Lie Algebras.

Kevin McGerty[1]

Michaelmas 2022[2]

[1]Please contact me at mcgerty@maths.ox.ac.uk if you find any errors or ambiguities in these notes.
[2]Last updated: 4 November, 2022.

Contents

Sections, proofs, or individual Remarks which are marked with a (*) are non-examinable.

Notation

- $\mathbb{N}$: the natural numbers.
- $\mathbb{N}X$: if $X \subseteq V$ is a subset of a vector space V then $\mathbb{N}X = \{\sum_{i \in I} n_i x_i : n_i \in \mathbb{N}\}$, where I any finite indexing set such that $\{x_i : i \in I\} \subseteq X$.
- $[.,.]$: a Lie bracket, see Definition 2.1.2.
- $\mathfrak{gl}(V)$: for a vector space V, the Lie algebra which is $\mathrm{End}(V)$ as a vector space, equipped with the commutator as its Lie bracket.
- $\mathfrak{sl}(V)$: the Lie subalgebra of $\mathfrak{gl}(V)$ constisting of endomorphisms of trace 0.
- An *ideal* in a Lie algebra $\mathfrak{g}$ is a subspace $\mathfrak{a}$ such that $[a,x] \in \mathfrak{a}$ for all $x \in \mathfrak{g}, a \in \mathfrak{a}$. See Definition 2.2.1.
- $N_{\mathfrak{g}}(\mathfrak{a})$: the normalizer of a subalgebra $\mathfrak{a}$ in a Lie algebra $\mathfrak{g}$. See Definition 2.2.6.
- $(C^k(\mathfrak{g}))_{k \geq 0}$: The *lower central series* of a Lie algebra, see Definition 5.3.2
- $(D^k(\mathfrak{g}))_{k \geq 0}$: The *derived series* of a Lie algebra, see Definition 5.1.3
- $\kappa = \kappa^{\mathfrak{g}}$: The Killing form, an invariant symmetric bilinear form on a Lie algebra given by:

$$\kappa(x,y) = \mathrm{tr}_{\mathfrak{g}}(\mathrm{ad}(x)\mathrm{ad}(y)).$$

- $\mathrm{rad}(B)$: If B is a symmetric bilinear form on a vector space V then

$$\mathrm{rad}(B) = \{v \in V : B(v,w) = 0, \forall w \in V\}.$$

- $\mathrm{rad}(\mathfrak{g})$: If $\mathfrak{g}$ is a Lie algebra, then $\mathrm{rad}(\mathfrak{g})$ is the maximal solvable ideal in $\mathfrak{g}$.
- (V,Φ): an abstract root system. See Definiton 7.3.2.

0.1 Modifications

(i) Chapter 1: Added a few more details about the examples discussed in lecture 1, but this is just for curiosity.

(ii) Chapter 2: Clarified the definition of a k-algebra to tidy up the unital and non-unital cases: the current version gives a definition of a k-algebra structure on a k-vector space A with or without a unit. In the case where A has a unit, then the definition becomes equivalent to the existence of a homomorphism of rings $\mathsf{k} \to A$ whose image lies in the centre of A.

(iii) Chapter 5: corrected some typos in Section 5.1 and subsection 5.3.2 now matches the account of representations of nilpotent Lie algebras give in lectures. (The original online notes had a mistake here.)

Chapter 1

*Background

In this section I use some material, like multivariable analysis, which is not necessary for the main body of the course, but if you know it, or are happy to rely on notions from Prelims multivariable calculus for which you have not been given a rigorous definition, it will help to put the material of this course in a broader context. For those worried about such things, fear not, it is non-examinable.

1.1 From group actions to group representations

In mathematics, group actions give a way of encoding the symmetries of a space or physical system. Formally these are defined as follows: an action of a group G on a space[1] X is a map $a\colon G \times X \to X$, written $(g.x) \mapsto a(g,x)$ or more commonly $(g,x) \mapsto g.x$ which satisfies the properties

1. $e.x = x$, for all $x \in X$, where $e \in G$ is the identity;
2. $(g_1 g_2).x = g_1.(g_2.x)$ for all $g_1, g_2 \in G$ and $x \in X$.

Natural examples of actions are that of the group of rigid motions SO_3 on the unit sphere $\{x \in \mathbb{R}^3 : \|x\| = 1\}$, or the general linear group $\mathrm{GL}_n(\mathbb{R})$ on $\mathbb{R}^n$.

Whenever a group acts on a space X, there is a resulting linear action (a representation) on the vector space of functions on X. Indeed if $\mathrm{Fun}(X)$ denotes the vector space of real-valued functions on X, then the formula

$$g(f)(x) = f(g^{-1}.x), \quad \forall g \in G, f \in \mathrm{Fun}(X), x \in X,$$

defines a representation of G on $\mathrm{Fun}(X)$. If X and G have more structure. *e.g.* that of a topological space or smooth manifold, then this action may also preserve the subspaces of say continuous, or differentiable functions. Lie algebras arise as the "infinitesimal version" of group actions, which loosely speaking means they are what we get by trying to differentiate group actions.

Example 1.1.1. Take for example the natural action of the circle S^1 by rotations on the plane $\mathbb{R}^2$. This action can be written explicitly using matrices:

$$g(t) = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}$$

where we have smoothly parametrized the circle S^1 using the trigonometric functions. Note that for this parametrization, $g(t)^{-1} = g(-t)$. The induced action on $\mathrm{Fun}(\mathbb{R}^2)$ restricts to an action on

[1] I'm being deliberately vague here about what a "space" is, X could just be a set, but it could also have a more geometric nature, such as a topological space or submanifold of $\mathbb{R}^n$.

$C^\infty(\mathbb{R}^2)$ the space of smooth (*i.e.* infinitely differentiable) functions on $\mathbb{R}^2$. Using our parametrization, it makes sense to differentiate this action at the identity element (*i.e.* at $t = 0$) to get an operation $\nu\colon C^\infty(\mathbb{R}^2) \to C^\infty(\mathbb{R}^2)$, where if $z = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, then ν is given by

$$\begin{aligned}
\nu(f) &= \frac{d}{dt}(f(g(-t).z))_{|t=0} \\
&= -Df_z \circ g'(0).(z) \\
&= -\begin{pmatrix} \partial_x f & \partial_y f \end{pmatrix} \begin{pmatrix} -\sin(t) & -\cos(t) \\ \cos(t) & -\sin(t) \end{pmatrix}_{|t=0} \begin{pmatrix} x \\ y \end{pmatrix} \\
&= y\partial_x - x\partial_y.
\end{aligned}$$

The operator we obtained in this example, $\nu = y\partial_x - x\partial_y$ is a $C^\infty(\mathbb{R}^2)$- linear combination of ∂_x and ∂_y. Such an operation is called a *derivation*, and we now discuss their structure. We will work with the spaces $\mathbb{R}^n$ for the rest of this section, but everything we say also applies, *mutatis mutandis* to the context of smooth manifolds.

Definition 1.1.2. For any positive integer n, an $\mathbb{R}$-linear operator $\nu\colon C^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ is said to be a *derivation* if, for any $f_1, f_2 \in C^\infty(\mathbb{R}^n)$ it satisfies

$$\nu(f_1.f_2) = \nu(f_1).f_2 + f_1.\nu(f_2). \tag{1.1}$$

The next Lemma shows that the previous, somewhat formal, definition, actually results in a class of objects with a very concrete description. When working in $\mathbb{R}^n$ we will denote the partial derivative of f in the direction of the i-th standard basis vector by $\partial_i f$ (in preference to the notation $\partial f/\partial x_i$ you may have seen more often).

Lemma 1.1.3. *If $\nu\colon C^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ is a derivation, then there exist unique smooth functions $a_1, \dots, a_n \in C^\infty(\mathbb{R}^n)$ such that for all $f \in C^\infty(\mathbb{R}^n)$ we have*

$$\nu(f) = \sum_{j=1}^{n} a_j \partial_j(f).$$

Proof. Let ν be a derivation. Since $1^2 = 1 \in \mathbb{R}$, we must have $\nu(1) = \nu(1).1 + 1.\nu(1)$, that is, $2\nu(1) = \nu(1)$ so that $\nu(1) = 0$, and hence by linearity $\nu(\lambda) = 0$ for all $\lambda \in \mathbb{R}$, *i.e.* any derivation vanishes on the subspace of constant functions $\mathbb{R} \subseteq C^\infty(\mathbb{R}^n)$.

Now suppose that f is an arbitrary smooth function, and fix a point $c \in \mathbb{R}^n$. For $x \in \mathbb{R}^n$, let $\gamma_x\colon [0,1] \to \mathbb{R}^n$ be the line-segment path $\gamma_x(t) = c + t(x - c)$ from c to x, and let $h_x(t) = f(\gamma_x(t))$. Now by the fundamental theorem of calculus $f(x) - f(c) = h(1) - h((0) = \int_0^1 h'(t)\mathrm{d}t$ and since by chain rule we have $h'(t) = Df_{\gamma_x(t)}(x - c)$, we obtain

$$f(x) = f(c) + \int_0^1 \sum_{j=1}^{n} \partial_j f(\gamma_x(t))(x_j - c_j)\mathrm{d}t = f(c) + \sum_{j=1}^{n} g_j(x).(x_j - c_j)$$

where $g_j(x) = \int_0^1 \partial_j f(c + t(x - c))\mathrm{d}t \in C^\infty(\mathbb{R}^n)$ and clearly $g_j(c) = \partial_j f(c)$. Since the functions $(x_j - c_j)$ vanish at $x = c$ for all j, it follows from the (1.1) that $\nu(g_j.(x_j - c_j))(c) = g_j(c)\nu(x_j - c_j)$. Using this and the fact that $\nu(\mathbb{R}) = 0$, it follows that if we set $a_j = \nu(x_j) = \nu(x_j - c_j)$ then

$$\nu(f)(c) = \sum_{j=1}^{n} a_j(c)\partial_j f(c).$$

Since $c \in \mathbb{R}^n$ was arbitrary, it follows that $\nu(f) = \sum_{j=1}^n a_j \partial_j f$ as required. □

It follows that to give a derivation is the same as to give an n-tuple of functions $(a_1, \dots, a_n)$, or in other words a smooth function $a \colon \mathbb{R}^n \to \mathbb{R}^n$.

Definition 1.1.4. A *vector field* on $X = \mathbb{R}^n$ (or, with a bit more work, any manifold) is a (smooth) function $v \colon \mathbb{R}^n \to \mathbb{R}^n$, which one can think of as giving the infinitesimal direction of a flow (*e.g.* of a fluid, or an electric field say). The set of vector fields forms a vector space which we denote by Θ_X. Such fields can be made to act on functions $f \colon X \to \mathbb{R}$ by differentiation. If $v = (a_1, a_2, \dots, a_n)$ in standard coordinates (here $a_i : \mathbb{R}^n \to \mathbb{R}$), then set

$$v(f) = \sum_{i=1}^{n} a_i \frac{\partial f}{\partial x_i}.$$

By the previous Lemma, this yields a bijection between vector fields and derivations on $C^\infty(\mathbb{R}^n)$.

Heuristically, we think of the infinitesimal version of a group action as the collection of derivations on smooth functions we obtain by "differentiating the group action at the identity element". (For the circle the collection of vector fields we get are just the scalar multiples of the vector field v, but for actions of larger group this will yield a larger space of derivations). It turns out this set of derivations forms a vector space, but it also has a kind of "product" which is a sort of infinitesimal remnant of the group multiplication[2]. The next definition seeks to formalize the notion of an "infinitesimal symmetry".

Note that if we compose two derivations $v_1 \circ v_2$ we again get an operator on functions, but it is not given by a vector field, since it involves second order differential operators. However, it is easy to check using the symmetry of mixed partial derivatives that if v_1, v_2 are derivations, then $[v_1, v_2] = v_1 \circ v_2 - v_2 \circ v_1$ is again a derivation. Thus the space Θ_X of vector fields on X is equipped with a natural product[3] $[.,.]$ which is called a *Lie bracket*. The derivatives of a group action give subalgebras of the algebra Θ_X.

Example 1.1.5. Consider the action of $\mathrm{SO}_3(\mathbb{R})$ on $\mathbb{R}^3$. This is the group of orientation-preserving linear isometries of $\mathbb{R}^3$. It is well-known that any element of $g \in \mathrm{SO}_3(\mathbb{R})$ is a rotation about an axis L through the origin by some angle, θ let us say. Then there is a continuous path γ in $\mathrm{SO}_3(\mathbb{R})$ from the identity to g which, for $t \in [0,1]$ is the rotation by $t.\theta$ about that axis.

This path is smooth and extends to t in an open interval containing $t = 0$, so it makes sense to associate to it the derivation $f \mapsto \frac{d}{dt}(f(\gamma(-t)(x)))$. Picking an orthonormal basis $\{e_1, e_2, e_3\}$ which is positively oriented, with e_3 lying along the axis of rotation of g and e_1 and e_2 on the plane perpendicular to e_3, then a calculation almost identical to the one above in the case of the circle shows that v is a scalar multiple of $x_2\partial_1 - x_1\partial_2$, where the scalar depends on the angle θ.

But since, for each $g \in \mathrm{SO}_3(\mathbb{R})$, the derivation v_g we obtain in this way, is determined up to scaling by the axis of rotation, and if we conjugate g by an element of $h \in \mathrm{SO}_3(\mathbb{R})$, then hgh^{-1} is a rotation by the same angle around the axis $h(L)$ and $h\gamma(t)h^{-1}$ is a path from the identity to hgh^{-1}. Applying the chain rule as for the case of a circle, noting that a linear map is its own derivative, it follows that the derivation obtained from using the rotation hgh^{-1} in place of g is obtained from that for g simply by applying h. It follows from this that the linear span of all such derivations is in fact a 3-dimensional vector space $\mathfrak{g} = \mathrm{span}_{\mathbb{R}}\{x\partial_y - y\partial_x, y\partial_z - z\partial_y, z\partial_x - x\partial_z\}$, and moreover it is then not hard to check that $\mathfrak{g}$ is closed under the bracket operations $[\cdot,\cdot]$. (This also gives a non-trivial example of a 3-dimensional Lie algebra).

[2] To be a bit more precise, it comes from the conjugation action of the group on itself.

[3] This is in the weakest sense, in that it is a bilinear map $\Theta_X \times \Theta_X \to \Theta_X$. It is not even associative – the axiom it does satisfy is discussed shortly.

Chapter 2

Lie algebras: Definition and Basic properties

2.1 Definitions and Examples

The definition of a Lie algebra is an abstraction of the above example of the product on vector fields. It is purely algebraic, so it makes sense over any field k. We begin, however, with an even more basic definition:

Definition 2.1.1. Let R be a commutative ring[1]. An *R-algebra* is a pair $(A, \star)$ consisting of an R-module A and an R-bilinear map $\star\colon A \times A \to A$, that is, for all $a_1, a_2, b_1, b_2 \in A$ and $r \in R$, the operation $\star$ satisfies:

$$(r.a_1 + a_2) \star b_1 = r.(a_1 \star b_1) + (a_2 \star b_1),$$
$$a_1 \star (r.b_1 + b_2) = r.a_1 \star b_1 + (a_1 \star b_2).$$

We say that $(A, \star)$ is *unital* (or *has a unit*) if there is an element $1_A \in A$ such that $1_A \star a = a \star 1_A = a$ for all $a \in A$. Note that if it exits, the multiplicative unit is unique. We say that $(A, \star)$ is *associative* if $a \star (b \star c) = (a \star b) \star c$ for all $a, b, c \in A$. When A is associative, we will normally suppress the operation $\star$ and so, for any $a, b \in A$, write ab rather than $a \star b$ for the value of the bilinear map on the pair (a, b).

Note that an associative $\mathbb{Z}$-algebra (*i.e.* letting $R = \mathbb{Z}$ the integers) is just a ring. In this course we will usually assume that R is a field, which we will denote by k.

Definition 2.1.2. A Lie algebra over a field k is a pair $(\mathfrak{g}, [.,.]_{\mathfrak{g}})$ consisting of a k-vector space $\mathfrak{g}$, along with a bilinear operation $[.,.]_{\mathfrak{g}}\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ taking values in $\mathfrak{g}$ known as a Lie bracket, which satisfies the following axioms:

1. $[.,.]_{\mathfrak{g}}$ is alternating, *i.e.* $[x, x]_{\mathfrak{g}} = 0$ for all $x \in \mathfrak{g}$.
2. The Lie bracket satisfies the *Jacobi Identity*: that is, for all $x, y, z \in \mathfrak{g}$ we have:

$$[x, [y, z]_{\mathfrak{g}}]_{\mathfrak{g}} + [y, [z, x]_{\mathfrak{g}}]_{\mathfrak{g}} + [z, [x, y]_{\mathfrak{g}}]_{\mathfrak{g}} = 0.$$

Remark 2.1.3. It is straight-forward to check directly from the definition that the commutator bracket $[v_1, v_2] = v_1 \circ v_2 - v_2 \circ v_1$ which we put on the space of vector fields Θ_X satisfies the above conditions.

Note that by considering the bracket $[x+y, x+y]_{\mathfrak{g}}$ it is easy to see that the alternating condition implies that for all $x, y \in L$ we have $[x, y]_{\mathfrak{g}} = -[y, x]_{\mathfrak{g}}$, that is $[.,.]_{\mathfrak{g}}$ is skew-symmetric. If $\mathrm{char}(\mathsf{k}) \neq 2$,

[1]All *commutative* rings in this course will have a multiplicative identity.

the alternating condition is equivalent to skew-symmetry. Note that a Lie algebra is a k-algebra in the sense of Definition 2.1.1 where the product $[.,.]$ is neither commutative nor associative, and moreover it does not have a unit (*i.e.* a multiplicative identity)[2]. We will normally simply write $[.,.]$ and reserve use the decorated bracket only for emphasis or where there is the potential for confusion.

Definition 2.1.4. Let $(\mathfrak{g}_1, [.,.]_1)$ and $(\mathfrak{g}_2, [.,.]_2)$ be Lie algebras. A k-linear map $\phi\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ is said to be a *homomorphism* of Lie algebras if it respects the Lie brackets, that is:

$$\phi([a,b]_1) = [\phi(a), \phi(b)]_2 \quad \forall a, b \in \mathfrak{g}_1.$$

An *isomorphism* of Lie algebras is a bijective homomorphism, since, just as for group homomorphisms and linear maps, the (set-theoretic) inverse of a Lie algebra homomorphism is automatically itself a Lie algebra homomorphism.

Example 2.1.5.

1. If V is any vector space then setting the Lie bracket $[.,.]$ to be zero we get a (not very interesting) Lie algebra. Such Lie algebras are called *abelian* Lie algebras.

2. If A is an (associative) k-algebra, then A can be given the structure of a k-Lie algebra, where if $a, b \in A$ then we set $[a,b] = a.b - b.a$, the *commutator* of a and b. The commutator bracket is clearly alternating, and checking the Jacobi identity is a fundamental calculation. Indeed we have

$$\begin{aligned}[x,[y,z]] &= x(yz - zy) - (yz - zy)x = xyz - xzy - yzx + zyx \\ &= (xyz - yzx) + (zyx - yzx)\end{aligned}$$

   where,in the final expression, we have paired terms which can be obtained from each other by cycling x, y and z. Since the terms in these pairs have opposite signs, it is then clear that adding the three expressions obtained by cycling x, y and z gives zero.

3. For a more down-to-earth example, take $\mathfrak{g} = \mathfrak{gl}_n$ the k-vector space of $n \times n$ matrices with entries in k. It is easy to check that this is a Lie algebra for the commutator product:

$$[X, Y] = X.Y - Y.X.$$

   Slightly more abstractly, if V is a vector space, then we will write $\mathfrak{gl}(V)$ for the Lie algebra $\mathrm{End}(V)$ equipped with the commutator product as for matrices.

4. If $\mathfrak{g}$ is a Lie algebra and $N < \mathfrak{g}$ is a k-subspace of $\mathfrak{g}$ on which the restriction of the Lie bracket takes values in N, so that it induces a bilinear form $[.,.]_N\colon N \times N \to N$, then $(N, [.,.]_N)$ is clearly a Lie algebra, and we say N is a (Lie) *subalgebra* of $\mathfrak{g}$.

5. Let $\mathfrak{sl}_n = \{X \in \mathfrak{gl}_n : \mathrm{tr}(X) = 0\}$ be the space of $n \times n$ matrices with trace zero. It is easy to check that $\mathfrak{sl}_n$ is a Lie subalgebra of $\mathfrak{gl}_n$ (even though it is *not* a subalgebra of the associative algebra $\mathrm{End}(V)$). More generally we say any Lie subalgebra of $\mathfrak{gl}(V)$ for a vector space V is a *linear Lie algebra*.

