.

And then there are affine Kac–Moody superalgebras, affine Kac–Moody (super)groups and the quantum (super)groups obtained by deforming their universal enveloping (super)algebras. Even more generally, much of the beauty of string theories can be attributed to what physicists call *conformal field theory* in which the Lorentz/Poincaré symmetries of special relativity are extended to conformal symmetries on, *eg*. two-dimensional Riemann surfaces. The symmetries underlying conformal field theories have been axiomatised in several ways, one of which is called a *vertex operator (super)algebra*.

Along with the mathematical gadgets mentioned above, vertex operator (super)algebras not only play a central role in modern mathematical physics, they also regularly astonish mathematicians from all areas (from algebraic geometers to number theorists to category theorists) with their amazing properties. I believe it's fair to say that this explosion of Lie theory is 21st century mathematics that is still being uncovered. There are surely many more beautiful seams to mine here and further connections to "classical" mathematics to understand. Hopefully, you'll get a chance to try your hand at this yourself one day.

— *fin* —