6. Generalising the example of vector fields a bit, if A is a k-algebra and $\delta\colon A \to A$ is a k-linear map, then we say δ is a k-*derivation* if it satisfies the Leibniz rule, that is, if:

$$\delta(a.b) = \delta(a).b + a.\delta(b), \quad \forall a, b \in A.$$

   It is easy to see by a direct calculation that if $\mathrm{Der}_{\mathsf{k}}(A)$ denotes the k-vector space of k-derivations on A, then $\mathrm{Der}_{\mathsf{k}}(A)$ is stable under taking commutators, that is, if

$$[\delta_1, \delta_2] = \delta_1 \circ \delta_2 - \delta_2 \circ \delta_1.$$

[2]This makes them sound awful. However, as we will see this is not the way to think about them!

then $[\delta_1, \delta_2] \in \mathrm{Der}_{\mathsf{k}}(A)$. Indeed

$$\begin{aligned}(\delta_1 \circ \delta_2 - \delta_2 \circ \delta_1)(a.b) &= \delta_1\left(\delta_2(a).b + a.\delta_2(b)\right) - \delta_2\left(\delta_1(a).b + a\delta_2(b)\right) \\ &= \delta_1\delta_2(a).b + \delta_2(a).\delta_1(b) + \delta_1(a).\delta_2(b) + a.\delta_2(\delta_1(b)) \\ &\quad - \delta_2\delta_1(a).b - \delta_1(a).\delta_2(b) - \delta_2(a).\delta_1(b) - a.\delta_2\delta_1(b)) \\ &= [\delta_1, \delta_2](a).b + a.[\delta_1, \delta_2](b).\end{aligned}$$

7. If A is an associative k-algebra, then if $a \in A$ the operation of taking commutator with a is a derivation. That is, if $\delta_a \colon A \to A$ is given by $\delta_a(b) = [a, b]$ for any $b \in A$, then $\delta_a \in \mathrm{Der}_{\mathsf{k}}(A)$. Indeed
$$\delta_a(b).c + b.\delta_a(c) = (ab - ba)c + b(ac - ca) = a.(bc) - (bc).a = \delta_a(b.c)$$
The map $\Delta \colon A \to \mathrm{Der}_{\mathsf{k}}(A)$ given by $a \mapsto \delta_a$ is compatible with commutators:
$$\delta_{[a,b]} = [\delta_a, \delta_b].$$
In fact slightly more is true: if $\partial \in \mathrm{Der}_{\mathsf{k}}(A)$ and $b \in A$ then $[\partial, \delta_b] = \delta_{\partial(b)}$. (Applying this to $\partial = \delta_a$ gives the compatibility with commutators). Indeed for all $c \in \mathfrak{g}$ we have
$$[\partial, \delta_b](c) = \partial(bc - cb) - (b\partial(c) - \partial(c).b) = \partial(b).c - c.\partial(b) = \delta_{\partial(b)}(c).$$
Thus in particular, the map $\Delta \colon A \to \mathrm{Der}_{\mathsf{k}}(A)$ is a homomorphism of Lie algebras (where A and $\mathrm{Der}_{\mathsf{k}}(A)$ are equipped with the commutator bracket).

8. Given a Lie algebra $\mathfrak{g}$ we let $\mathrm{Der}_{\mathsf{k}}(\mathfrak{g}) = \{\phi \in \mathfrak{gl}(\mathfrak{g}) : \phi([x, y]) = [\phi(x), y] + [x, \phi(y)]\}$. It is a Lie subalgebra of $\mathfrak{gl}(\mathfrak{g})$ (indeed the proof above that $\mathrm{Der}_{\mathsf{k}}(A)$ is a Lie algebra only requires the product on A to be bilinear).

9. One way of interpreting the Jacobi identity is that, assuming the alternating property, it is equivalent to the condition that, for any $x \in \mathfrak{g}$, the operation $\mathrm{ad}(x) \in \mathfrak{gl}(\mathfrak{g})$ given by $\mathrm{ad}(x)(y) = [x, y]$ lies in $\mathrm{Der}_{\mathsf{k}}(\mathfrak{g})$. Indeed
$$\begin{aligned}\mathrm{ad}(x)([y, z]) &= [\mathrm{ad}(x)(y), z] + [y, \mathrm{ad}(x)(z)] \\ &\iff [x, [y, z]] = [[x, y], z] + [y, [x, z]] \\ &\iff [x, [y, z]] - [y, [x, z]] - [[x, y], z] = 0 \\ &\iff [x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0\end{aligned}$$
where the equivalence between the third and fourth equalities follows from the alternating property of a Lie bracket.

10. The Jacobi identity is also equivalent, again assuming the alternating property, to the fact that $\mathrm{ad} \colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is a homomorphism of Lie algebras: Indeed, for all $x, y, z \in \mathfrak{g}$ we have
$$\begin{aligned}[\mathrm{ad}(x), \mathrm{ad}(y)](z) &= [x, [y, z]] - [y, [x, z]] \\ &= [x, [y, z]] + [y, [z, x]] \\ &= -[z, [x, y]] \\ &= \mathrm{ad}([x, y])(z).\end{aligned}$$
where the second and fourth equality uses the alternating property, and the third the Jacobi identity.

**Remark* 2.1.6. Combining (8) and (9) in the above example we see that the adjoint representation $x \mapsto \mathrm{ad}(x)$ is in fact a Lie algebra homomorphism from $\mathfrak{g}$ to $\mathrm{Der}_{\mathsf{k}}(\mathfrak{g})$. This is, in a sense, where the Jacobi identity comes from: very roughly, the conjugation action of G on itself yields a group homomorphism $G \to \mathrm{GL}(\mathfrak{g})$ (since conjugation preserves the identity $e \in G$) whose image lies in $\mathrm{Aut}(\mathfrak{g})$. The adjoint representation of $\mathfrak{g}$ is then the derivative of this action yields the adjoint representation ad which hence should have image in $\mathrm{Der}_{\mathsf{k}}(\mathfrak{g})$.

2.2 Ideals and isomorphism theorems

As one might expect if a Lie algebra is suppose to be an "infinitesimal" version of a Lie group, most notions for groups have analogues in the context of Lie algebras. It might be worth noting, however, that the linear structure of a Lie algebra comes from the basic properties of the derivative: it is the Lie bracket which reflects the "infinitesimal" versions of properties of a group. The existence of both the linear structure and the Lie bracket means that many of the notions we consider for a Lie algebra also have natural analogues for a ring (which is an algebra object equipped with an addition and an (associative) multiplication.

Definition 2.2.1. An *ideal* in a Lie algebra $(\mathfrak{g}, [.,.]_{\mathfrak{g}})$ is a subspace $\mathfrak{a}$ such that for all $x \in \mathfrak{g}$ and $a \in \mathfrak{a}$ we have $[a,x]_{\mathfrak{g}} \in \mathfrak{a}$. It is easy to check that if $\phi\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ is a homomorphism, then

$$\ker(\phi) = \{a \in \mathfrak{g}_1 : \phi(a) = 0\}$$

is an ideal of $\mathfrak{g}_1$.

Remark 2.2.2. Notice that because a Lie bracket is alternating, the condition that, for all $x \in \mathfrak{g}$ and $a \in \mathfrak{a}$ one has $[a,x] \in \mathfrak{a}$, is equivalent to the condition that $[x,a] \in \mathfrak{a}$ for all $x \in \mathfrak{g}$, $a \in \mathfrak{a}$. Thus, similarly to commutative rings, the notions of a left, right or two-sided ideal in a Lie algebra are all the same.

Just as for rings, in fact any ideal is the kernel of a Lie algebra homomorphism:

Theorem 2.2.3. *The first isomorphism theorem for Lie algebras:*

1. *Let $\mathfrak{a}$ be an ideal in a Lie algebra $\mathfrak{g}$. Let $\mathfrak{g}/\mathfrak{a}$ be the quotient space and let $q\colon \mathfrak{g} \to \mathfrak{g}/\mathfrak{a}$ be the quotient map (of vector spaces). Then there is a unique Lie bracket on $\mathfrak{g}/\mathfrak{a}$ with respect to which q is a homomorphism of Lie algebras, that is, for all $x, y \in \mathfrak{g}$*

$$[q(x), q(y)] = q([x,y]), \quad \textit{i.e.} \quad [x+\mathfrak{a}, y+\mathfrak{a}] = [x,y] + \mathfrak{a}.$$

   Moreover, if $\phi\colon \mathfrak{g} \to \mathfrak{k}$ is a Lie algebra homomorphism such that $\phi(\mathfrak{a}) = 0$, then ϕ induces a homomorphism $\bar{\phi}\colon \mathfrak{g}/\mathfrak{a} \to \mathfrak{k}$ such that $\bar{\phi} \circ q = \phi$, so that $ker(\bar{\phi}) = ker(\phi)/\mathfrak{a}$.

2. *Let $\phi\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ be a homomorphism of Lie algebras. The subspace $\phi(\mathfrak{g}_1) = im(\phi)$ is a subalgebra of $\mathfrak{g}_2$ and ϕ induces an isomorphism $\bar{\phi}\colon \mathfrak{g}/ker(\phi) \to im(\phi)$.*

Proof. The proof is almost identical to the proof in the case of rings. The key point is to see that the coset $[x,y] + \mathfrak{a}$ is independent of the choice of representative for the cosets $x + \mathfrak{a}$, $y + \mathfrak{a}$, and the condition that $\mathfrak{a}$ is an ideal ensures this. □

Definition 2.2.4. If V, W are subspaces of a Lie algebra $\mathfrak{g}$, then write $[V,W]$ for the linear span of the elements $\{[v,w] : v \in V, w \in W\}$. Notice that if I, J are ideals in $\mathfrak{g}$ then so is $[I,J]$. Indeed to check this, note that by part 8) of Example 2.1.5, if $z \in \mathfrak{g}$, $x \in I, y \in J$ then we have

$$[z,[x,y]] = \mathrm{ad}(z)([x,y]) = [\mathrm{ad}(z)(x), y] + [x, \mathrm{ad}(z)(y)] \in [I,J]$$

since $\mathrm{ad}(z)(x) = [z,x] \in I$ if $x \in \mathfrak{g}$, and similarly $\mathrm{ad}(z)(y) = [z,y] \in J$.

Remark 2.2.5. If I and J are ideals in a Lie algebra $\mathfrak{g}$ then it is easy to check that their intersection $I \cap J$ is again an ideal in $\mathfrak{g}$, and we have $[I,J] \subseteq I \cap J$. (Thus $[I,J]$ is the Lie algebra analogue of the product of ideals in a commutative ring.) Similarly, it is easy to see that the linear sum $I + J$ of I and J is also an ideal[3].

[3] Note however that the linear sum of two subalgebras is *not* necessarily a subalgebra.

Definition 2.2.6. Let $\mathfrak{g}$ be a Lie algebra and let $\mathfrak{a} \leq \mathfrak{g}$ be a subalgebra. If we let

$$N_{\mathfrak{g}}(\mathfrak{a}) = \{x \in \mathfrak{g} : \mathrm{ad}(x)(\mathfrak{a}) \subseteq \mathfrak{a}\} = \{x \in \mathfrak{g} : \mathrm{ad}(a)(x) \in \mathfrak{a}, \forall a \in \mathfrak{a}\},$$

then $N_{\mathfrak{g}}(\mathfrak{a}) \supseteq \mathfrak{a}$ is a subalgebra of $\mathfrak{g}$, as one can check using the formulation of the Jacobi identity given in Definition 2.2.4. It is the largest subalgebra of $\mathfrak{g}$ within which $\mathfrak{a}$ is an ideal.

Remark 2.2.7. Note that because the Lie bracket is skew-symmetric, we do not need to consider notions of left, right and two-sided ideals, as they will all coincide. If a nontrivial Lie algebra has no nontrivial ideals we say that it is *almost simple*. It it is in addition not Abelian, *i.e.* the Lie bracket is not identically zero, then we say that it is *simple*.

Just as for groups and rings, one can deduce the usual stable of isomorphism theorems from the first isomorphism theorem.

Theorem 2.2.8.

1. *If* $\mathfrak{h}$ *is a subalgebra of* $\mathfrak{g}$ *and* I *is an ideal in* $\mathfrak{g}$ *then* $\mathfrak{h} + I$ *is a subalgebra of* $\mathfrak{g}$ *(containing* I *as an ideal)* $\mathfrak{h} \cap I$ *is an ideal in* $\mathfrak{h}$*, and*

$$(\mathfrak{h} + I)/I \cong \mathfrak{h}/(\mathfrak{h} \cap I).$$

2. *If* $J \subset I \subset \mathfrak{g}$ *are ideals of* $\mathfrak{g}$ *then we have:*

$$(\mathfrak{g}/J)/(I/J) \cong \mathfrak{g}/I.$$

Proof. The proofs are identical to the corresponding results for groups. We give a proof of (2) as an example. Since $J \subseteq I$ the quotient map $\mathfrak{g}\colon \mathfrak{g} \to \mathfrak{g}/I$, which has kernel I, induces a map $\bar{q}\colon \mathfrak{g}/J \to \mathfrak{g}/I$. The kernel of this map is by definition $\{x + J : x + I = I\}$, that is, I/J. The result follows. ◻

Chapter 3

Representations of Lie algebras

3.1 Definition and examples

Just as for finite groups (or indeed groups in general) one way of studying Lie algebras is to try and understand how they can act on other (usually more concrete) objects. For Lie algebras, since they are already vector spaces over k, it is natural to study their action on linear spaces, or in other words, "representations". Formally we make the following definition.

Definition 3.1.1. A *representation* of a Lie algebra $\mathfrak{g}$ is a vector space V equipped with a homomorphism of Lie algebras $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$. In other words, ρ is a linear map such that

$$\rho([x,y]) = \rho(x) \circ \rho(y) - \rho(y) \circ \rho(x)$$

where $\circ$ denotes composition of linear maps. We may also refer to a representation of $\mathfrak{g}$ as a $\mathfrak{g}$-module. A representation is *faithful* if $\ker(\rho) = 0$. When there is no danger of confusion we will normally suppress ρ in our notation, and write $x(v)$ rather than $\rho(x)(v)$, for $x \in \mathfrak{g}, v \in V$.

If (V, ρ) and (W, σ) are $\mathfrak{g}$-representations, we say that $\phi\colon V \to W$ is a $\mathfrak{g}$-homomorphism (or homomorphism of $\mathfrak{g}$-representations) if $\phi \circ \rho(x) = \sigma(x) \circ \phi$ for all $x \in \mathfrak{g}$. We will sometimes write $\mathrm{Rep}(\mathfrak{g})$ for the collection[1] of representations of $\mathfrak{g}$.

We will study representation of various classes of Lie algebras in this course, but for the moment we will just give some basic examples.

Example 3.1.2. 1. If $\mathfrak{g} = \mathfrak{gl}(V)$ for V a vector space, then the identity map $\mathfrak{gl}(V) \to \mathfrak{gl}(V)$ is a representation of $\mathfrak{gl}(V)$ on V, which is known as the vector representation. Clearly any subalgebra $\mathfrak{g}$ of $\mathfrak{gl}(V)$ also inherits V as a representation, where then the map ρ is just the inclusion map.

2. Given an arbitrary Lie algebra $\mathfrak{g}$, there is a natural representation ad of $\mathfrak{g}$ on $\mathfrak{g}$ itself known as the adjoint representation. The homomorphism from $\mathfrak{g}$ to $\mathfrak{gl}(\mathfrak{g})$ is given by

$$\mathrm{ad}(x)(y) = [x,y], \quad \forall x, y \in \mathfrak{g}.$$

Indeed, as noted in Example 2.1.5, the fact that this map is a homomorphism of Lie algebras is just a rephrasing[2] of the Jacobi identity. Note that while the vector representation is clearly faithful, in general the adjoint representation is not. Indeed the kernel is known as the *centre* of $\mathfrak{g}$:

$$\mathfrak{z}(\mathfrak{g}) = \{x \in \mathfrak{g} : [x,y] = 0, \forall y \in \mathfrak{g}\}.$$

Note that if $x \in \mathfrak{z}(\mathfrak{g})$ then for any representation $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ the endomorphism $\rho(x)$ commutes with all the elements $\rho(y) \in \mathrm{End}(V)$ for all $y \in \mathfrak{g}$.

[1] If you take the Category Theory course, $\mathrm{Rep}(\mathfrak{g})$ is a category whose objects are representations of $\mathfrak{g}$ and whose morphisms are $\mathfrak{g}$-homomorphisms.

[2] It's also (for some people) a useful way of remembering what the Jacobi identity says.

3. If $\mathfrak{g}$ is any Lie algebra, then the zero map $\mathfrak{g} \to \mathfrak{gl}_1$ is a Lie algebra homomorphism. The corresponding representation is called the *trivial representation*. It is the Lie algebra analogue of the trivial representation for a group (which send every group element to the identity).

4. If (V, ρ) is a representation of $\mathfrak{g}$, we say that a subspace $U < V$ is a *subrepresentation* if $\phi(x)(U) \subseteq U$ for all $x \in \mathfrak{g}$. It follows immediately that ϕ restricts to give a homomorphism from $\mathfrak{g}$ to $\mathfrak{gl}(U)$, hence $(U, \phi_{|U})$ is again a representation of $\mathfrak{g}$. Note also that if $\{V_i : i \in I\}$ are a collection of invariant subspaces, their sum $\sum_{i \in I} V_i$ is clearly also invariant, and so again a subrepresentation.

5. If (V, ρ) is a $\mathfrak{g}$-representation, then it contains a natural subrepresentation

$$V^{\mathfrak{g}} = \{v \in V : \rho(x)(v) = 0, \forall x \in \mathfrak{g}\}$$

known as the $\mathfrak{g}$-*invariants* in V.

We now give a different kind of example, and consider the question of classifying all possible representations of the simplest Lie algebra $\mathfrak{gl}_1(\mathsf{k})$.

Example 3.1.3. Giving a representation of $\mathfrak{gl}_1$ is equivalent to giving a vector space equipped with a linear map. Indeed as a vector space $\mathfrak{gl}_1 = \mathsf{k}$, hence if (V, ρ) is a representation of $\mathfrak{gl}_1$ we obtain a linear endomorphism of V by taking $\rho(1)$. Since every other element of $\mathfrak{gl}_1$ is a scalar multiple of 1 this completely determines the representation, and this correspondence is clearly reversible.

If we assume k is algebraically closed, then you know the classification of linear endomorphisms is given by the Jordan canonical form. From this you can see that the only irreducible representations of $\mathfrak{gl}_1$ are the one-dimensional ones, while indecomposable representations correspond to linear maps with a single Jordan block.

Example 3.1.4. If L is a one-dimensional vector space and $\mathfrak{g}$ a Lie algebra, a representation of $\mathfrak{g}$ on L is a homomorphism of Lie algebras $\lambda \colon \mathfrak{g} \to \mathfrak{gl}(L) \cong k$, where the isomorphism $\mathfrak{gl}(L) \cong k$ is canonical – it is given by the scalar multiplication map. We write $\mathfrak{gl}_1 = \mathfrak{gl}_1(\mathsf{k})$ for k viewed as an Abelian Lie algebra, so that a one-dimensional representation of $\mathfrak{g}$ on L is given by a Lie algebra homomorphism $\lambda \colon \mathfrak{g} \to \mathfrak{gl}_1$ (identifying $\mathfrak{gl}_1$ with $\mathfrak{gl}(L)$ via the canonical isomorphism above).

Since $\mathfrak{gl}_1$ is abelian, the condition that λ is a homomorphism of Lie algebras is precisely the requirement that $\lambda(D(\mathfrak{g})) = 0$, thus the isomorphism classes of one-dimensional representations of $\mathfrak{g}$ are canonically identified with $(\mathfrak{g}/D(\mathfrak{g}))^* \cong D(\mathfrak{g})^0 \subseteq \mathfrak{g}^*$. Here, $D(\mathfrak{g})^0$ denotes the elements of $\mathfrak{g}^*$ which vanish on $D(\mathfrak{g})$. It is canonically identified with $(\mathfrak{g}/D(\mathfrak{g}))^*$ by the transpose of the quotient map $q \colon \mathfrak{g} \to \mathfrak{g}/D(\mathfrak{g})$.). We will write k_λ for the $\mathfrak{g}$-representation (k, λ), where we view λ as an element of $\mathfrak{g}^*$ which vanishes on $D(\mathfrak{g})$). In particular, if $\lambda = 0 \in \mathfrak{g}^*$ then k_0 is the trivial representation of $\mathfrak{g}$.

We end this section with some terminology which will be useful later.

Definition 3.1.5. A representation is said to be *irreducible* if it has no proper non-zero subrepresentations, and it is said to be *completely reducible* if it is isomorphic to a direct sum of irreducible representations. A representation V is said to be *indecomposable* if, whenever we have $V = U_1 \oplus U_2$ with U_1, U_2 subrepresentations, either $U_1 = V$ or $U_2 = V$ (and $U_2 = 0$, $U_1 = 0$ respectively).

3.2 Representations and constructions from linear algebra

There are a number of standard ways of constructing new representations from old, all of which have their analogues in the context of group representations. For example,

1. *Quotients*: Recall that if V is a k-vector space, and U is a subspace, then we may form the quotient vector space V/U. If $\phi \colon V \to V$ is an endomorphism of V which preserves U, that is if $\phi(U) \subseteq U$, then there is an induced map $\bar{\phi} \colon V/U \to V/U$. Applying this to representations of a Lie algebra $\mathfrak{g}$ we see that if V is a representation of $\mathfrak{g}$ and U is a subrepresentation we may always form the *quotient representation* V/U.

2. *Direct sums*: If (V,ρ) and (W,σ) are representations of $\mathfrak{g}$, then clearly $V \oplus W$, the direct sum of V and W, becomes a $\mathfrak{g}$-representation via the obvious homomorphism $\rho \oplus \sigma$.

3. *Hom spaces*: If (V,ρ) and (W,σ) are representations of $\mathfrak{g}$, then the vector space $\mathrm{Hom}(V,W)$ of linear maps from V to W also has the structure of a $\mathfrak{g}$-representation via

$$x(\phi) = \sigma(x) \circ \phi - \phi \circ \rho(x), \quad \forall x \in \mathfrak{g}, \phi \in \mathrm{Hom}(V,W). \tag{3.1}$$

One can check that this gives $\mathrm{Hom}(V,W)$ the structure of a $\mathfrak{g}$-representations, *i.e.* that it gives a Lie algebra homomorphism from $\theta\colon \mathfrak{g} \to \mathfrak{gl}(\mathrm{Hom}(V,W))$ by a direct calculation. Another, slightly quicker way, goes as follows: as already noted, $\rho \oplus \sigma$ defines a homomorphism from $\mathfrak{g}$ to $\mathfrak{gl}(V \oplus W)$. Now we may decompose $\mathfrak{gl}(V \oplus W) = \mathrm{Hom}_{\mathsf{k}}(V \oplus W, V \oplus W)$ into four summands which we write in the format of 2×2 matrix:

$$\mathfrak{gl}(V \oplus W) = \begin{pmatrix} \mathrm{Hom}_{\mathsf{k}}(V,V) = \mathfrak{gl}(V) & \mathrm{Hom}_{\mathsf{k}}(W,V) \\ \mathrm{Hom}_{\mathsf{k}}(V,W) & \mathrm{Hom}_{\mathsf{k}}(W,W) = \mathfrak{gl}(W) \end{pmatrix} \supseteq \begin{pmatrix} \rho(\mathfrak{g}) & 0 \\ 0 & \sigma(\mathfrak{g}) \end{pmatrix},$$

where the right-most containment shows the image of $(\rho \oplus \sigma)$ in terms of this decomposition: since it preserves the summands V and W, it is contained in the diagonal "blocks". Composing $(\rho \oplus \sigma)$ with the adjoint representation of $\mathfrak{gl}(V \oplus W)$ yields a homomorphism

$$\vartheta\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{gl}(V \oplus W)), \quad x \mapsto \mathrm{ad}\begin{pmatrix} \rho(x) & 0 \\ 0 & \sigma(x) \end{pmatrix}, \quad (\forall x \in \mathfrak{g}),$$

giving $\mathfrak{gl}(V \oplus W)$ the structure of a $\mathfrak{g}$-representation for which each direct of the four direct summands is a subrepresentation. The action on $\mathrm{Hom}_{\mathsf{k}}(V,W)$ is given by

$$\begin{pmatrix} \rho(x) & 0 \\ 0 & \sigma(x) \end{pmatrix}\begin{pmatrix} 0 & 0 \\ \phi & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ \phi & 0 \end{pmatrix}\begin{pmatrix} \rho(x) & 0 \\ 0 & \sigma(x) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \sigma(x) \circ \phi - \phi \circ \rho(x) & 0 \end{pmatrix}.$$

It follows that the formula in (3.1) gives an action of $\mathfrak{g}$ on $\mathrm{Hom}_{\mathsf{k}}(V,W)$, that is, the structure of a $\mathfrak{g}$-representation.

4. *Duals*: An important special case of this is where $W = \mathsf{k}$ is the trivial representation (as above, so that the map $\sigma\colon \mathfrak{g} \to \mathfrak{gl}(\mathsf{k})$ is the zero map). This allows us to give $V^* = \mathrm{Hom}(V, \mathsf{k})$, the dual space of V, a natural structure of $\mathfrak{g}$-representation where (since $\sigma = 0$) the action of $x \in \mathfrak{g}$ on $f \in V^*$ is given by $\rho^*\colon \mathfrak{g} \to \mathfrak{gl}(V^*)$ where

$$\rho^*(x)(f) = -f \circ \rho(x) \quad (f \in V^*).$$

If $\alpha\colon V \to V$ is any linear map, recall that the transpose map $\alpha^t\colon V^* \to V^*$ is defined by $\alpha^t(f) = f \circ \alpha$, thus our definition of the action of $x \in \mathfrak{g}$ on V^* is just[3] $-\rho(x)^t$. This makes it clear that the action of $\mathfrak{g}$ on V^* is compatible with the standard constructions on dual spaces, *e.g.* if U is a subrepresentation of V, the U^0 the annihilator of U will be a subrepresentation of V^*, and moreover, the natural isomorphism of V with V^{**} is an isomorphism of $\mathfrak{g}$-representations.

3.3 Representations and multilinear algebra: tensor products

An important method for constructing representations of a finite group G arises from the fact that if V and W are G-representations, then so is $V \otimes W$. It turns out that the same is true for representations of a Lie algebra.[4] In this section we show that the same is true for representations of Lie algebras. First, however, let us review the group case:

[3]Note that the minus sign is crucial to ensure this is a Lie algebra homomorphism – concretely this amounts to noticing that $A \mapsto -A^t$ preserves the commutator bracket on $n \times n$ matrices.

[4]As an intuitive guide, if Lie algebras are supposed to be some kind of "infinitesimal" version of a group, then their representations ought to have the same properties

3.3.1 Tensor products and group representations In fact we will consider two related questions one can consider:

(i) Suppose that G_1 and G_2 are groups, then we can form their product $G_1 \times G_2$ (as a set, this is the Cartesian product of G_1 and G_2, and the group operation is given componentwise). If we are given representations (V_1, ρ_1) and (V_2, ρ_2) of G_1 and G_2 respectively, can we build a representation of $G_1 \times G_2$ from them?

(ii) If we take $G_1 = G_2 = G$, can we build a new representation of G from V and W?

It is relatively easy to see that if the answer to the first question is positive then so is the answer to the second, because of (W, σ) is the representation produced by a solution to (i), then $\sigma \circ \Delta$ give a representation of G on W, where $\Delta \colon G \to G \times G$ given by $\Delta(g) = (g, g)$ (this map is usually called the "diagonal" map or embedding of G into $G \times G$).

One answer to the first question is given by the tensor product: the vector space $V \otimes W$ is naturally a representation of $G_1 \times G_2$. To see why, note that the homomorphisms ρ_1 and ρ_2 combine to give a homomorphism $(\rho_1, \rho_2) \colon G_1 \times G_2 \to \mathrm{GL}(V_1) \times \mathrm{GL}(V_2)$, and so to obtain an action of $G_1 \times G_2$ on $V_1 \otimes V_2$ it suffices to produce an action of $\mathrm{GL}(V_1) \times \mathrm{GL}(V_2)$ on $V \otimes W$.

To see what is required, one just needs to understand what data needs to be given in order to define a group homomorphism from $\mathrm{GL}(V) \times \mathrm{GL}(W) \to \mathrm{GL}(V \otimes W)$. The answer is given in the following simple Lemma.

Lemma 3.3.1. *Let $\theta \colon G_1 \times G_2 \to H$ be a group homomorphism and let $\theta_1 : G_1 \to H$ be given by $\theta_1(g_1) = \theta(g_1, e_{G_2})$, and similarly for $\theta_2 : G_2 \to H$. Then θ is determined by the pair of homomorphisms θ_1, θ_2 and any such pair yields a homomorphism of $G_1 \times G_2$ provided $\theta_1(g_1)\theta_2(g_2) = \theta_2(g_2)\theta_1(g_1)$ for all $g_1 \in G_1$ and $g_2 \in G_2$.*

Proof. For simplicity of notation, we identify G_1 and G_2 with their images in $G_1 \times G_2$ under the maps $g_1 \mapsto (g_1, e_{G_2})$ and $g_2 \mapsto (e_{G_1}, g_2)$ respectively. To see that the pair (θ_1, θ_2) determine θ uniquely, note that

$$\theta(g_1, g_2) = \theta(g_1, e_{G_2}).\theta(e_{G_1}, g_2) = \theta_1(g_1)\theta_2(g_2).$$

On the other hand, provided θ_1, θ_2 satisfy $\theta_1(g_1).\theta_2(g_2) = \theta_2(g_2)\theta_1(g_1)$, then it is easy to see that $\theta(g_1, g_2) = \theta_1(g_1)\theta_2(g_2)$ is a group homomorphism the the map $\theta \colon G_1 \times G_2 \to H$ given by $\theta(g_1, g_2) = \theta_1(g_1)\theta_2(g_2)$ is a group homomorphism. Indeed if $(g_1, g_2), (k_1, k_2) \in G_1 \times G_2$ then

$$\begin{aligned}
\theta((g_1, g_2).(k_1, k_2)) &= \theta((g_1k_1, g_2k_2)) := \theta_1(g_1k_1)\theta_2(g_2k_2) \\
&= \theta_1(g_1)\theta_1(k_1)\theta_2(g_2)\theta_2(k_2) = \theta_1(g_1)\theta_2(g_2)\theta_1(k_1)\theta_2(k_2) \\
&= \theta((g_1, g_2)).\theta(k_1, k_2).
\end{aligned}$$

□

Thus to give a homomorphism from $\mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \to \mathrm{GL}(V_1 \otimes V_2)$ we must give homomorphisms from $\mathrm{GL}(V_i) \to \mathrm{GL}(V_1 \otimes V_2)$ for $i = 1, 2$ whose images centralise each other. But this is easy to find using the basic compatibility between linear maps and the tensor product shown in I.2.2: if $\alpha_1, \alpha_2 \in \mathrm{Hom}(V_1, V_1)$ and $\beta_1, \beta_2 \in \mathrm{Hom}(V_2.V_2)$, then they induce linear maps $\alpha_1 \otimes \beta_1$ and $\alpha_2 \otimes \beta_2$ from $V \otimes W$ to itself which satisfy the identity

$$(\alpha_1 \otimes \beta_1) \circ (\alpha_2 \otimes \beta_2) = (\alpha_1 \circ \alpha_2) \otimes (\beta_1 \circ \beta_2).$$

In particular, for any $\alpha \in \mathrm{Hom}(V, V)$ and $\beta \in \mathrm{Hom}(W, W)$, we have

$$(\alpha \otimes 1_W) \circ (1_V \otimes \beta) = \alpha \otimes \beta = (1_V \otimes \beta) \circ (\alpha \otimes 1_W), \tag{3.2}$$

and hence the map $\tau(\alpha, \beta) = \alpha \otimes \beta$ gives the required action of $\mathrm{GL}(V) \times \mathrm{GL}(W)$ on $V \otimes W$.

3.3.2 Tensor products and Lie algebra representations Suppose now that $\mathfrak{g}_1, \mathfrak{g}_2$ are Lie algebras and (V_1, ρ_1) and (V_2, ρ_2) are representations of $\mathfrak{g}_1$ and $\mathfrak{g}_2$ respectively. We may consider the corresponding pair of questions for Lie algebras that we considered for groups above: can we build a representation of $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ from V_1 and V_2, and when $\mathfrak{g}_1 = \mathfrak{g}_2 = \mathfrak{g}$ can we also obtain a $\mathfrak{g}$-representation?

The map $x \mapsto (x, 0) + (0, x) = (x, x)$ is a homomorphism of Lie algebras $\mathfrak{g} \to \mathfrak{g} \oplus \mathfrak{g}$, thus using this "diagonal" map, we can, as above, reduce the second question to the first, namely can we build a representation of $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ from V_1 and V_2? Again, our candidate is $V_1 \otimes V_2$, and since $(\rho_1, \rho_2) \colon \mathfrak{g}_1 \oplus \mathfrak{g}_2 \to \mathfrak{gl}(V_1) \oplus \mathfrak{gl}(V_2)$ is a Lie algebra homomorphism, this question also reduces to whether there is a natural Lie algebra homomorphism $\mathfrak{gl}(V_1) \oplus \mathfrak{gl}(V_2) \to \mathfrak{gl}(V_1 \otimes V_2)$.

As with the group case, we first consider what it means to give a representation of a direct sum of Lie algebra $\mathfrak{g}_1 \oplus \mathfrak{g}_2$:

Lemma 3.3.2. *Let $\mathfrak{g}_1, \mathfrak{g}_2$ be Lie algebras and suppose that $\alpha_i \colon \mathfrak{g}_i \to \mathfrak{gl}(U)$ are Lie algebra homomorphisms. Then $\beta \colon \mathfrak{g}_1 \oplus \mathfrak{g}_2 \to \mathfrak{gl}(U)$ given by $\beta(x_1, x_2) = \alpha_1(x_1) + \alpha_2(x_2)$ is a Lie algebra homomorphism if and only if $[\alpha_1(x), \alpha_2(y)] = 0$ for all $(x, y) \in \mathfrak{g}_1 \times \mathfrak{g}_2$.*

Proof. This is a direct calculation. For all $(x_1, x_2), (y_1, y_2) \in \mathfrak{g}_1 \oplus \mathfrak{g}_2$ we have

$$\begin{aligned}
[\beta(x_1, x_2), \beta(y_1, y_2)] &= [\alpha_1(x_1) + \alpha_2(x_2), \alpha_1(y_1) + \alpha_2(y_2)] \\
&= [\alpha_1(x_1), \alpha_1(y_1)] + [\alpha_1(x_1), \alpha_2(y_2)] + [\alpha_2(x_2), \alpha_1(y_1)] + [\alpha_2(x_2), \alpha_2(y_2)] \\
&= [\alpha_1(x_1), \alpha_1(y_1)] + [\alpha_2(x_2), \alpha_2(y_2)] \\
&= \alpha_1([x_1, y_1]) + \alpha_2([x_2, y_2]) \\
&= \beta(([x_1, y_1], [x_2, y_2]) \\
&= \beta([(x_1, x_2), (y_1, y_2)])
\end{aligned}$$

where in passing from the second to the third equality we use the hypothesis applied to $(x.y) = (x_1, y_2)$ and $(y_1, x_2) \in \mathfrak{g}_1 \times \mathfrak{g}_2$. The converse follows similarly. □

Equipped with this observation, it follows that we again simply need, for $\alpha \in \mathfrak{gl}(V)$ and $\beta \in \mathfrak{gl}(W)$, to give action maps τ_V and τ_W on $V \otimes W$ which commute with each other. But by (3.2), we have

$$(\alpha \otimes 1) \circ (1 \otimes \beta) - (1 \otimes \beta) \circ (\alpha \otimes 1) = (\alpha \otimes \beta) - (\alpha \otimes \beta) = 0.$$

It therefore follows from Lemma 3.3.2 that $\eta_V(\alpha) = \alpha \otimes 1$ and $\eta_W(\beta) = 1 \otimes \beta$ give representations of $\mathfrak{gl}(V)$ and $\mathfrak{gl}(W)$ on $V \otimes W$ which commute with each other, and hence induce a representation of $\mathfrak{gl}(V) \oplus \mathfrak{gl}(W)$ on $V \otimes W$ as required.

Now returning to the general setting.

Definition 3.3.3. If (V, ρ) and (W, σ) are $\mathfrak{g}$-representations for an arbitrary Lie algebra $\mathfrak{g}$ then $V \otimes W$ becomes a $\mathfrak{g}$ representation via the composition

$$\mathfrak{g} \xrightarrow{\Delta} \mathfrak{g} \oplus \mathfrak{g} \xrightarrow{\rho \oplus \sigma} \mathfrak{gl}(V) \oplus \mathfrak{gl}(W) \xrightarrow{\tau_V \oplus \tau_W} \mathfrak{gl}(V \otimes W)$$

More explicitly (and this is the only formula you really need to remember from this section!) $V \otimes W$ becomes a $\mathfrak{g}$-representation via the map $\rho \otimes \sigma \colon \mathfrak{g} \to \mathfrak{gl}(V \otimes W)$ where

$$(\rho \otimes \sigma)(x)(v \otimes w) = \rho(x)(v) \otimes w + v \otimes \sigma(x)(w), \quad \forall v \in V, w \in W. \tag{3.3}$$

Remark 3.3.4. The discussion in the section is an attempt to explain how one might discover the action of a Lie algebra $\mathfrak{g}$ on a tensor product. On the other hand, if one simply guessed the formula in Equation (3.3), it is straight-forward to check directly that it does indeed give a Lie algebra homomorphism from $\mathfrak{g}$ to $\mathfrak{gl}(V \otimes W)$. It is a good exercise to do this computation for oneself.

Remark 3.3.5. In fact $\mathfrak{gl}(V_1) \otimes \mathfrak{gl}(V_2) \cong (V_1^* \otimes V_1) \otimes (V_2^* \otimes V_2) \cong (V_1 \otimes V_2)^* \otimes (V_1 \otimes V_2)$, so that $\mathfrak{gl}(V_1) \otimes \mathfrak{gl}(V_2) \cong \mathfrak{gl}(V_1 \otimes V_2)$, it can be checked that this implies that the image of the group homomorphism $\mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \to \mathrm{GL}(V_1 \otimes V_2)$ is equal to $P(V_1, V_2)$, the elements of $\mathrm{GL}(V_1 \otimes V_2)$ which preserve "pure tensors", that is, $g \in P(V_1, V_2)$ if for all $v_1 \in V_1, v_2 \in V_2$ we have $g(v_1 \otimes v_2) = w_1 \otimes w_2$ for some $w_1 \in V_1$, $w_2 \in V_2$.

3.3.3 Homomorphisms, $\mathfrak{g}$-homomorphism, and tensor products

The properties asserted of the maps described in this section are proved in detail in Appendix I.2.

An immediate consequence of the above definition is that, just as for group representations, if V and W are $\mathfrak{g}$-representations, then the isomorphism $\sigma\colon V \otimes W \to W \otimes V$ given by $\sigma(v \otimes w) = w \otimes v$, $(v \in V, w \in W)$ is compatible with the action of $\mathfrak{g}$ and hence induces an isomorphism of $\mathfrak{g}$-representations.

Let V and W be k-vector spaces. There is a natural linear map $\theta\colon V^* \otimes W \to \mathrm{Hom}(V, W)$, given by $\theta(f \otimes w) = f.w$ where $(f.w)(v) = f(v).w$ for all $v \in V, f \in V^*$ and $w \in W$. This map is injective, and its image is precisely the space of finite-rank linear maps[5] from V to W. In particular, if $\dim(V) < \infty$ then we have $\mathrm{End}(V) \cong V^* \otimes V$. Similarly, there is a natural map $m\colon V^* \otimes W^* \to (V \otimes W)^*$, where

$$m(f \otimes g)(v \otimes w) = f(v).g(w), \quad \forall v \in V, w \in W, f \in V^*, g \in W^*.$$

The map m is also injective and hence, by considering dimensions, it is an isomorphism when V and W are finite-dimensional. This tensor product description of $\mathrm{End}(V) = \mathrm{Hom}(V, V)$ gives a natural description of the trace map: Notice that we have a natural bilinear map $V^* \times V \to \mathsf{k}$ given by $(f, v) \mapsto f(v)$. By the universal property of the tensor product, this induces a linear map $\iota\colon V^* \otimes V \to \mathsf{k}$. Under the identification with $\mathrm{Hom}(V, V)$ this map is identified with the trace of a linear map.

Remark 3.3.6. It is worth noticing that this gives a coordinate-free way of defining the trace, and also some explanation for why one needs some finiteness condition in order for the trace to be defined.

Remark 3.3.7. If $\mathfrak{g}$ is a Lie algebra and V and W are $\mathfrak{g}$-representations, then the maps defined above – the trace map and the map identifying $V^* \otimes W$ with the finite rank linear maps from V to W, are maps of $\mathfrak{g}$-representations.

Example 3.3.8. If $\mathfrak{g}$ is a Lie algebra and (V, ρ) is a $\mathfrak{g}$-representation, then ρ induces a natural bilinear map $\mathfrak{g} \times V \to V$, namely $(x, v) \mapsto \rho(x)(v)$. By the universal property of tensor products this yields a linear map $\tilde{\rho}\colon \mathfrak{g} \otimes V \to V$. We claim this map is a homomorphism of $\mathfrak{g}$ representations (where $\mathfrak{g}$ is viewed as the adjoint representation). To see this, $a\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g} \otimes V)$ denote the action of $\mathfrak{g}$ on $\mathfrak{g} \otimes V$. Then we have, for $x, g \in \mathfrak{g}, v \in V$,

$$a(x)(g \otimes v) = [x, g] \otimes v + g \otimes \rho(x)(v),$$

and under $\tilde{\rho}$ this maps to

$$\rho([x, g])(v) + \rho(g) \circ \rho(x)(v) = [\rho(x), \rho(g)] + \rho(g)\rho(x)(v) = \rho(x) \circ \rho(g)(v) = \rho(x)(\tilde{\rho}(g \otimes v))$$

and hence $\tilde{\rho} \circ a = \rho \circ \tilde{\rho}$ as required.

It is also easy to check from the definitions that the natural map $\theta\colon V^* \otimes W \to \mathrm{Hom}(V, W)$ defined in Lemma 3.3 is also a map of $\mathfrak{g}$-representations, as is the contraction map $\iota\colon V^* \otimes V \to \mathsf{k}$, where we view k as the trivial representation of $\mathfrak{g}$. For example, for ι we have:

$$\iota(x(f \otimes v)) = \iota(x(f) \otimes v + f \otimes x(v)) = -f(x(v)) + f(x(v)) = 0, \quad \forall x \in \mathfrak{g}, v \in V, f \in V^*.$$

Thus all the maps between tensor products of vector spaces discuss in Appendix I.2 yield maps of $\mathfrak{g}$-representations.

[5] That is, the linear maps from V to W which have finite-dimensional image.

3.3.4 Tensoring with one-dimensional representations Let (L, ρ) be a one-dimensional representation of $\mathfrak{g}$, then $\mathfrak{gl}(L) \cong k = \mathfrak{gl}_1(\mathsf{k})$. It follows that ρ is identified with an element of $(\mathfrak{g}/D(\mathfrak{g}))^*$, and this clearly identifies (L, ρ) up to isomorphism. Conversely, if $\lambda \in (\mathfrak{g}/D(\mathfrak{g}))^*$, we write k_λ for the action of $\mathfrak{g}$ on k given by λ.

If (V, ρ) is any $\mathfrak{g}$-representation, then by Example I.8, we have an isomorphism of vector spaces $V \otimes \mathsf{k}_\lambda \to V$ given by the map $v \otimes \lambda \mapsto \lambda.v$. Via this map, one can think of the $\mathfrak{g}$-representation $V \otimes \mathsf{k}_\lambda$ as the same vector space V but now equipped with a new action ρ_λ of $\mathfrak{g}$, where $\rho_\lambda(x) = \rho(x) + \lambda(x).I_V$ (where we write I_V for the identity map.) Note that, in particular, if $\lambda, \mu \in (\mathfrak{g}/D(\mathfrak{g}))^*$ then this shows that $\mathsf{k}_\lambda \otimes \mathsf{k}_\mu \cong \mathsf{k}_{\lambda+\mu}$.

Chapter 4

Classifying Lie algebras

The goal of this course is to study the structure of Lie algebras, and attempt to classify them. The most ambitious "classification" result would be to give a description of all finite-dimensional Lie algebras up to isomorphism. In very low dimensions this is actually possible: For dimension 1 clearly there is a unique (up to isomorphism) Lie algebra since the alternating condition demands that the bracket is zero. In dimension two, one can again have an abelian Lie algebra, but there is another possibility: if $\mathfrak{g}$ has a basis $\{e,f\}$ then we may set $[e,f] = f$, and this completely determines the Lie algebra structure. All two-dimensional Lie algebras which are not abelian are isomorphic to this one (check this). It is also possible to classify three-dimensional Lie algebras, but it becomes rapidly intractable to do this in general as the dimension increases.

This reveals an essential tension in seeking any kind of classification result for mathematical objects: a classification result should describe all such objects (or at least those in a natural, and likely reasonably "large" class) up to some notion of equivalence. Clearly, using a stricter notion of equivalence will mean any classification theorem you can prove will provide finer information about the objects you are studying, but this must be balanced against the intrinsic complexity of the objects which may make such a classification (even for quite small classes) extremely complicated. Hence it is likely reasonable to accept a somewhat crude notion of equivalence in order to be have any chance of obtaining a classification theorem which has a relatively simple statement.

4.1 Classification by composition factors

Our approach will follow the strategy often used in finite groups: In that context, the famous Jordan-Hölder theorem shows that any finite group can be given by gluing together finite *simple* groups, in the sense that we may find an decreasing chain of subgroups

$$G = G_0 \rhd G_1 \rhd \ldots G_{n-1} \rhd G_n = \{e\},$$

where, for each i, $(1 \leq i \leq n)$, the subgroup G_i is a normal in G_{i-1} and $S_i = G_{i-1}/G_i$ is simple. That such a filtration of G exists is easy to prove by induction. The non-trivial part of the theorem is that, for any fixed finite simple group H, the number of S_i which are isomorphic to H is independent of the choice filtration. This is usually phrased as saying that the multiplicity with which a *composition factor* S_i occurs in the sequence $\{G_{i-1}/G_i : 1 \leq i \leq n\}$ is well-defined.

One can thus give a somewhat crude classification of finite groups, where one considers two finite groups to be equivalent if they have the same composition factors, by giving a classification of finite *simple* groups. But even the question of classifying finite simple groups is not at all obviously tractable, and answering it was one of the spectacular mathematical achievements of the second half of the twentieth century.

For Lie algebras, we can attempt something similar. In fact, it turns out that, at least in characteristic zero, we obtain a far more complete answer about the structure of an arbitrary finite-dimensional Lie algebra than one could hope to obtain in a Part C course on finite group theory.

One aspect of this finer information will reveal a sharp distinction between $\mathfrak{gl}_1$ and the non-abelian Lie algebras which have no proper ideals, which is one reason for the following definition:

Definition 4.1.1. A non-zero Lie algebra $\mathfrak{g}$ is said to be *almost simple*[1] if it has no proper ideals. If $\mathfrak{g}$ is almost simple and $\dim(\mathfrak{g}) > 1$ then we say that $\mathfrak{g}$ is *simple*. Equivalently, an almost simple Lie algebra is simple if it is non-abelian. Thus the only almost simple Lie algebra which is not simple is $\mathfrak{gl}_1$.

Here we give a proof of the Jordan-Holder theorem for finite dimensional Lie algebras over a field of arbitrary characteristic. The proof mirrors the case of finite groups. As we will see later, Cartan's criteria will give stronger results (though only in when working over fields of characteristic zero), so this result is only included for completeness.

Definition 4.1.2. A *composition series* for a finite dimensional Lie algbera $\mathfrak{g}$ is a chain

$$C = (\mathfrak{g} = \mathfrak{g}_0 \triangleright \mathfrak{g}_1 \triangleright \dots \triangleright \mathfrak{g}_r = 0)$$

of subalgebras such that, for $1 \leq i \leq r$, the subalgebra $\mathfrak{g}_i$ is an ideal in $\mathfrak{g}_{i-1}$ and the quotient $\mathfrak{g}_i/\mathfrak{g}_{i-1}$ is almost simple. The quotients $\mathfrak{g}_i/\mathfrak{g}_{i-1}$ are called the *composition factors* of the composition series. (Note that the $\mathfrak{g}_i$ are *not* necessarily ideals in $\mathfrak{g}$.)

It is straight-forward to check by induction on $\dim(\mathfrak{g})$ that any finite-dimensional Lie algebra has a composition series (one needs to use isomorphism theorems in the same manner as the proof in the context of finite groups – see Sheet 0 for more details). The following Lemma shows moreover that the property of possessing a composition series is inherited by ideals and quotients:

Lemma 4.1.3. *Suppose that* $\mathfrak{g}$ *has a composition series* $C = (\mathfrak{g} = \mathfrak{g}_0 \triangleright \mathfrak{g}_1 \triangleright \dots \triangleright \mathfrak{g}_n = 0)$ *and let* $\mathfrak{a}$ *be an ideal of* $\mathfrak{g}$. *Then* C *induces a composition series for* $\mathfrak{a}$*, and a composition series for the quotient* $\mathfrak{g}/\mathfrak{a}$.

Proof. Consider the intersection of $C \cap \mathfrak{a}$ of C with $\mathfrak{a}$, that is, $C \cap \mathfrak{a}$ is the sequence $(\mathfrak{a} = \mathfrak{a}_0 \geq \mathfrak{a}_1 \geq \dots \geq \mathfrak{a}_n = 0)$, where $\mathfrak{a}_i = \mathfrak{a} \cap \mathfrak{g}_i$. Note that its terms, while nested, need not be strictly decreasing. Since $\mathfrak{g}_i$ is an ideal in $\mathfrak{g}_{i-1}$, it is clear that $\mathfrak{a}_i$ is an ideal in $\mathfrak{a}_{i-1}$, and by the second isomorphism theorem $\mathfrak{a}_{i-1}/\mathfrak{a}_i \cong (\mathfrak{a} \cap \mathfrak{g}_{i-1} + \mathfrak{g}_i)/\mathfrak{g}_i$. Since $\mathfrak{g}_{i-1}/\mathfrak{g}_i$ is almost simple, it follows $\mathfrak{a}_{i-1}/\mathfrak{a}_i$ is either zero (that is $\mathfrak{a}_{i-1} = \mathfrak{a}_i$) or $\mathfrak{a}_{i-1}/\mathfrak{a}_i \cong \mathfrak{g}_{i-1}/\mathfrak{g}_i$. It follows that if we let $C_\mathfrak{a}$ be the sequence of subalgebras of $\mathfrak{a}$ obtained by omitting repetitions from $C \cap \mathfrak{a}$, then $C_\mathfrak{a}$ is a composition series for $\mathfrak{a}$. When $\mathfrak{a}$ is an ideal, then we may consider the sequence $C/\mathfrak{a} = ((\mathfrak{g}_i + \mathfrak{a})/\mathfrak{a})_{i=0}^r$, then the third isomorphism theorem shows that

$$((\mathfrak{g}_{i-1} + \mathfrak{a})/\mathfrak{a})/((\mathfrak{g}_i + \mathfrak{a})/\mathfrak{a}) \cong (\mathfrak{g}_{i-1} + \mathfrak{a})/(\mathfrak{g}_i + \mathfrak{a}),$$

which is either zero or isomorphic to $\mathfrak{g}_{i-1}/\mathfrak{g}_i$ (since it is the image of $\mathfrak{g}_{i-1}$ under $p\colon \mathfrak{g}/\mathfrak{g}_i \to \mathfrak{g}/(\mathfrak{g}_i + \mathfrak{a})$) the quotient map. Thus we see that, again by removing repetitions from $C/\mathfrak{a}$, we obtain a composition series $\bar{C}_\mathfrak{a}$ for $\mathfrak{g}/\mathfrak{a}$. □

4.2 The Jordan-Hölder Theorem for Lie algebras

The proofs in this subsection are not examinable.

Definition 4.2.1. If $\mathfrak{s}$ is an almost simple Lie algebra and $C = (\mathfrak{g}_i)_{i=0}^r$ is a composition series for a finite-dimensional Lie algebra $\mathfrak{g}$, define the multiplicity of $\mathfrak{s}$ in C to be $[\mathfrak{s}, C] = \#\{i \in \{1, \dots r\} : \mathfrak{s} \cong \mathfrak{g}_{i-1}/\mathfrak{g}_i\}$

The Jordan-Hölder Theorem shows that these multiplicities are in fact independent of C, so that we may define the multiplicity $[\mathfrak{s}, \mathfrak{g}]$ of $\mathfrak{s}$ in $\mathfrak{g}$ to be the number of times it occurs as a composition factor in any composition series of $\mathfrak{g}$.

[1]This is not standard terminology, but it is convenient to use here.

Lemma 4.2.2. *Let $\mathfrak{g}$ be a Lie algebra and let $\mathfrak{h}$ and $\mathfrak{m}$ be proper ideals such that $\mathfrak{h} \cap \mathfrak{m} = \{0\}$ and $\mathfrak{h}, \mathfrak{m}, \mathfrak{g}/\mathfrak{m}, \mathfrak{g}/\mathfrak{h}$ are all almost simple. Then $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{h}$ is a direct sum of almost simple ideals.*

Proof. Consider the map $q \colon \mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$. Since $\mathfrak{m} \cap \mathfrak{h} = \{0\}$, the restriction of q to $\mathfrak{m}$ is injective. But since $\mathfrak{g}/\mathfrak{h}$ is almost simple it follows that $\mathfrak{m} \cong \mathfrak{g}/\mathfrak{h}$. Moreover, since $[\mathfrak{h}, \mathfrak{m}] \subseteq \mathfrak{h} \cap \mathfrak{m} = \{0\}$, it follows that $\mathfrak{g} \cong \mathfrak{h} \oplus \mathfrak{g}/\mathfrak{h} \cong \mathfrak{h} \oplus \mathfrak{m}$. □

We can now prove the Jordan-Hölder Theorem:

Theorem 4.2.3. *(Jordan-Hölder Theorem) Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. Then if $\mathfrak{s}$ is an almost simple Lie algebra and $C_1 = (\mathfrak{g}_i)_{i=0}^r$ and $C_2 = (\mathfrak{h}_j)_{j=0}^s$ are composition series for $\mathfrak{g}$, we have $[\mathfrak{s}, C_1] = [\mathfrak{s}, C_2]$. In other words, the composition factors occuring in C_1 are, up to isomorphism and permutation, the same as those occuring in C_2.*

Proof. Let us prove this by induction on the dimension of $\mathfrak{g}$, the result being clear for $\dim(\mathfrak{g}) = 1$. Suppose therefore that $\mathfrak{g}$ has two composition series

$$C_1 = (0 = \mathfrak{h}_0 < \mathfrak{h}_1 < \dots < \mathfrak{h}_r = \mathfrak{g}); \quad C_2 = (0 = \mathfrak{m}_0 < \mathfrak{m}_1 < \dots < \mathfrak{m}_s = \mathfrak{g}).$$

If $\mathfrak{m}_{r-1} = \mathfrak{h}_{s-1}$, then applying induction to this ideal and the two composition series of it given by truncating the two series for $\mathfrak{g}$, thus it follows that $r-1 = s-1$, and the isomorphism classes of composition factors $\{\mathfrak{h}_{i+1}/\mathfrak{h}_i : 0 \leq i \leq r-2\}$ and $\{\mathfrak{m}_{j+1}/\mathfrak{m}_j : 0 \leq j \leq s-2\}$ are equal up to permutation and the remaining composition factors for the two series $\mathfrak{g}/\mathfrak{m}_{r-1}$, and $\mathfrak{g}/\mathfrak{h}_{s-1}$ are equal to each other, so we are done.

Now suppose that $\mathfrak{h}_{r-1} \neq \mathfrak{m}_{s-1}$, so that $\mathfrak{k} = \mathfrak{h}_{r-1} \cap \mathfrak{m}_{s-1}$ is a proper ideal in $\mathfrak{h}_{r-1}$ and $\mathfrak{m}_{s-1}$. Again by induction, the theorem holds for each of $\mathfrak{k}$, $\mathfrak{h}_{r-1}, \mathfrak{m}_{s-1}$. In particular, all composition series for $\mathfrak{h}_{r-1}$ have length $r-1$ and all composition series for $\mathfrak{m}_{s-1}$ have length $s-1$. Now by Lemma 4.1.3, the intersection

$$C_1 \cap \mathfrak{m}_{s-1} = (0 < \mathfrak{h}_0 \cap \mathfrak{m}_{s-1} < \dots < \mathfrak{h}_{r-1} \cap \mathfrak{m}_{s-1} = \mathfrak{k} < \mathfrak{m}_{s-1}).$$

yields, after removing repetitions where necessary, a composition series for $\mathfrak{k}$ and a composition series for $\mathfrak{m}_{s-1}$ which is one term longer. It follows that all composition series for $\mathfrak{k}$ have length $s-2$. However if we similarly consider $C_2 \cap \mathfrak{h}_{r-1}$ we see that $s-2 = r-2$ so that $s = r$, and hence the induced composition series of $\mathfrak{k}$ have no repetitions, and so the isomorphism classes of composition factors of C_1 and C_2 coincide except perhaps those in $\mathfrak{g}/\mathfrak{k}$. But these coincide by Lemma 4.2.2, so the proof is complete.

□

As noted above, it will turn out that in characteristic zero, the simple Lie algebras will all occur at "the top" of the composition series of a finite-dimensional Lie algebra, as a direct sum. The almost simple Lie algebra $\mathfrak{gl}_1$, however, can be glued to itself in non-trivial ways. Thus our study of the structure of Lie algebras therefore begins by examining Lie algebras which have only one isomorphism class of composition factor, namely $\mathfrak{gl}_1$. Before we do that, however, it seems useful to introduce the formalism of exact sequences:

4.3 Exact sequences of Lie algebras

Definition 4.3.1. We say that the sequence of Lie algebras and Lie algebra homomorphisms

$$\mathfrak{g}_1 \xrightarrow{i} \mathfrak{g} \xrightarrow{q} \mathfrak{g}_2$$

is *exact at* $\mathfrak{g}$ if $\mathrm{im}(i) = \ker(q)$. A sequence of maps

$$0 \longrightarrow \mathfrak{g}_1 \xrightarrow{i} \mathfrak{g} \xrightarrow{q} \mathfrak{g}_2 \longrightarrow 0$$

is called a *short exact sequence* if it is exact at each of $\mathfrak{g}_1$, $\mathfrak{g}$ and $\mathfrak{g}_2$, so that i is injective, q is surjective and $\mathrm{im}(i) = \ker(q)$. In this case, we say that $\mathfrak{g}$ is an *extension* of $\mathfrak{g}_2$ by $\mathfrak{g}_1$. The existence of a composition series for a finite-dimensional Lie algebra shows that any such Lie algebra is constructed through successive extensions by almost simple Lie algebras.

Two kinds of extensions of Lie algebras will arise naturally in this course:

4.3.1 Split extensions

Definition 4.3.2. An extension of Lie algebras

$$0 \longrightarrow \mathfrak{g}_1 \xrightarrow{i} \mathfrak{g} \xrightarrow{q} \mathfrak{g}_2 \longrightarrow 0$$

is said to be *split* if there is a homomorphism of Lie algebras $s\colon \mathfrak{g}_2 \to \mathfrak{g}$ such that $q \circ s = \mathrm{id}_{\mathfrak{g}_2}$.

Notice that in this case the image $s(\mathfrak{g}_2)$ of the splitting map s is a subalgebra of $\mathfrak{g}$ which is isomorphic to $\mathfrak{g}_2$ and is complementary to $i(\mathfrak{g}_1)$, in the sense that $\mathfrak{g} = i(\mathfrak{g}_1) \oplus s(\mathfrak{g}_2)$ as vector spaces. Indeed the homomorphism s is determined by $s(\mathfrak{g}_2)$ its image, because it is the inverse of $q_{|s(\mathfrak{g})}$, the restriction of q to that image. Moreover, since $i(\mathfrak{g}_1)$ is an ideal of $\mathfrak{g}$, the adjoint action of $\mathfrak{g}$ preserves $i(\mathfrak{g}_1)$, and so it restricts to give an action of $s(\mathfrak{g}_2)$ on $i(\mathfrak{g}_1)$. This completely describes the Lie bracket on $\mathfrak{g}$: For any $x, y \in \mathfrak{g}$, there are unique $x_1, y_1 \in \mathfrak{g}_1$ and $x_2, y_2 \in \mathfrak{g}_2$ such that $x = i(x_1) + s(x_2), y = i(y_1) + s(y_2)$. Then

$$\begin{aligned}[x, y] &= [i(x_1) + s(x_2), i(y_1) + s(y_2)] \\ &= i([x_1, x_2]) + \mathrm{ad}(s(x_2))(i(y_1)) - \mathrm{ad}(s(y_2))(i(x_1)) + s([x_2, y_2]).\end{aligned}$$

This motivates the following definition:

Definition 4.3.3. Suppose that $\mathfrak{g}$, $\mathfrak{h}$ are Lie algebras, and we have a homomorphism $\phi\colon \mathfrak{g} \to \mathrm{Der}_{\mathsf{k}}(\mathfrak{h})$, the Lie algebra of derivations[2] on $\mathfrak{h}$. Then it is straight-forward to check that we can form a new Lie algebra $\mathfrak{h} \rtimes \mathfrak{g}$, the *semi-direct product*[3] of $\mathfrak{g}$ and $\mathfrak{h}$ by ϕ which as a vector space is just $\mathfrak{g} \oplus \mathfrak{h}$, and where the Lie bracket is given by:

$$[(x_1, y_1), (x_2, y_2)] = ([x_1, x_2] + \phi(y_1)(x_2) - \phi(y_2)(x_1), [y_1, y_2]),$$

where $x_1, x_2 \in \mathfrak{h}, y_1, y_2 \in \mathfrak{g}$. The Lie algebra $\mathfrak{h}$, viewed as the subspace $\{(x, 0) : x \in \mathfrak{h}\}$ of $\mathfrak{h} \rtimes \mathfrak{g}$, is clearly an ideal of $\mathfrak{h} \rtimes \mathfrak{g}$. Since it does not intersect $\mathfrak{h}$, the quotient map $q\colon \mathfrak{h} \rtimes \mathfrak{g} \to (\mathfrak{h} \rtimes \mathfrak{g})/\mathfrak{h}$ induces an isomorphism $\mathfrak{g} \to (\mathfrak{h} \rtimes \mathfrak{g})/\mathfrak{h}$, hence $\mathfrak{h} \rtimes \mathfrak{g}$ is a split extension of $\mathfrak{g}$ by $\mathfrak{h}$. It is not difficult to check that any split extension is of this form.

Remark 4.3.4. In general, there may be many ways to split an exact sequence of Lie algebras (see Problem Sheet 1).

Example 4.3.5. Let $\mathfrak{s}_2$ be the 2-dimensional Lie algebra with basis $\{x, y\}$ and Lie bracket given by $[x, y] = y$. Then $\mathsf{k}.y$ is an ideal in $\mathfrak{s}_2$, and $\mathfrak{s}_2/\mathsf{k}.y$ is 1-dimensional, hence we have a short exact sequence:

$$0 \longrightarrow \mathfrak{gl}_1 \xrightarrow{i} \mathfrak{s}_2 \xrightarrow{q} \mathfrak{gl}_1 \longrightarrow 0$$

where $i(\lambda) = \lambda.y$ and $q(ax + by) = a$, for all $a, b, \lambda \in \mathsf{k}$. Now the map $s(\lambda) = \lambda.x$ is a Lie algebra homomorphism, hence the extension is split.

Note that $\mathrm{Der}_{\mathsf{k}}(\mathfrak{a}) = \mathfrak{gl}(\mathfrak{a})$ for an Abelian Lie algebra $\mathfrak{a}$, and so $\mathrm{Der}_{\mathsf{k}}(\mathfrak{gl}_1) = \mathfrak{gl}(\mathfrak{gl}_1) = \mathfrak{gl}_1$, and the map from $\mathfrak{gl}_1$ to $\mathrm{Der}_{\mathsf{k}}(\mathfrak{gl}_1)$ describing $\mathfrak{s}_2$ as a semi-direct product corresponds to the identity map under this identification.

[2]Recall that the derivations of a Lie algebra are the linear maps $\alpha\colon \mathfrak{h} \to \mathfrak{h}$ such that $\alpha([x, y]) = [\alpha(x), y] + [x, \alpha(y)]$.

[3]This is the Lie algebra analogue of the semidirect product of groups, where you build a group $H \rtimes G$ via a map from G to the automorphisms (rather than derivations) of H.

Remark 4.3.6. A short exact sequence of the form

$$0 \longrightarrow \mathfrak{g}_1 \xrightarrow{i} \mathfrak{g} \xrightarrow{q} \mathfrak{gl}_1 \longrightarrow 0$$

is automatically split. Indeed if we pick any $x \in \mathfrak{g}$ with $q(x) = 1 \in \mathfrak{gl}_1(\mathsf{k})$ then setting $s(\lambda) = \lambda.x$ it is immediate that $q \circ s = \mathrm{id}$. But since a Lie bracket is alternating, it always vanishes on any line, and hence s is a Lie algebra homomorphism. It follows that $\mathfrak{g}$ is a semidirect product $\mathfrak{g}_1 \rtimes \mathfrak{gl}_1(\mathsf{k})$.

Remark 4.3.7. If you took Part A Groups, you should note the analogy with the notion of a short exact sequence of groups, which is a sequence

$$1 \longrightarrow G_1 \xrightarrow{i} G \xrightarrow{q} G_2 \longrightarrow 1$$

where we write 1 for the trivial group (rather than 0 for the trivial Lie algebra). Exactness at G means that $\mathrm{im}(i) = \ker(q)$, and similarly at G_1 and G_2, so that i is injective and q is surjective. In Part A Groups you show that this sequence is split, that is, there exists a splitting map $s\colon G_2 \to G$ such that $q \circ s = \mathrm{id}_{G_2}$, if and only if $G \cong G_1 \rtimes G_2$.

4.3.2 Central extensions Another type of extension which plays an important role in our study of Lie algebras is a *central extension*. In this case, the Lie algebra $\mathfrak{g}_1$ in the sequence of Definition 4.3.1 is assumed to be central in $\mathfrak{g}$, that is $\mathfrak{g}_1 \subseteq \mathfrak{z}(\mathfrak{g})$, and hence in particular $\mathfrak{g}_1$ is Abelian. Picking a *linear* splitting $s\colon \mathfrak{g}_2 \to \mathfrak{g}$, we can write any $x, y \in \mathfrak{g}$ uniquely in the form $x = i(x_1) + s(x_2), y = i(x_2) + s(y_2)$, respectively. Thus, as $i(\mathfrak{g}_1)$ is central, the Lie bracket on $\mathfrak{g}$ is given by

$$[x,y] = [i(x_1) + s(x_2), i(x_2) + s(y_2)] = [s(x_2), s(y_2)] = i(\alpha(x_2, y_2)) + s([x_2, y_2])$$

where $\alpha(x,y) = ([x,y])_1$, that is, $i(\alpha(x_2,y_2))$ is the component of $[s(x_2), s(y_2)]$ in $i(\mathfrak{g}_1)$.

Definition 4.3.8. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra and let $\mathfrak{z}$ be a vector space. A *2-cocycle* on $\mathfrak{g}$ taking values in the vector space $\mathfrak{z}$ is a map $\alpha\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{z}$ satisfying the conditions:

i) $\alpha(x,x) = 0$, for all $x \in \mathfrak{g}$, (*i.e.* α is *alternating*);

ii) $\alpha(x,[y,z]) + \alpha(y,[z,x]) + \alpha(z,[x,y]) = 0, \quad \forall x,y,z \in \mathfrak{g}$.

Given such a cocycle, one can define a Lie algebra structure on the vector space $\mathfrak{z} \oplus \mathfrak{g}$ by setting

$$[(z_1, x_1), (z_2, x_2)] = (\alpha(x_1, x_2), [x_1, x_2]).$$

The resulting Lie algebra is a central extension of $\mathfrak{g}$. Picking a vector-space basis of $\mathfrak{z}$, say $\{e_1, \dots, e_k\}$, and writing α in terms of its components with respect to this basis, that is, $\alpha(x,y) = \sum_{j=1}^{k} \alpha_j(x,y).e_j$ one can immediately reduce the study of 2-cocycles to the study of k-valued 2-cocycles.

Example 4.3.9. It is straight-forward to understand central extensions of a Lie algebra $\mathfrak{g}$ by $\mathfrak{gl}_1$ in low dimensions. If $\mathfrak{g}$ is 1-dimensional, then the fact that α is alternating forces it to vanish, and hence the only central extension of $\mathfrak{gl}_1$ by $\mathfrak{gl}_1$ is the abelian Lie algebra $\mathfrak{gl}_1^{\oplus 2}$.

If $\dim(\mathfrak{g}) = 2$, then if $\mathfrak{g}$ is abelian then condition (*ii*) is automatically satisfied, and there is a unique non-zero alternating bilinear form up to isomorphism: if $\mathfrak{g}$ has basis $\{x, y\}$, then $\alpha(x,y) = 1 = -\alpha(y,x)$, defines a central extension of $\mathfrak{g}$.

Remark 4.3.10. Split and central extensions are in a loose sense complementary to each other: An extension of $\mathfrak{g}_2$ by $\mathfrak{g}_1$ which is both central and split is just the direct sum $\mathfrak{g}_1 \oplus \mathfrak{g}_2$, where $\mathfrak{g}_1 \cong \mathfrak{gl}_1^{\oplus k}$ and $k = \dim_{\mathsf{k}}(\mathfrak{g}_1)$

Chapter 5

Something from nothing: Solvable and nilpotent Lie algebras

Conventions: *From this point onwards in these notes, we will assume that all Lie algebras and all representations are finite-dimensional over* k*, unless the contrary is explicitly stated.*

We now begin to study particular classes of Lie algebras. The first class we study, solvable Lie algebras, in terms of the discussion on classification of Lie algebras in the previous section, can be given as the class of Lie algebras which can be built using only $\mathfrak{gl}_1$, the simplest Lie algebra[1] which possesses only the structure of the base field k and the trivial Lie bracket.

5.1 Definition and basic properties

Definition 5.1.1. A Lie algebra $\mathfrak{g}$ is *solvable* if its only composition factor is $\mathfrak{gl}_1(\mathsf{k})$. This is equivalent to the condition that $\mathfrak{g}$ has a nested sequence of subalgebras

$$\mathfrak{g} = \mathfrak{g}_0 \supsetneq \mathfrak{g}_1 \supsetneq \dots \supsetneq \mathfrak{g}_d = \{0\},$$

where $\mathfrak{g}_{k+1}$ is an ideal in $\mathfrak{g}_k$ and $\mathfrak{g}_k/\mathfrak{g}_{k+1}$ is abelian for each k $(0 \leq k \leq d-1)$. Indeed if such a sequence of subalgebras exists, any refinement of it to a composition series will have $\mathfrak{gl}_1(\mathsf{k})$ as its only composition factor, and conversely, a composition series with $\mathfrak{gl}_1(\mathsf{k})$ as its only composition factor is an example of such a sequence of subalgebras.

If $\mathfrak{g} = \mathfrak{g}_0 \supset \mathfrak{g}_1 \supset \dots \supset \mathfrak{g}_n = \{0\}$ is a composition series for $\mathfrak{g}$ with $\mathfrak{g}_k/\mathfrak{g}_{k+1} \cong \mathfrak{gl}_1$ for each $k \in \{0, 1, \dots, n-1\}$, so that $\dim(\mathfrak{g}) = n$, then we have $\mathfrak{g}_{n-1} \cong \mathfrak{gl}_1$, and, for each $k \in \{0, 1, \dots, n-1\}$, we have a short exact sequence

$$0 \longrightarrow \mathfrak{g}_{k+1} \xrightarrow{\iota_{k+1}} \mathfrak{g}_k \xrightarrow{q_k} \mathfrak{gl}_1 \longrightarrow 0$$

where ι_{k+1} is the inclusion map and q_k the quotient map. Thus $\mathfrak{g}_{k-1}$ is an extension of $\mathfrak{gl}_1$ by $\mathfrak{g}_k$. By Remark 4.3.7, this short exact sequence must split, and so $\mathfrak{g}_k$ is a semidirect product of $\mathfrak{g}_{k-1}$ by $\mathfrak{gl}_1(\mathsf{k})$, and so solvable Lie algebras are precisely the Lie algebras one obtains from the zero Lie algebra by taking iterated semidirect products with $\mathfrak{gl}_1(\mathsf{k})$.

Example 5.1.2. Example 4.3.5 shows that $\mathfrak{s}_2$, the 2-dimensional non-abelian Lie algebra, is solvable.

Definition 5.1.3. We can rephrase the condition that a Lie algebra $\mathfrak{g}$ is solvable in terms of a decreasing sequence of ideals in $\mathfrak{g}$: The *derived subalgebra*[2] $D(\mathfrak{g})$ of $\mathfrak{g}$ is defined to be $[\mathfrak{g}, \mathfrak{g}]$ (an ideal in $\mathfrak{g}$ since $\mathfrak{g}$ is). Inductively we define $D^k(\mathfrak{g}) = D(D^{k-1}(\mathfrak{g})) = [D^{k-1}(\mathfrak{g}), D^{k-1}(\mathfrak{g})]$ for each $k \geq 1$. The sequence of ideals $(D^k(\mathfrak{g}))_{k\geq 0}$ is called the *derived series* of $\mathfrak{g}$. Note that, since $\mathfrak{g}$ is an ideal in $\mathfrak{g}$, it follows by induction on k that $D^k(\mathfrak{g}) = [D^{k-1}(\mathfrak{g}), D^{k-1}(\mathfrak{g})]$ is an ideal in $\mathfrak{g}$.

[1] Hence starting with nothing...

[2] Oddly, it is not known as the derived ideal, even though it is indeed an ideal.

Lemma 5.1.4. *Let* $\mathfrak{g}$ *be a Lie algebra. Then* $D(\mathfrak{g})$ *is the smallest ideal in* $\mathfrak{g}$ *such that* $\mathfrak{g}/D(\mathfrak{g})$ *is abelian. In particular,* $\mathfrak{g}$ *is solvable precisely when the derived series* $(D^k(\mathfrak{g}))_{k\geq 1}$ *satisfies* $D^k(\mathfrak{g}) = 0$ *for sufficiently large* k*.*

Proof. For the first claim, suppose that I is an ideal for which $\mathfrak{g}/I$ is abelian. Then, for all $x, y \in \mathfrak{g}$, we must have $[x, y] \in I$, and hence $D(\mathfrak{g}) \subseteq I$. Since this also shows $\mathfrak{g}/D(\mathfrak{g})$ is abelian, the claim follows.

Next note that we have a short exact sequence

$$0 \longrightarrow D(\mathfrak{g}) \longrightarrow \mathfrak{g} \longrightarrow \mathfrak{g}/D\mathfrak{g} \longrightarrow 0$$

that is, $\mathfrak{g}$ is an extension of the abelian Lie algebra $\mathfrak{g}/D(\mathfrak{g})$ by $D(\mathfrak{g})$. It follows that if $D^k(\mathfrak{g}) = \{0\}$ for some k, then $\mathfrak{g}$ has a filtration by ideals for which the subquotients are abelian, so it is certainly solvable. Conversely, if $\mathfrak{g}$ is solvable, so that we have a nested sequence of subalgebras $\mathfrak{g} = \mathfrak{g}_0 \supset \mathfrak{g}_1 \supset \ldots \supset \mathfrak{g}_n = \{0\}$, where $\mathfrak{g}_{i+1}$ is an ideal in $\mathfrak{g}_i$ and $\mathfrak{g}_i/\mathfrak{g}_{i+1}$ is abelian. But then $D(\mathfrak{g}_i) = [\mathfrak{g}_i, \mathfrak{g}_i] \subseteq \mathfrak{g}_{i+1}$, and so since $\mathfrak{g} = \mathfrak{g}_0$, by induction it follows that $D^k(\mathfrak{g}) \subseteq \mathfrak{g}_k$, and hence for $k \geq n$ we have $D^k(\mathfrak{g}) = 0$. □

Remark 5.1.5. Because the terms of the derived series are ideals in $\mathfrak{g}$, it follows that if $\mathfrak{g}$ is solvable, then there is a filtration of $\mathfrak{g}$ by ideals not just subalgebras which are each an ideal in the previous term of the filtration. In particular, if $\mathfrak{g}$ is solvable, it follows $\mathfrak{g}$ has an non-trivial abelian ideal, since the last non-zero term of the derived series must be such an ideal.

We now give a fundamental family of solvable Lie algebras. To describe them we need the following definition:

Definition 5.1.6. Let V be a vector space, and let $\mathcal{F} = (F_i)_{i=0}^k$ be a *flag* in V, that is

$$\mathcal{F} = (0 = F_0 \subset F_1 \subset F_2 \subset \ldots \subset F_k = V)$$

is a nested sequence of subspaces with $\dim(F_{i-1}) < \dim(F_i)$ for $1 \leq i \leq k$. If $\mathcal{F}^1$ and $\mathcal{F}^2$ are flags in V then we say that $\mathcal{F}^2$ is a *refinement* of $\mathcal{F}^1$ if every subspace in $\mathcal{F}^1$ occurs in $\mathcal{F}^2$. If $\dim(F_i) = i$ for all i (so that $\dim(V) = k$) then $\mathcal{F}$ is called a *complete flag* (as it cannot be refined any further). It is clear (since any linearly independent set can be extended to a basis) that any flag can be refined to a complete flag.

Example 5.1.7. Let V be a finite dimensional vector space and $\mathcal{F} = (0 = F_0 < F_1 < \ldots < F_n = V)$ be a complete flag in V. Let, for $0 \leq k \leq n$,

$$\mathfrak{b}^k_{\mathcal{F}} = \{x \in \mathfrak{gl}(V) : x(F_i) \subseteq F_{i-k},\ \forall i, k \leq i \leq n\},$$

Thus if we set $\mathfrak{b}_{\mathcal{F}} = \mathfrak{b}^0_{\mathcal{F}}$, then $\mathfrak{b}_{\mathcal{F}}$ is the associative subalgebra of $\mathrm{End}(V)$ consisting of the endomorphisms which preserve all of the subspaces in the complete flag $\mathcal{F}$.

i) $D(\mathfrak{b}_{\mathcal{F}}) \subseteq \mathfrak{b}^1_{\mathcal{F}}$,

ii) $[\mathfrak{b}^k_{\mathcal{F}}, \mathfrak{b}^l_{\mathcal{F}}] \subseteq \mathfrak{b}^{k+l}_{\mathcal{F}}, \quad \forall k, l \geq 1.$

Given the claims, it follows by an easy induction that $D^k(\mathfrak{b}_{\mathcal{F}}) \subseteq \mathfrak{b}^{2^{k-1}}_{\mathcal{F}}$, and hence, since $\mathfrak{b}^m_{\mathcal{F}} = 0$ for $m \geq \dim(V)$, it follows that $\mathfrak{b}_{\mathcal{F}}$ is solvable.

For *i*), suppose first that $x, y \in \mathfrak{b}_{\mathcal{F}}$ and consider $[x, y]$. We need to show that $[x, y](F_i) \subset F_{i-1}$ for each i, $1 \leq i \leq n$. Since $x, y \in \mathfrak{b}_{\mathcal{F}}$, certainly we have $[x, y](F_i) \subseteq F_i$ for all $i, 1 \leq i \leq n$, thus it is enough to show that the map $\overline{[x, y]}$ induced by $[x, y]$ on F_i/F_{i-1} is zero. But this map is the commutator of the maps induced by x and y in $\mathrm{End}(F_i/F_{i-1})$, which since F_i/F_{i-1} is one-dimensional, is abelian, so that all commutators are zero.

For *ii*) note that if $x \in \mathfrak{b}^k_{\mathcal{F}}$, $y \in \mathfrak{b}^l_{\mathcal{F}}$ where $k, l \geq 0$, then the compositions xy and yx both lie in $\mathfrak{b}^{k+l}_{\mathcal{F}}$. It follows that $[x, y] \in \mathfrak{b}^{k+l}_{\mathcal{F}}$ as required.

We will see shortly that, in characteristic zero, any solvable linear Lie algebra $\mathfrak{g} \subset \mathfrak{gl}(V)$, where V is finite dimensional, is a subalgebra of $\mathfrak{b}_{\mathcal{F}}$ for some complete flag $\mathcal{F}$. We next note some basic properties of solvable Lie algebras.

Lemma 5.1.8. *Let $\mathfrak{g}$ be a Lie algebra, $\phi\colon \mathfrak{g} \to \mathfrak{h}$ a homomorphism of Lie algebras.*

1. *We have $\phi(D^k\mathfrak{g}) = D^k(\phi(\mathfrak{g}))$. In particular $\phi(\mathfrak{g})$ is solvable if $\mathfrak{g}$ is, thus any quotient of a solvable Lie algebra is solvable.*
2. *If $\mathfrak{g}$ is solvable then so are all subalgebras of $\mathfrak{g}$.*
3. *If $im(\phi)$ and $ker(\phi)$ are solvable then so is $\mathfrak{g}$. Thus if I is an ideal and I and $\mathfrak{g}/I$ are solvable, so is $\mathfrak{g}$.*

Proof. The first two statements are immediate from the definitions. For the third, note that if $\mathrm{im}(\phi)$ is solvable, then for some N we have $D^N\mathrm{im}(\phi) = \{0\}$, so that by part (1) we have $D^N(\mathfrak{g}) \subset \ker(\phi)$, hence if $D^M\ker(\phi) = \{0\}$ we must have $D^{N+M}\mathfrak{g} = \{0\}$ as required. □

5.2 Representations of solvable Lie algebras

In this section we will assume that our field k is algebraically closed of characteristic zero.

5.2.1 Lie's theorem Our first goal is the following theorem:

Theorem 5.2.1. *(Lie's theorem) Let $\mathfrak{g}$ be a solvable Lie algebra and V is a $\mathfrak{g}$-representation. Then there is a homomorphism $\lambda\colon \mathfrak{g} \to \mathfrak{gl}_1(\mathsf{k})$ and a nonzero vector $v \in V$ such that $x(v) = \lambda(x).v$ for all $x \in \mathfrak{g}$. Equivalently, any finite-dimensional irreducible representation of a solvable Lie algebra is one-dimensional.*

Proof. We first explain the equivalence asserted in the last sentence of the statement. Note that the existence of a non-zero $v \in V$ such that $x(v) = \lambda(x).v$ for all $x \in \mathfrak{g}$ is equivalent to the assertion that the line $\mathsf{k}.v$ is a subrepresentation of V. Thus the statement of the theorem shows that any representation contains a one-dimensional subrepresentation, and hence any irreducible representation must itself be one-dimensional. Since any representation contains an irreducible representation, the equivalence follows.

To establish the statement, we use induction on $\dim(\mathfrak{g})$. If $\dim(\mathfrak{g}) = 1$, then $\mathfrak{g} = \mathsf{k}.x$ for any nonzero $x \in \mathfrak{g}$, and since k is algebraically closed, x has an eigenvector in V and we are done. For $\dim(\mathfrak{g}) > 1$, since $\mathfrak{g}$ is solvable, we may write it as an extension

$$0 \longrightarrow I \longrightarrow \mathfrak{g} \longrightarrow \mathfrak{gl}_1 \longrightarrow 0$$

where I is an ideal of codimension[3] 1. Picking $x \notin I$, we may thus write $\mathfrak{g} = \mathsf{k}.x \oplus I$ (a direct sum of vector spaces, but a semi-direct product of I and $\mathfrak{gl}_1(\mathsf{k})$ as Lie algebras).

Now by induction, V contains a non-zero vector $v \in V$ with the property that $\mathsf{k}.v$ is preserved by the action of I. Thus, for any $h \in I$ we have $h(v) = \nu(h).v$ for some homomorphism of Lie algebras $\nu\colon I \to \mathfrak{gl}_1$.

Let $U = \{w \in V : h(w) = \nu(h).w, \forall h \in I\}$, a non-zero subspace of V since by definition $v \in U$. Now recall $\mathfrak{g} = \mathsf{k}.x \oplus I$. Provided the action of x preserves U, then any eigenvector for x in U satisfy the conditions of the theorem. Indeed if $u \in U$ is such an eigenvector, so that $x(u) = \mu.u$ for some $\mu \in \mathsf{k}$ then, for any $a \in \mathsf{k}$ and $h \in I$, we define $\lambda(a.x + h) = a.\mu + \nu(h)$.

To see if this is indeed the case, note that we have

$$\begin{aligned} h(x(w)) &= [h,x](w) + xh(w) \\ &= \lambda([h,x])(w) + \lambda(h).x(w) \end{aligned}$$

[3] Recall that a subspace U of a vector space V has codimension d if $\dim(V/U) = d$.

Thus if we are to conclude that the action of x preserves U, and thus complete the proof, it remains to show that $\lambda([h,x]) = 0$, that is, we must show that λ vanishes on $[\mathfrak{g}, I]$. This is the content of the next Lemma.

□

Remark 5.2.2. The following Lemma completes the proof of Lie's theorem. It relies on a trick which permeates the course, namely that one can often compute a trace in two different ways to obtain important information. One way will be by observing that one is computing the trace of a commutator, which is therefore zero. The other will, in one fashion or another, follow from consideration of the generalised eigenspaces of the linear map in question.

Lemma 5.2.3. *(Lie's Lemma) Let $\mathfrak{g}$ be a Lie algebra and let $I \subset \mathfrak{g}$ be an ideal, and V a finite dimensional representation. Suppose $v \in V$ is a vector such that $x(v) = \lambda(x).v$ for all $x \in I$, where $\lambda\colon I \to \mathfrak{gl}_1(\mathsf{k})$. Then λ vanishes on $[\mathfrak{g}, I] \subset I$.*

Proof. Let $x \in \mathfrak{g}$. For each $m \in \mathbb{N}$, let $W_m = \operatorname{span}\{v, x(v), \dots, x^m(v)\}$. The W_m form a nested sequence of subspaces of V. We claim that $hx^m(v) \in \lambda(h)x^m v + W_{m-1}$ for all $h \in I$ and $m \geq 0$. Using induction on m, the claim being immediate for $m = 0$, note that

$$\begin{aligned} hx^m(v) &= [h,x]x^{m-1}(v) + xhx^{m-1}(v) \\ &\in (\lambda([h,x])x^{m-1}v + W_{m-2}) + x(\lambda(h)x^{m-1}(v) + W_{m-2}) \\ &\in \lambda(h)x^m(v) + W_{m-1}, \end{aligned}$$

where in the second equality we use induction on m for both $h, [h,x] \in I$.

Now since V is finite dimensional, there is a maximal n such that the vectors $\{v, x(v), \dots, x^n(v)\}$ are linearly independent, and so $W_m = W_n$ for all $m \geq n$. It then follows that W_n is preserved by x, and from the claim it follows that W_n is also preserved by every $h \in I$. Moreover, the claim also shows that for any $h \in I$ the matrix of $[x,h]$ with respect to the basis $\{v, x(v) \dots, x^n(v)\}$ of W_n is upper triangular with diagonal entries all equal to $\lambda([x,h])$. It follows that $\operatorname{tr}([x,h]) = (n+1)\lambda([x,h])$. Since the trace of a commutator is zero[4], it follows that $(n+1)\lambda([x,h]) = 0$, and so since $\operatorname{char}(\mathsf{k}) = 0$ we conclude that $\lambda([x,h]) = 0$. □

Corollary 5.2.4. *Let $\mathfrak{g}$ be a solvable Lie algebra and let (V, ρ) be a $\mathfrak{g}$-representation. Then there is a complete flag $\mathcal{F} = (V = F_0 \supset F_1 \supset \dots \supset F_d = \{0\})$ where each F_i is a $\mathfrak{g}$-subrepresentation. In particular, if $\mathfrak{g}$ is solvable, then it has a composition series each of whose terms is an ideal in all of $\mathfrak{g}$.*

Proof. Take any composition series for V. Since Lie's theorem shows that the irreducible representations of $\mathfrak{g}$ are all one-dimensional, the resulting chain of subrepresentations will form a complete flag. The final sentence follows by applying this to the adjoint representation $(\mathfrak{g}, \mathrm{ad})$, since $I \leq \mathfrak{g}$ is an ideal in $\mathfrak{g}$ is and only if it is a subrepresentation of the adjoint representation. □

Remark 5.2.5. Note that we can rephrase the statement of the Corollary in terms of Example 5.1.7: If $\mathfrak{g}$ is solvable and (V, ρ) is any $\mathfrak{g}$-representation, then by Lemma 5.1.8, the image $\mathfrak{g}_1 = \rho(\mathfrak{g})$ is solvable. Then Corollary 5.2.4 shows that there is a complete flag $\mathcal{F}$ in V such that $\mathfrak{g}_1 \subseteq \mathfrak{b}_{\mathcal{F}}$

Recall from Example 3.1.4 that the isomorphism classes of one-dimensional representations of a Lie algebra $\mathfrak{g}$ are given by the elements of $(\mathfrak{g}/D(\mathfrak{g}))^*$: if $\lambda \in (\mathfrak{g}/D\mathfrak{g})^*$, via the identification $(\mathfrak{g}/D(\mathfrak{g}))^* \cong D(\mathfrak{g})^0$, we may view λ as a linear map $\lambda\colon \mathfrak{g} \to \mathsf{k}$ vanishing on $D(\mathfrak{g})$, and that vanishing precisely ensures it is a homomorphism from $\mathfrak{g}$ to $\mathfrak{gl}_1(\mathsf{k})$. Recall that we write k_λ for the representation (k, λ).

[4]It is important here that $\rho([x,h])$ is the commutator of $\rho(x)$ and $\rho(h)$ *both of which preserve* W_n – by the claim in the case of $\rho(h)$, and by our choice of n in the case of $\rho(x)$ – in order to conclude the trace is zero.

Definition 5.2.6. We will refer to an element of $(\mathfrak{g}/D(\mathfrak{g}))^*$ (equivalently, an isomorphism class of 1-dimensional $\mathfrak{g}$-representations) as a *weight* of $\mathfrak{g}$. In the case where $\mathfrak{g}$ is solvable, Lie's theorem shows that the weights are exactly the isomorphism classes of irreducible $\mathfrak{g}$-representations. Note that if $\lambda, \mu \in (\mathfrak{g}/D(\mathfrak{g}))^*$, then the addition in the vector space $(\mathfrak{g}/D(\mathfrak{g}))^*$ makes it an abelian group. This abelian group structure can also be seen from the point of view of one-dimensional representations: since the tensor product of 1-dimensional vector spaces is 1-dimensional, the tensor product restricts to an operation on 1-dimensional vector spaces. This gives the set of one-dimensional representations the structure of an abelian group: it is abelian because the map induced by interchange factors gives an isomorphism of $\mathfrak{g}$-representations $L_1 \otimes L_2 \cong L_2 \otimes L_1$ and if L is any one-dimensional representation then $L \otimes L^* \cong \mathsf{k}_0$ via the evaluation (or contraction) map induced by the natural bilinear pairing $L \times L^* \to \mathsf{k}$.

Since a direct calculation shows that $\mathsf{k}_\lambda \times \mathsf{k}_\mu \cong \mathsf{k}_{\lambda+\mu}$, this abelian group structure becomes the vector addition under the identification of the set of isomorphism classes of 1-dimensional representations with $(\mathfrak{g}/D(\mathfrak{g}))^*$.

5.3 Nilpotent Lie algebras

In this section we continue our study of Lie algebras which are built from $\mathfrak{gl}_1$, but now by using *central extensions* rather than arbitrary extensions.

Definition 5.3.1. A Lie algebra $\mathfrak{g}$ is said to be *nilpotent* if it can be obtained from 0, the trivial Lie algebra, by iterated central extensions. If $\mathfrak{g}$ can be obtained by precisely k iterated extentions, we say $\mathfrak{g}$ is *k-step nilpotent*. Thus, for example, a Lie algebra is 1-step nilpotent if and only if it is abelian.

To make this more concrete, suppose that $\mathfrak{g}$ is a nilpotent Lie algebra. Then, for some $k \geq 0$ there are Abelian Lie algebras $(\mathfrak{c}_i)_{i=0}^k$ and, for each $i \geq 1$ a short exact sequence

$$0 \longrightarrow \mathfrak{c}_i \xrightarrow{p_i} \mathfrak{g}_i \xrightarrow{q_i} \mathfrak{g}_{i-1} \longrightarrow 0$$

where $\mathfrak{g}_0 = \mathfrak{c}_0$ and $\mathfrak{c}_i \subseteq \mathfrak{z}(\mathfrak{g}_i)$, that is, $\mathfrak{g}_i$ is a central extension of $\mathfrak{g}_{i-1}$ by $\mathfrak{c}_i$. and $\mathfrak{g} = \mathfrak{g}_k$. It follows that $q_k : \mathfrak{g} = \mathfrak{g}_k \to \mathfrak{g}_{k-1}$, and if we set $Q_i = q_i \circ q_{i+1} \circ \ldots \circ q_k$, then $Q_i : \mathfrak{g} \to \mathfrak{g}_i$ exhibits $\mathfrak{g}_i$ as a quotient of $\mathfrak{g}$. Set $\mathfrak{q}_i = \ker(Q_i)$, so that if we set $\mathfrak{q}_0 = \mathfrak{g}$, then $(\mathfrak{q}_i)_{i=0}^k$ gives a descending sequence of ideals in $\mathfrak{g}$, and $\mathfrak{q}_i/\mathfrak{q}_{i-1} \cong \mathfrak{c}_i$ is central in $\mathfrak{g}/\mathfrak{q}_{i-1}$. The sequence of central extensions constructing $\mathfrak{g}$ can thus be reconstructed from the sequence of ideals $(\mathfrak{q}_i)_{i=0}^k$.

Definition 5.3.2. For $\mathfrak{g}$ a Lie algebra, let $C^0(\mathfrak{g}) = \mathfrak{g}$, and $C^i(\mathfrak{g}) = [\mathfrak{g}, C^{i-1}(\mathfrak{g})]$ for $i \geq 1$. This sequence of ideals of $\mathfrak{g}$ is called the *lower central series* of $\mathfrak{g}$.

Remark 5.3.3. Notice that $C^1(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}]$ is the *derived subalgebra*[5] of $\mathfrak{g}$ and, as we have seen, this is also denoted[6] $D(\mathfrak{g})$ and sometimes $\mathfrak{g}'$.

Proposition 5.3.4. *Suppose that $\mathfrak{g}$ is nilpotent and $(\mathfrak{q}_i)_{i=0}^k$ the sequence of ideals associated to a realization of $\mathfrak{g}$ as an iterated sequence of central extensions. Then*

1. *For each $i \geq 0$ we have $C^i(\mathfrak{g}) \subseteq \mathfrak{q}_i$ and hence $C^k(\mathfrak{g}) = 0$.*
2. *Conversely, if $\mathfrak{g}$ is such that, for some $N \geq 0$ we have $C^N(\mathfrak{g}) = 0$, then $\mathfrak{g}$ is at most N-step nilpotent.*

Proof. Suppose $\mathfrak{g}$ is any Lie algebra, and $\mathfrak{b} \subseteq \mathfrak{a}$ are ideals in $\mathfrak{g}$. If $\mathfrak{a}/\mathfrak{b}$ is central in $\mathfrak{g}/\mathfrak{b}$, then for any $x \in \mathfrak{g}$ and $y \in \mathfrak{a}$ we must have $[x, y] \in \mathfrak{b}$ and hence $[\mathfrak{g}, \mathfrak{a}] \subseteq \mathfrak{b}$. Since $\mathfrak{a}/[\mathfrak{g}, \mathfrak{a}]$ is certainly central in $\mathfrak{g}/[\mathfrak{g}, \mathfrak{a}]$ it follows that $[\mathfrak{g}, \mathfrak{a}]$ is the smallest ideal of $\mathfrak{g}$ contained in $\mathfrak{a}$ for which $\mathfrak{a}$ becomes central in the quotient algebra.

[5] Oddly, not as the *derived ideal* even though it is an ideal.

[6] Partly just to cause confusion, but also because it comes up a lot, playing slightly different roles, which leads to the different notation. We'll see it again shortly in a slightly different guise.

Applying this observation to $C^i(\mathfrak{g})$ inductively yields 1). For the converse. observe that the previous paragraph also shows that

$$0 \longrightarrow C^k(\mathfrak{g})/C^{k+1}(\mathfrak{g}) \longrightarrow \mathfrak{g}/C^{k+1}(\mathfrak{g}) \longrightarrow \mathfrak{g}/C^k(\mathfrak{g}) \longrightarrow 0$$

shows that $\mathfrak{g}/C^{i+1}(\mathfrak{g})$ is a central extension of $\mathfrak{g}/C^i(\mathfrak{g})$. It follows that if $C^N(\mathfrak{g}) = 0$ for some N then $\mathfrak{g}$ is at most N-step nilpotent. □

Lemma 5.3.5. *Let $\mathfrak{g}$ be a Lie algebra. Then*

1. *If $\mathfrak{g}$ is nilpotent, any subalgebra or quotient of $\mathfrak{g}$ is nilpotent.*
2. *If $\mathfrak{g}$ is nilpotent, then the centre $\mathfrak{z}(\mathfrak{g})$ is non-zero if $\mathfrak{g}$ is. Moreover, $\mathfrak{g}/\mathfrak{z}(\mathfrak{g})$ is nilpotent if and only if $\mathfrak{g}$ is.*

Proof. For 1) we use induction on $\dim(\mathfrak{g})$. If $\mathfrak{g}$ is Abelian, the result is trivial, so we may suppose that $\mathfrak{g}$ is a central extension

$$0 \longrightarrow \mathfrak{c} \longrightarrow \mathfrak{g} \longrightarrow \mathfrak{q} \longrightarrow 0$$

where $\mathfrak{c}$ is central. If $\mathfrak{h}$ is a subalgebra, then we obtain an induced short exact sequence

$$0 \longrightarrow \mathfrak{c} \cap \mathfrak{h} \longrightarrow \mathfrak{h} \longrightarrow (\mathfrak{h} + \mathfrak{c})/\mathfrak{c} \longrightarrow 0$$

But since $\dim(\mathfrak{q}) < \dim(\mathfrak{g})$, by induction $(\mathfrak{h} + \mathfrak{c})/\mathfrak{c})$ is nilpotent as it is a subalgebra of $\mathfrak{g}/\mathfrak{c} \cong \mathfrak{q}$. Hence $\mathfrak{h}$ is nilpotent also (as it is either isomorphic to $(\mathfrak{h} + \mathfrak{c})/\mathfrak{c}$ or it is a central extension of it).

Part (2) is trivial since a non-trivial central extension always has a non-trivial centre.

□

Remark 5.3.6. Notice that if $\mathfrak{a}$ is an arbitrary ideal in $\mathfrak{g}$, and $\mathfrak{a}$ and $\mathfrak{g}/\mathfrak{a}$ are nilpotent it does *not* follow that $\mathfrak{g}$ is nilpotent. Indeed recall from Example 4.3.5 the non-Abelian 2-dimensional Lie algebra $\mathfrak{s}_2$, with basis $\{x, y\}$ where $[x, y] = y$. Then $\mathsf{k}.y$ is a 1-dimensional ideal in $\mathfrak{s}_2$ but it is not central. Indeed $\mathfrak{z}(\mathfrak{s}_2) = 0$ so $\mathfrak{s}_2$ is not nilpotent, even though the ideal $\mathsf{k}.y$ and the quotient $\mathfrak{s}_2/\mathsf{k}.y$ are (since they are both abelian). Note that this shows that $\mathfrak{s}_2$ cannot be written as a central extension of $\mathfrak{gl}_1$ by itself.

Remark 5.3.7. The characterisation of the property of nilpotence in terms of the lower central series is similar to the characterisation of solvable Lie algebras in terms of the derived series. This is one reason it is commonly used. There is, however, another nature nested sequence of ideals which can be used to characterize nilpotence: If $\mathfrak{g}$ is any Lie algebra, set $Z^0(\mathfrak{g}) = \mathfrak{g}$, and, assuming $Z^k(\mathfrak{g})$ is defined, let $q_k \colon \mathfrak{g} \to \mathfrak{g}/Z^k(\mathfrak{g})$ be the quotient map, and set $Z^{k+1}(\mathfrak{g}) = q_k^{-1}(\mathfrak{z}(\mathfrak{g}_k))$. This process yields an *increasing* sequence of ideals of $\mathfrak{g}$ known as the *upper central series*. If it exhausts $\mathfrak{g}$, that is, if for some $n \geq 0$ we have $Z^k(\mathfrak{g}) = \mathfrak{g}$ for all k large enough, the $\mathfrak{g}$ is nilpotent. If $\mathfrak{g}$ is not nilpotent, the upper central series will stabilize at a maximal nilpotent ideal of $\mathfrak{g}$.)

We now wish to show that simple considerations from linear algebra give us a large supply of nilpotent Lie algebras.

Definition 5.3.8. Let V be a vector space and suppose that $\mathcal{F}$ is a flag in V. We set

$$\mathfrak{n}_{\mathcal{F}} = \mathfrak{b}^1_{\mathcal{F}} = \{x \in \mathfrak{gl}(V) : x(F_i) \subseteq F_{i-1}, \forall i \in \{1, 2, \dots, k\}\}.$$

It is easy to see that $\mathfrak{n}_{\mathcal{F}}$ is an associative subalgebra of $\mathrm{End}(V)$, and hence a Lie subalgebra of $\mathfrak{gl}(V)$. Moreover, if $\mathcal{F}'$ refines $\mathcal{F}$, then $\mathfrak{n}_{\mathcal{F}} \leq \mathfrak{n}_{\mathcal{F}'}$.

Lemma 5.3.9. *Suppose that $\mathcal{F}$ is a flag in a finite-dimensional vector space V. Then the Lie algebra $\mathfrak{n}_{\mathcal{F}} \subseteq \mathfrak{gl}(V)$ is nilpotent.*

Proof. For convenience, let us write $\mathfrak{n}$ for the Lie algebra $\mathfrak{n}_{\mathcal{F}}$. For each positive integer k, let $\mathfrak{n}^k$ be the subspace

$$\mathfrak{n}^k = \{x \in \mathfrak{gl}(V) : x(F_i) \subset F_{i-k}\}$$

(where we let $0 = F_l$ for all $l \leq 0$). Then clearly $\mathfrak{n}^k \subseteq \mathfrak{n}$, and $\mathfrak{n}^k = 0$ for any $k \geq n$. Thus if we can show $C^k(\mathfrak{n}) \subseteq \mathfrak{n}^{k+1}$ for each $k \geq 0$, it will follow that $\mathfrak{n}$ is nilpotent. The claim is immediate for $k = 0$, thus by induction we may assume that $C^k(\mathfrak{n}) \subseteq \mathfrak{n}^{k+1}$. But if $x \in \mathfrak{n}$ and $y \in \mathfrak{n}^{k+1}$, we have $xy(F_i) \subset x(F_{i-k-1}) \subset F_{i-k-2}$, and similarly $yx(F_i) \subset F_{i-k-2}$, thus certainly $[x,y] \in \mathfrak{n}^{k+2}$ and so $C^{k+1}(\mathfrak{n}) \subseteq \mathfrak{n}^{k+2}$ as required. In fact you can check that $C^k(\mathfrak{n}) = \mathfrak{n}^{k+1}$, so that $\mathfrak{n}$ is $(n-1)$-step nilpotent *i.e.* $C^{n-2}(\mathfrak{n}) \neq 0$, and $C^{n-1}(\mathfrak{n}) = 0$. □

Example 5.3.10. When $\mathcal{F}$ is a complete flag, so that $\dim(V) = n$, if we pick a basis $\{e_1, e_2, \dots, e_n\}$ of V such that $F_k = \mathrm{span}(e_1, e_2, \dots, e_k\}$, then the matrix A representing an element $x \in \mathfrak{n} = \mathfrak{n}_{\mathcal{F}}$ with respect to this basis is strictly upper triangular, that is, $a_{ij} = 0$ for all $i \geq j$. It follows that $\dim(\mathfrak{n}) = \binom{n}{2}$. When $n = 2$ we just get the 1-dimensional Lie algebra $\mathfrak{gl}_1$, thus the first nontrivial case is when $n = 3$. In this case $\mathfrak{n}$ is a 3-dimensional 2-step nilpotent Lie algebra.

If we pick a basis $\{e_1, e_2, \dots, e_n\}$ of V such that $F_i = \mathrm{span}(e_1, \dots, e_i)$, then $\mathfrak{gl}(V)$ gets identified with $\mathfrak{gl}_n$ and $\mathfrak{b}_{\mathcal{F}}$ corresponds to the subalgebra $\mathfrak{b}_n$ of upper triangular matrices. It is straight-forward to show by considering the subalgebra $\mathfrak{t}_n$ of diagonal matrices that $\mathfrak{b}_n$ is not nilpotent.

Remark 5.3.11. Notice that in the previous example, unlike in Lemma 5.3.9, it is essential that $\mathcal{F}$ is a complete flag. If $\mathcal{F}$ is not a complete flag the corresponding subalgebra $\mathfrak{b}_{\mathcal{F}}$ will not be solvable.

Remark 5.3.12. Note that the subalgebra $\mathfrak{t} \subset \mathfrak{gl}_n$ of diagonal matrices is nilpotent, since it is abelian, but the only nilpotent endomorphism of V is $\mathfrak{t}$ is 0. Thus a nilpotent linear Lie algebra need not consist of nilpotent endomorphisms. It turns out that, in some sense, the example of $\mathfrak{t}$ is the *only* way in which a nilpotent Lie algebra $\mathfrak{n} \subseteq \mathfrak{gl}(V)$ can fail to consist of nilpotent endomorphisms. We will make this precise in 5.3.2.

5.3.1 Nilpotent representations

Definition 5.3.13. Let $\mathfrak{g}$ be a Lie algebra and (V, ρ) a representation of $\mathfrak{g}$. We say that (V, ρ) is *nilpotent* if, for all $x \in \mathfrak{g}$, the endomorphism $\rho(x) \in \mathfrak{gl}(V)$ is a nilpotent linear map (that is, for some $n \geq 1$, $\rho(x)^n = 0$).

Lemma 5.3.14. *Let A be an associative algebra, and suppose $a, b \in A$ are nilpotent i.e. for some $n > 0$, we have $a^n = b^n = 0$. Then if a and b commute, $a + b$ is also nilpotent.*

Proof. This follows from the binomial theorem: Indeed we have

$$(a+b)^N = \sum_{k=0}^{N} \binom{N}{k} a^k b^{n-k}.$$

But now if $N \geq 2n$, then we must have either $k \geq n$ or $N - k \geq n$, hence in either case, each of the terms on the left-hand side vanishes, hence so does the right-hand side, and hence $a + b$ is nilpotent as required. □

Lemma 5.3.15. *Suppose $\mathfrak{g}$ is a Lie algebra and (V, ρ) and (W, σ) are representation of $\mathfrak{g}$.*

1. *If $x \in \mathfrak{g}$ is such that both $\rho(x)$ and $\sigma(x)$ are nilpotent, then the action of x on $V \otimes W$ is also nilpotent. Moreover, the action of x on V^* is also nilpotent. Thus if V and W are nilpotent, so are V^*, $V \otimes W$ and $Hom(V, W) \cong V^* \otimes W$.*

2. *If V is nilpotent, then any subrepresentation and any quotient representation of V is also nilpotent.*

Proof. By definition, x on $V \otimes W$ is by the endomorphsim $\rho(x) \otimes 1_W + 1_V \otimes \sigma(x)$. Since the two terms in this sum commute, the claim follows from Lemma 5.3.14 (taking $A = \mathrm{End}(V \otimes W)$.)

To see that x acts nilpotently on V^*, note that if $f \in V^*$, then

$$x^n(f)(v) = (-1)^n f(\rho(x)^n)(v)) = \pm f(0) = 0, \quad \forall v \in V, f \in V^*.$$

Part (2) follows from the fact that if $U \leq V$ and we write $\mathrm{End}_U(V) = \{\alpha \in \mathrm{End}(V) : \alpha(U) \subseteq U\}$ then the restriction map $r\colon \mathrm{End}_U(V) \to \mathrm{End}(U)$ and the quotient map $q\colon \mathrm{End}_U(V) \to \mathrm{End}(V/U)$ are both compatible with composition. □

The next proposition is the key result in this section. For the proof we will need the notion of the *normalizer* $N_{\mathfrak{g}}(\mathfrak{a})$ of a subalgebra $\mathfrak{a}$ of a Lie algebra $\mathfrak{g}$ given in Definition 2.2.6. We have

$$N_{\mathfrak{g}}(\mathfrak{a}) = \{x \in \mathfrak{g} : [x, a] \in \mathfrak{a}, \forall a \in \mathfrak{a}\},$$

so that $N_{\mathfrak{g}}(\mathfrak{h})$ is the largest subalgebra of $\mathfrak{g}$ in which $\mathfrak{a}$ is an ideal.

Proposition 5.3.16. *Let $\mathfrak{g}$ be a Lie algebra, and let (V, ρ) be a nilpotent representation of $\mathfrak{g}$.*

1. *The invariant subspace*
$$V^{\mathfrak{g}} = \{v \in V : \rho(x)(v) = 0, \forall x \in \mathfrak{g}\}$$
is non-zero.

2. *There is a complete flag $\mathcal{F}$ in V such that $\mathfrak{g} \subseteq \mathfrak{n}_{\mathcal{F}}$. In particular, the image $\rho(\mathfrak{g})$ is a nilpotent Lie algebra.*

Proof. We use induction on $d = \dim(\mathfrak{g})$, the case $d = 1$ being clear. Now if ρ is not faithful, *i.e.* $\ker(\rho) \neq 0$, then $\dim(\rho(\mathfrak{g})) < \dim(\mathfrak{g})$, and we are done by induction applied to the image $\rho(\mathfrak{g})$, hence we may assume ρ gives an embedding of $\mathfrak{g}$ into $\mathfrak{gl}(V)$ as a subalgebra, and we may thus identify $\mathfrak{g}$ with its image in the rest of this proof.

Now let $\mathcal{S} = \{\mathfrak{b} \subsetneq \mathfrak{g} : \mathfrak{b} \text{ is a proper subalgebra of } \mathfrak{g}\}$ denote the set of proper subalgebras of $\mathfrak{g}$, and pick $\mathfrak{a} \in \mathcal{S}$. Now by Lemma 5.3.15, $\mathfrak{a} \subsetneq \mathfrak{g} \subseteq \mathfrak{gl}(V) = V^* \otimes V$ are all nilpotent representations of $\mathfrak{a}$, since the restriction of (V, ρ) to $\mathfrak{a}$ is. But then, by the same Lemma, $\mathfrak{g}/\mathfrak{a}$ is also a nilpotent representation, and since $\dim(\mathfrak{a}) < \dim(\mathfrak{g})$, it follows by induction that the $\mathfrak{a}$-invariants $(\mathfrak{g}/\mathfrak{a})^{\mathfrak{a}}$ form a non-zero subrepresentation. Let $x \in \mathfrak{g}$ be such that $0 \neq x + \mathfrak{a} \in (\mathfrak{g}/\mathfrak{a})^{\mathfrak{a}}$. Then $\mathrm{ad}(a)(x) \in \mathfrak{a}$ for all $a \in \mathfrak{a}$, or equivalently, since $\mathrm{ad}(a)(x) = -\mathrm{ad}(x)(a)$, for all $a \in \mathfrak{a}$, we have $\mathrm{ad}(x)(a) \in \mathfrak{a}$, that is, $x \in N_{\mathfrak{g}}(\mathfrak{a})$. Thus the normalizer of $\mathfrak{a}$ is a subalgebra of $\mathfrak{g}$ which is strictly larger than $\mathfrak{a}$.

Thus if we take $\mathfrak{a} \in \mathcal{S}$ of maximal dimension, we must have $N_{\mathfrak{g}}(\mathfrak{a}) = \mathfrak{g}$, that is $\mathfrak{a}$ is an ideal in $\mathfrak{g}$. But then if $z \in \mathfrak{g} \backslash \mathfrak{a}$, it is easy to see that $\mathsf{k}.z \oplus \mathfrak{a}$ is a subalgebra[7] of $\mathfrak{g}$, hence again by maximality, we must have $\mathfrak{g} = \mathsf{k}.z \oplus \mathfrak{a}$. By induction, we know that $V^{\mathfrak{a}} = \{v \in V : a(v) = 0, \forall a \in \mathfrak{a}\}$ is a nonzero subspace of V. We claim that z preserves $V^{\mathfrak{a}}$. Indeed

$$a(z(v)) = [a, z](v) + z(a(v)) = 0, \quad \forall a \in \mathfrak{a}, v \in V^{\mathfrak{a}},$$

since $[a, z] \in \mathfrak{a}$. But the restriction of z to $V^{\mathfrak{a}}$ is nilpotent, so the subspace $U = \{v \in V^{\mathfrak{a}} : z(v) = 0\}$ is nonzero. Since $U = V^{\mathfrak{g}}$ we are done.

For the second part, let $C = (0 < F_1 < \ldots < F_m = V)$ be a composition series for V. It suffices to show that each of the composition factors are trivial. But if $1 \leq k \leq m$, then F_k is a subrepresentation of V and hence it is nilpotent. Similarly $Q_k = F_k/F_{k-1}$, as a quotient of F_k must be nilpotent. But then by part (1), its invariants $Q_k^{\mathfrak{g}}$ are a non-zero subrepresentation of Q_k, and since Q_k is simple it follows that Q_k is the trivial representation as required.

□

Corollary 5.3.17. *(Engel's theorem.) A Lie algebra $\mathfrak{g}$ is nilpotent if and only if $ad(x)$ is nilpotent for every $x \in \mathfrak{g}$, i.e the adjoint representation is nilpotent.*

[7] One way to see this is to note that $\mathsf{k}.z \oplus \mathfrak{a}$ is a line– *i.e.* one-dimensional subspace– of $\mathfrak{g}/\mathfrak{a}$ and any such subspace is a subalgebra, because, by the alternating property, the Lie bracket vanishes on lines. Note in particular that the direct sum is one of vector spaces, not Lie algebras.

Proof. First note that $\mathfrak{g}^{\mathfrak{g}}$, the invariants of $\mathfrak{g}$ in its adjoint representation, is precisely the centre $\mathfrak{z}(\mathfrak{g})$ of $\mathfrak{g}$, and if $\mathfrak{g}$ is nilpotent, $\mathfrak{z}(\mathfrak{g})$ is nontrivial whenever $\mathfrak{g}$ is. In other words, if we let $\mathfrak{z}_1 = \mathfrak{z}(\mathfrak{g})$ and inductively let $\mathfrak{z}_k$ be the preimage in $\mathfrak{g}$ of $\mathfrak{z}(\mathfrak{g}/\mathfrak{z}_{k-1})$, we obtain a strictly increasing sequence of subalgebras of $\mathfrak{g}$ for which $\mathrm{ad}(x)(\mathfrak{z}_k) \subseteq \mathfrak{z}_{k-1}$. Since $\dim(\mathfrak{g}) < \infty$, this sequence terminates with $\mathfrak{z}_n = \mathfrak{g}$ for some $n \geq 1$, and hence it follows that $\mathrm{ad}(x)$ acts nilpotently on $\mathfrak{g}$ as required.

To show the converse, first note that $\mathfrak{g}$ is evidently a central extension of $\mathfrak{g}/\mathfrak{z}(\mathfrak{g})$, and $\mathfrak{z}(\mathfrak{g}) = \ker(\mathrm{ad})$ is the kernel of the adjoint representation, thus it suffices to show that the image of $\mathfrak{g}$ in $\mathfrak{gl}(\mathfrak{g})$ is nilpotent. But by assumption, $\mathrm{ad}(x)$ is nilpotent for every $x \in \mathfrak{g}$, hence by part 2) of Theorem 5.3.16, there is a flag $\mathcal{F}$ in $\mathfrak{g}$ such that $\mathrm{ad}(\mathfrak{g}) \leq \mathfrak{n}_{\mathcal{F}}$. But recalling that $\mathfrak{z}(\mathfrak{g}) = \mathfrak{g}^{\mathfrak{g}}$ as above, it then follows that $\mathfrak{g}$ is nilpotent as required. □

5.3.2 Representations of nilpotent Lie algebras *In this section we assume that* k *is an algebraically closed field of characteristic zero.*

Definition 5.3.18. Let $\mathfrak{g}$ be a Lie algebra and let $\mathcal{S}$ be a set of irreducible representation of $\mathfrak{g}$. Let

$$\mathrm{Rep}_{\mathcal{S}}(\mathfrak{g}) = \{V \in \mathrm{Rep}\mathfrak{g} : [T : V] > 0 \text{ if and only if } \exists S \in \mathcal{S}, T \cong S\}$$
$$\mathrm{Rep}_{\mathcal{S}}(\mathfrak{g}, V) = \{W \leq V : W \in \mathrm{Rep}_{\mathcal{S}}(\mathfrak{g})\}.$$

If $\mathcal{S} = \{S\}$ then we will write $\mathrm{Rep}_S(\mathfrak{g}), \mathrm{Rep}_S(\mathfrak{g}, V)$ rather than $\mathrm{Rep}_{\{S\}}(\mathfrak{g}), \mathrm{Rep}_{\{S\}}(\mathfrak{g}, V)$ respectively.

Proposition 5.3.19. *Let* $\mathfrak{g}$ *be a Lie algebra and* (V, ρ) *a representation of* $\mathfrak{g}$*. If* $\mathcal{S}$ *is a set of irreducible* $\mathfrak{g}$*-representation then* $\mathrm{Rep}_{\mathcal{S}}(\mathfrak{g}, V)$ *has a unique element* $V_{\mathcal{S}}$ *which is maximal with respect to containment, that is* $V_{\mathcal{S}} \in \mathrm{Rep}_{\mathcal{S}}(\mathfrak{g}, V)$ *and if* $U \in \mathrm{Rep}_{\mathcal{S}}(\mathfrak{g}, V)$ *then* $U \leq V_{\mathcal{S}}$.

Proof. First note that if it exists, such a maximal element is automatically unique, since if W_1, W_2 are both maximal with respect to containment we must have $W_1 \leq W_2 \leq W_1$ and hence $W_1 = W_2$.

Next note that if $V_1, V_2 \in \mathcal{U}_S$ then $V_1 + V_2 \in \mathcal{U}_S$. Indeed by the second isomorphism theorem, $(V_1 + V_2)/V_1 \cong V_2/(V_1 \cap V_2)$, so that any composition factor of $V_1 + V_2$ must be a composition factor of V_1 or of $V_2/(V_1 \cap V_2)$, and hence is a composition factor of V_1 or V_2. Now pick $W \in \mathcal{U}_S$ with $\dim(W) \geq \dim(U)$ for all $U \in \mathcal{U}_{[S]}$ (such a W exists if V is finite-dimensional, as we always assume). We claim that W is maximal for containment. Indeed if $U \in \mathcal{U}_S$ then we have just shown that $W + U \in \mathcal{U}_S$, hence $\dim(W) \leq \dim(W + U) \leq \dim(W)$ by our choice of W, and hence $U \leq W$ and W is maximal for containment as required. Thus $W = V_{\mathcal{S}}$ is the unique maximal subrepresentation in $\mathcal{U}_{\mathcal{S}}$. □

Definition 5.3.20. Recall that the isomorphism classes of 1-dimensional representations of $\mathfrak{g}$ can be identified with $D(\mathfrak{g})^0 \subseteq \mathfrak{g}$, and given $\lambda \in D(\mathfrak{g})^0$, we write k_λ for the 1-dimensional representation (k, λ). Given a $\mathfrak{g}$-representation (V, ρ), we will write V_λ and $\mathrm{Rep}_\lambda(\mathfrak{g}, V)$ instead of V_{k_λ} and $\mathrm{Rep}_{\mathsf{k}_\lambda}(\mathfrak{g}, V)$. When $\lambda \in D(\mathfrak{g})^0$ we will refer to V_λ as the *λ-generalised weight space* of V.[8] If V is a finite-dimensional representation of a Lie algebra $\mathfrak{g}$, let

$$\Psi_V = \{\lambda \in D(\mathfrak{g})^0 : \lambda \text{ is a composition factor of } V\}$$

Thus Ψ_V is the finite set of the one-dimensional representations of V which occur as composition factors of V. If $\mathfrak{g}$ is solvable and $\mathrm{char}(\mathsf{k}) = 0$ then by Lie's Theorem Ψ_V contains all the composition factors of V.

If $\varphi \colon \mathfrak{g}_1 \to \mathfrak{g}_2$ and (V, ρ) is a representation of $\mathfrak{g}_2$, then $(V, \varphi^*(\rho))$ is a representation of $\mathfrak{g}_1$, where $\varphi^*(\rho) = \rho \circ \varphi$. Since $\varphi(D(\mathfrak{g}_1)) \subseteq D(\mathfrak{g}_2)$, the transpose $\varphi^\top : \mathfrak{g}_2^* \to \mathfrak{g}_1^*$ restricts to give a map $\varphi^\top : D(\mathfrak{g}_2)^0 \to D(\mathfrak{g}_1)^0$, and $\Psi_{\varphi^*(V)} = \varphi^\top(\Psi_V)$. Now if $x \in \mathfrak{g}$ and $i_x \colon \mathfrak{gl}1 \to \mathfrak{g}$ is the homomorphism $i_x(t) = t.x$ ($\forall t \in \mathsf{k} = \mathfrak{gl}1$), and $\lambda \in D(\mathfrak{g})^0$ then $i_x^\top(\lambda) = \lambda(x)$. The weights of the $\mathfrak{gl}1$-representation $\rho \circ i_x$ are just the eigenvalues of $\rho(x)$, as in Example 3.1.3, it follows that the eigenvalues of $\rho(x)$ are $\{\lambda(x) : \lambda \in \Psi_V\}$, and the μ-generalised eigenspace of $\rho(x)$ is $\bigoplus_{\lambda \in \Psi_V : \lambda(x) = \mu} V_\lambda$.

[8] In lectures these were sometimes referred to as ”weight spaces”, whereas it is more standard to reserve the term “weight space” for a direct sum of copies of k_λ for some $\lambda \in D(\mathfrak{g})^0$.

Lemma 5.3.21. *Suppose that $\mathfrak{g}$ is a Lie algebra and $\lambda, \mu \in D(\mathfrak{g})^0$ are weights of $\mathfrak{g}$. If V and W are finite-dimensional representations of $\mathfrak{g}$ then*

i) $V_\lambda \otimes W_\mu \subseteq (V \otimes W)_{\lambda+\mu}$.

ii) If $\phi \colon V \to W$ is a homomorphism of $\mathfrak{g}$-representation, then $\phi(V_\lambda) \subseteq W_\lambda$.

Proof. For part (i), we may assume that $V = V_\lambda$ and $W = W_\mu$, hence there are composition series $(F_k)_{k=0}^r$ and $(G_l)_{l=0}^s$, where $F_k/F_{k+1} \cong \mathsf{k}_\lambda$ for each k, and $G_l/G_{l+1} \cong \mathsf{k}_\mu$, for all $l \in \{0, 1, \dots, r\}$ and $k \in \{0, 1 \dots, s\}$. Pick bases $\{e_i : 0 \leq i \leq r-1\}$ and $\{f_j : 0 \leq j \leq s-1\}$ of V and W respectively such that $F_k = \langle \{e_i : i \geq k\} \rangle_{\mathsf{k}}$ and $G_l = \langle \{f_j : j \geq l\} \rangle_{\mathsf{k}}$. If we set $H_k = \sum_{r+s=k} F_r \otimes G_s$, then H_k is a subrepresentation of $V \otimes W$ and we have $x(e_k) \otimes f_l = \lambda(x) e_k \otimes f_l + F_{k+1} \otimes G_l$ and $e_k \otimes x(f_l) = \mu(x).e_k \otimes f_l + F_k \otimes G_{l+1}$ hence

$$x(e_k \otimes f_l) = x(e_k) \otimes f_l + e_k \otimes x(f_l) \in (\lambda + \mu)(e_k \otimes f_l) + H_{k+l+1} \tag{5.1}$$

and thus $H_k/H_{k+1} \cong \mathsf{k}_{\lambda+\mu}^{\dim(H_k)-\dim(H_{k+1})}$. It follows $V \otimes W$ has $\mathsf{k}_{\lambda+\mu}$ as its unique composition factor.

For part (ii), since $V_\lambda \in \mathrm{Rep}_\lambda(\mathfrak{g})$, and $\phi(V_\lambda) \cong V_\lambda/\ker(\phi_{|V_\lambda})$ is isomorphic to a quotient of V_λ, it lies in $\mathrm{Rep}_\lambda(\mathfrak{g}, W)$ and so by the maximality of W_λ it follows that $\phi(V_\lambda) \subseteq W_\lambda$. □

The adjoint representation of a nilpotent Lie algebra $\mathfrak{g}$ has the trivial representation as its only composition factor, that is, $\mathfrak{g} = \mathfrak{g}_0$. This has the following important consequence:

Proposition 5.3.22. *Let $\mathfrak{g}$ be a nilpotent Lie algebra, $\mathfrak{h} \subseteq \mathfrak{g}$ be a subalgebra of $\mathfrak{g}$, and (V, ρ) a representation of $\mathfrak{g}$. Then if $\mu \in (\mathfrak{h}/D(\mathfrak{h}))^* \cong D(\mathfrak{h})^0/\mathfrak{h}^0 \subseteq \mathfrak{g}^*/\mathfrak{h}^0 \cong \mathfrak{h}^*$ is a weight of $\mathfrak{h}$, and V_μ is the μ-isotypic subrepresentation of $Res_{\mathfrak{h}}^{\mathfrak{g}}(V)$, the restriction of V to $\mathfrak{h}$, then V_μ is a $\mathfrak{g}$-subrepresentation of V. In particular, taking $\mathfrak{h} = \mathsf{k}.x$ for $x \in \mathfrak{g}\backslash\{0\}$, any generalised eigenspace $V_{\mu,x}$ of $\rho(x)$ is a $\mathfrak{g}$-subrepresentation.*

Proof. Since $\mathfrak{g}$ is nilpotent, we have $\mathfrak{g} = \mathfrak{g}_0$ as an $\mathfrak{h}$-representation. But then by Lemma 5.3.21, we have $\mathfrak{g} \otimes V_\mu = \mathfrak{g}_0 \otimes V_\mu \subseteq (\mathfrak{g} \otimes V)_\mu$, and since the map $\tilde{a}_\rho \colon \mathfrak{g} \otimes V \to V$ given by $\tilde{a}_\rho(x \otimes v) = \rho(x)(v)$ is a homomorphism of $\mathfrak{h}$-representations by Example 3.3.8, it follows that $\tilde{a}_\rho(\mathfrak{g} \otimes V_\mu) = \rho(\mathfrak{g})(V_\mu) \subseteq V_\mu$, that is, V_μ is a $\mathfrak{g}$-subrepresentation as required. □

Definition 5.3.23. Let $\mathfrak{g}$ be a nilpotent Lie algebra and let (V, ρ) be a representation of $\mathfrak{g}$. Say $x \in \mathfrak{g}$ is V*-generic* if, for all $\lambda, \mu \in \Psi_V$ we have $\lambda(x) = \mu(x)$ if and only if $\lambda = \mu$.

If $D_V = \{\lambda - \mu : \lambda, \mu \in \Psi_V\}\backslash\{0\}$, then x is V-generic if and only if $x \notin \bigcup_{\nu \in D_V} \ker(\nu)$. If k is infinite,[9] it is an elementary exercise to show that a nonzero k-vector space cannot be written as the union of finitely many hyperplanes, hence V-generic elements of $\mathfrak{g}$ exist for any finite-dimensional $\mathfrak{g}$-representation V.

Theorem 5.3.24. *Let $\mathfrak{g}$ be a nilpotent Lie algebra and (V, ρ) a finite-dimensional representation of $\mathfrak{g}$. For each $\lambda \in (\mathfrak{g}/D\mathfrak{g})^*$, let*

$$W_\lambda = \bigcap_{x \in \mathfrak{g}} V_{\lambda(x),x}, \quad V_{\lambda(x),x} = \{v \in V : \exists n > 0 \textit{ such that } (\rho(x) - \lambda(x))^n(v) = 0\}.$$

If $x_0 \in \mathfrak{g}$ is V-generic, then we have $V_{\lambda(x_0),x_0} = V_\lambda = W_\lambda$ and hence $V = \bigoplus_\lambda V_\lambda$ is the direct sum of its (generalised) weight spaces.

Proof. Since $\mathfrak{g}$ is nilpotent, it is solvable, hence for any $\mathfrak{g}$-representation (U, σ) its composition factors all lie in Ψ_U and, as in Definition 5.3.20, if $x \in \mathfrak{g}$ then $\sigma(x)$ has spectrum $\{\lambda(x) : \lambda \in \Psi_U\}$. In particular, taking $U = V_\lambda$ we see that $\rho(x)_{|V_\lambda}$ has $\lambda(x)$ as its sole eigenvalue, that is, $V_\lambda \subseteq V_{\lambda(x),x}$. It follows that $V_\lambda \subseteq W_\lambda$.

[9]Any field k with $\mathrm{char}(\mathsf{k}) = 0$ contains a copy of $\mathbb{Q}$ and so is infinite. Alteratively, any algebraically closed field is infinite – *e.g.* take the n-th roots of some $\mu \in \mathsf{k}^\times$ where n is taken coprime to $\mathrm{char}(\mathsf{k})$.

Now if $x \in \mathfrak{g}$, we have $V = \bigoplus_{\lambda(x):\lambda \in \Psi_V} V_{\lambda(x),x}$. Moreover each summand $V_{\lambda(x),x}$ is a $\mathfrak{g}$-representation by Proposition 5.3.22, hence taking $U = V_{\lambda(x),x}$ we see that if k_ν is a composition factor, then $\nu(x) = \lambda(x)$. It follows that if we take x_0 to be V-generic, the generalised eigenspace $V_{\lambda(x_0),x_0}$ has λ as its unique composition factor, so that $V_{\lambda(x_0),x_0} \subseteq V_\lambda$. Hence $V_{\lambda(x_0),x_0} = V_\lambda = W_\lambda$ and $V = \bigoplus_{\lambda \in \Psi_V} V_\lambda$. □

5.3.3 Nilpotent Lie algebras as measurements: Cartan subalgebras

5.3.3.1 Cartan Subalgebras *In this section we work over an algebraically closed field* k*. In particular,* k *is infinite.*

Let $\mathfrak{g}$ be a Lie algebra. Recall that if $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$ then the normalizer $N_\mathfrak{g}(\mathfrak{h})$ of $\mathfrak{g}$ is

$$N_\mathfrak{g}(\mathfrak{h}) = \{x \in \mathfrak{g} : [x,h] \in \mathfrak{h}, \forall h \in \mathfrak{h}\}.$$

It follows immediately from the Jacobi identity that $N_\mathfrak{g}(\mathfrak{h})$ is a subalgebra, and clearly $N_\mathfrak{g}(\mathfrak{h})$ is the largest subalgebra of $\mathfrak{g}$ in which $\mathfrak{h}$ is an ideal.

Definition 5.3.25. A subalgebra $\mathfrak{h}$ is said to be a *Cartan subalgebra* if it is nilpotent and self-normalizing, that is, $N_\mathfrak{g}(\mathfrak{h}) = \mathfrak{h}$.

It is not clear from this definition whether a Lie algebra necessarily has a Cartan subalgebra. To show this, we need a few more definitions.

Definition 5.3.26. If $x \in \mathfrak{g}$, let $\mathfrak{g}_{0,x}$ be the generalized 0-eigenspace of $\operatorname{ad}(x)$, that is

$$\mathfrak{g}_{0,x} = \{y \in \mathfrak{g} : \exists N > 0 \text{ such that } \operatorname{ad}(x)^N(y) = 0\}$$

Note that we always have $x \in \mathfrak{g}_{0,x}$. We say that $x \in \mathfrak{g}$ is *regular* if $\mathfrak{g}_{0,x}$ is of minimal dimension.

Proposition 5.3.27. *1. If $x \in \mathfrak{g}$ is any element, then $\mathfrak{g}_{0,x}$ is a self-normalizing subalgebra of $\mathfrak{g}$.*

2. If $x \in \mathfrak{g}$ is a regular element, then $\mathfrak{g}_{0,x}$ is a nilpotent and so a Cartan subalgebra of $\mathfrak{g}$.

Proof. Part (1) is straight-forward: It follows immediately from Lemma 5.3.22 applied to the adjoint representation that $\mathfrak{h} = \mathfrak{g}_{0,x}$ is a subalgebra of $\mathfrak{g}$. To see that $\mathfrak{h}$ is a self-normalizing in $\mathfrak{g}$. Indeed if $z \in N_\mathfrak{g}(\mathfrak{h})$ then $[x,z] \in \mathfrak{h}$ (since certainly $x \in \mathfrak{h}$), so that for some n we have $\operatorname{ad}(x)^n([x,z]) = 0$, and hence $\operatorname{ad}(x)^{n+1}(z) = 0$ and $z \in \mathfrak{h}$ as required.

To establish part (2), assume that x is regular, and let $\mathfrak{h} = \mathfrak{g}_{0,x}$. To see that $\mathfrak{h}$ is nilpotent, by Engel's theorem it suffices to show that, for each $y \in \mathfrak{h}$, the map $\operatorname{ad}(y)$ is nilpotent as an endomorphism of $\mathfrak{h}$. To see this, we consider the characteristic polynomials of $\operatorname{ad}(y)$ on $\mathfrak{g}, \mathfrak{h}$ and $\mathfrak{g}/\mathfrak{h}$: Since $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, the characteristic polynomial $\chi^y(t) \in \mathsf{k}[t]$ of $\operatorname{ad}(y)$ on $\mathfrak{g}$ is the product of the characteristic polynomials of $\operatorname{ad}(y)$ on $\mathfrak{h}$ and $\mathfrak{g}/\mathfrak{h}$, which we will write as $\chi_1^y(t)$ and $\chi_2^y(t)$ respectively.

We may write $\chi^y(t) = \sum_{k=0}^n c_k(y)t^k$, where $n = \dim(\mathfrak{g})$. Pick $\{h_1, h_2, \dots, h_r\}$ a basis of $\mathfrak{h}$ (so that $\dim(\mathfrak{h}) = r$). Then if we write $y = \sum_{i=1}^r y_i h_i$, the coefficients $\{c_k(y)\}_{k=0}^n$ of $\chi^y(t)$ are polynomial functions of the coordinates $\{y_i : 1 \leq i \leq r\}$. Similarly we have

$$\chi_1^y(t) = \sum_{i=0}^r d_i(y)t^i, \quad \chi_2^y(t) = \sum_{j=0}^{n-r} e_j(y)t^j$$

where the $d_i, e_j \in \mathsf{k}[x_1, \dots, x_n]$ are polynomials and $d_i(y) = d_i(y_1, \dots, y_n)$ where $y = \sum_{i=1}^n y_i h_i$. Since $\operatorname{ad}(x)(x) = 0$, we have $x \in \mathfrak{g}_{0,x}$. But $\operatorname{ad}(x)$ is invertible on $\mathfrak{g}/\mathfrak{h}$, since all its eigenvalues are non-zero on $\mathfrak{g}/\mathfrak{h}$, hence $\chi_2^x(t)$ has $e_0(x) \neq 0$, and thus the polynomial e_0 is nonzero.

Now let $s = \min\{i : d_i(x_1, \dots, x_n) \neq 0\}$. Then we may write $\chi_1^y(t) = t^s \sum_{k=0}^{r-s} d_{s+k}(y)t^k$, and hence

$$\chi^y(t) = t^s(d_s + d_{s+1}t + \dots)(e_0 + e_1.t + \dots) = t^s d_s e_0 + \dots,$$

For any endomorphism of a vector space, the dimension of its λ-generalised eigenspace is the largest power of $(t-\lambda)$ dividing its characteristic polynomial. In particular this implies that, for any $y \in \mathfrak{h}$, we have $\dim(\mathfrak{g}_{0,y}) = \min\{i : c_i(y) \neq 0\}$. But since $e_0.d_s \in \mathsf{k}[x_1, \dots, x_n]$ is nonzero, there is some $z \in \mathfrak{h}$ such that $d_s(z).e_0(z) \neq 0$, and hence $\dim(\mathfrak{g}_{0,z}) = s$. Now by definition $s \leq r = \dim(\mathfrak{g}_{0,x})$, hence since x is regular, we must have $s = r$, and hence $\chi_1^y(t) = t^r$, for all $y \in \mathfrak{h}$. Hence every $\mathrm{ad}(y)$ is nilpotent on $\mathfrak{h}$, so that $\mathfrak{h}$ is a Cartan subalgebra as required.

□

In the course of the proof of the above Proposition we used the fact that the coefficients of the characteristic polynomial were polynomial functions of the coordinates of $y \in \mathfrak{h}$ with respect to a basis of $\mathfrak{h}$. This was crucial because, whereas the product of two arbitrary nonzero functions may well be zero, the product of two nonzero *polynomials* (over a field) is never zero. For completeness we give a proof[10]. (To apply it to the above, take $V = \mathfrak{g}$, $A = \mathfrak{h}$ and $\varphi = \mathrm{ad}$).

Lemma 5.3.28. *Suppose that V and A are finite dimensional vector spaces, $\varphi \colon A \to End(V)$ is a linear map, and $\{a_1, a_2, \dots, a_k\}$ is a basis of A. Let*

$$\chi_a(t) = \sum_{i=0}^{d} c_i(a)t^i \in \mathsf{k}[t]$$

be the characteristic polynomial of $\varphi(a) \in A$. Then if we write $a = \sum_{i=1}^k x_i a_i$, the coefficients $c_i(a)$ ($1 \leq i \leq d$) are polynomials in $\mathsf{k}[x_1, x_2, \dots, x_k]$.

Proof. Pick a basis of V so that we may identify $\mathrm{End}(V)$ with $\mathrm{Mat}_n(\mathsf{k})$ the space of $n \times n$ matrices. Then each $\varphi(a_i)$ is a matrix $(a_i^{jk})_{1\leq j,k\leq n}$, and if $a = \sum_{i=1}^k x_i a_i$, we have

$$\chi_a(t) = \det(tI_n - \sum_{i=1}^{k} x_i\varphi(a_i)),$$

which from the formula for the determinant clearly expands to give a polynomial in the x_i and t, which yields the result. □

Remark 5.3.29. As an aside, there's no reason one needs to pick a basis of a vector space V in order to talk about the space $\mathsf{k}[V]$ of k-valued polynomial functions on it. For example, one can define $\mathsf{k}[V]$ to be the subalgebra of all k-valued functions on V which is generated by V^* the space of functionals on V. (This is fine if k is algebraically closed at least, if that is not the case then one should be a bit more careful, e.g. recall if k is finite, then an element of $\mathsf{k}[t]$ is *not* a function on k).

Remark 5.3.30. Although we will not prove it in this course, any two Cartan subalgebras of $\mathfrak{g}$ are conjugate by an automorphism[11] of $\mathfrak{g}$, that is, given any two Cartan subalgebras $\mathfrak{h}_1, \mathfrak{h}_2$ there is an isomorphism $\alpha \colon \mathfrak{g} \to \mathfrak{g}$ such that $\alpha(\mathfrak{h}_1) = \mathfrak{h}_2$. One difficulty here is that, while the proof of Proposition 5.3.27 shows that, if $x \in \mathfrak{g}$ is regular, then $\mathfrak{g}_{x,0}$ is a Cartan subalgebra, it is not clear that any Cartan subalgebra is of that form.

5.3.3.2 The Cartan Decomposition

In this section we work over an algebraically closed field k of characteristic zero.

Our study of the representation theory of nilpotent Lie algebras can now be used to study the structure of an arbitrary Lie algebra. Indeed, if $\mathfrak{g}$ is any Lie algebra, we have shown that it contains a Cartan subalgebra $\mathfrak{h}$, and the restriction of the adjoint action makes $\mathfrak{g}$ into an $\mathfrak{h}$-representation. As such it decomposes into a direct sum

$$\mathfrak{g} = \bigoplus_{\lambda \in (\mathfrak{h}/D\mathfrak{h})^*} \mathfrak{g}_\lambda.$$

[10] If this all seems overly pedantic then feel free to ignore it.

[11] In fact, they are even conjugate by what is known as an *inner automorphism*.

Let $\Phi = \{\lambda \in (\mathfrak{h}/D(\mathfrak{h}))^* : \mathfrak{g}_\lambda \neq 0\}$ be the set of roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$. The next Lemma establishes some basic properties of this decomposition.

Lemma 5.3.31. *Let $\mathfrak{g}, \mathfrak{h}$ be as above, and let V be a (finite-dimensional) $\mathfrak{g}$-representation. As an $\mathfrak{h}$-representation, it decomposes*

$$V = \bigoplus_{\mu \in \Psi} V_\mu$$

into generalised $\mathfrak{h}$-weight spaces as in Theorem 5.3.24, where $\Psi = \{\lambda \in (\mathfrak{g}/D(\mathfrak{g}))^* : V_\lambda \neq 0\}$ is a finite subset of $(\mathfrak{g}/D(\mathfrak{g}))^*$. Then for any $\alpha, \mu \in (\mathfrak{g}/D(\mathfrak{g}))^*$ we have

$$\rho(\mathfrak{g}_\alpha)(V_\mu) \subseteq V_{\alpha+\mu}.$$

In particular, taking $(V, \rho) = (\mathfrak{g}, ad)$ we see that, for any $\alpha, \beta \in (\mathfrak{g}/D(\mathfrak{g}))^*$ we have $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$.

Proof. By Lemma 5.3.21, the $\mathfrak{h}$-representation $\mathfrak{g}_\lambda \otimes V_\mu$ has $\mathsf{k}_{\lambda+\mu}$ as its only composition factor. As in Example 3.3.8, the map $\tilde{\rho}\colon \mathfrak{g} \otimes V \to V$ given by $\tilde{\rho}(y \otimes v) = \rho(y)(v)$ is a homomorphisms of $\mathfrak{g}$-representations, and hence of $\mathfrak{h}$-representation, thus its image span$\{\rho(h)(v) : h \in \mathfrak{h}, v \in V\}$ lies in $V_{\alpha+\mu}$. □

Definition 5.3.32. By the previous Lemma, if $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$ then $\mathfrak{g}$ decomposes into a direct sum

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\lambda \neq 0} \mathfrak{g}_\lambda.$$

This is known as the *Cartan decomposition* of $\mathfrak{g}$. The set Φ of non-zero $\lambda \in (\mathfrak{h}/D\mathfrak{h})^*$ for which the subspace $\mathfrak{g}_\lambda$ is non-zero is called the set of *roots* of $\mathfrak{g}$, and the subspaces $\mathfrak{g}_\lambda$ are known[12] as the *root spaces* of $\mathfrak{g}$. Thus finally the Cartan decomposition becomes

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\lambda \in \Phi} \mathfrak{g}_\lambda.$$

By Remark 5.3.30 above, the Cartan decomposition of $\mathfrak{g}$ is unique up to automorphism.

[12] *i.e.* in the terminology for representations of nilpotent Lie algebras discussed above, the roots of $\mathfrak{g}$ are the weights of $\mathfrak{g}$ as an $\mathfrak{h}$-representation.

Chapter 6

Trace forms and semisimple Lie algebras

6.1 Trace forms and the Killing form

In this section we introduce certain symmetric bilinear forms, which will play an important role in the rest of the course. A brief review of the basic theory of symmetric bilinear forms[1] is given in §I.3 in Appendix 1 of these notes.

6.1.1 Bilinear forms Let $\mathrm{Bil}(V)$ be the space of bilinear forms on V, that is,

$$\mathrm{Bil}(V) = \{B\colon V \times V \to \mathsf{k} : B \text{ bilinear}\}.$$

From the definition of tensor products it follows that $\mathrm{Bil}(V)$ can be identified with $(V \otimes V)^*$. We say that a bilinear form is symmetric if $B(v,w) = B(w,v)$. In terms of tensor products, notice that the map $\sigma\colon V \times V \to V \times V$ given by $(v,w) \mapsto (w,v)$ induces an involution (which we will also denote by σ on $V \otimes V$ and the condition that B is symmetric is just the condition that $B \circ \sigma = B$.

If V is a $\mathfrak{g}$-representation, this means $\mathrm{Bil}(V)$ also has the structure of $\mathfrak{g}$-representation: explicitly, if $B \in \mathrm{Bil}(V)$, then it yields a linear map $b\colon V \otimes V \to \mathsf{k}$ by the universal property of tensor products, and if $y \in \mathfrak{g}$, it acts on B as follows:

$$\begin{aligned} y(B)(v,w) &= y(b)(v \otimes w) \\ &= -b(y(v \otimes w)) \\ &= -b(y(v) \otimes w + v \otimes y(w)) \\ &= -B(y(v),w) - B(v,y(w)). \end{aligned}$$

That is, B is invariant if $B(y(v),w) = -B(v,y(w))$ for all $v,w \in V$ and $y \in \mathfrak{g}$. Notice that the involution $\sigma \in \mathrm{End}(V \otimes V)$ commutes with the action of $\mathfrak{g}$ (this is a special case of the fact that, for any two $\mathfrak{g}$-representations, the map $\tau\colon V \otimes W \to W \otimes V$ given by $\tau(v \otimes w) = w \otimes v$ is a $\mathfrak{g}$-homomorphism). It follows that the action of $\mathfrak{g}$ preserves the space $\mathrm{SBil}(V)$ of symmetric bilinear forms[2]

If we apply this to $(V,\rho) = (\mathfrak{g}, \mathrm{ad})$, then the condition that $B \in \mathrm{Bil}(\mathfrak{g})^{\mathfrak{g}}$ is just that, for all $x,y,z \in \mathfrak{g}$,

$$0 = y(B)(x,z) = -B(\mathrm{ad}(y)(x),z) - B(x,\mathrm{ad}(y)(z)) = B([x,y],z) - B(x,[y,z]),$$

that is, $B([x,y],z) = B(x,[y,z])$.

[1]Part A Algebra focused more on positive definite and Hermitian forms, but there is a perfectly good theory of general symmetric bilinear forms.

[2]In Part B Representation Theory it was shown that if G is a group and (V,σ) is a G-representation, then $V \otimes V = \mathrm{Sym}^2(V) \oplus \mathrm{Alt}^2(V)$, where the two summands are the $+1$ and -1 eigenspaces for the involution $\tau\colon V \otimes V \to V \otimes V$ given by $\tau(v_1 \otimes v_2) = v_2 \otimes v_1$ $(v_1, v_2 \in V)$. The analogous result also holds for representations of a Lie algebra.

Definition 6.1.1. We say that a bilinear form B is *invariant* if it is an invariant vector for the action of $\mathfrak{g}$ on $\mathrm{Bil}(\mathfrak{g}) \cong (\mathfrak{g} \otimes \mathfrak{g})^*$, that is, if

$$B([x,y],z) = B(x,[y,z]), \quad \forall x,y,z \in \mathfrak{g}.$$

If $\alpha\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ is a homomorphism of Lie algebras, and B is a bilinear form on $\mathfrak{g}_2$, then we may "pull-back" B using α to obtain a bilinear form on $\mathfrak{g}_1$. Indeed viewing B as an element of $(\mathfrak{g}_2 \otimes \mathfrak{g}_2)^*$, we obtain an element $\alpha^*(B)$ of $(\mathfrak{g}_1 \otimes \mathfrak{g}_1)^*$ given by $\alpha^*(B)(x,y) = B(\alpha(x),\alpha(y))$. It is immediate from the definitions that if B is an invariant form for $\mathfrak{g}_2$, then $\alpha^*(B)$ is an invariant form for $\mathfrak{g}_1$.

It follows that if we can find an invariant form b_V on a general linear Lie algebra $\mathfrak{gl}(V)$, then any representation $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ of a Lie algebra $\mathfrak{g}$ on V will yield an invariant bilinear form $t_V = \rho^*(b_V)$ on $\mathfrak{g}$. The next Lemma shows that there is in fact a very natural invariant bilinear form, indeed an invariant symmetric bilinear form, on a general linear Lie algebra $\mathfrak{gl}(V)$:

Lemma 6.1.2. *Let V be a* k*-vector space. The trace form* $b_V\colon \mathfrak{gl}(V) \otimes \mathfrak{gl}(V) \to \mathsf{k}$ *given by*

$$b_V(a,b) = tr(a.b), \quad \forall a,b \in \mathfrak{gl}(V),$$

is an invariant symmetric bilinear form on $\mathfrak{gl}(V)$.

Proof. The invariance of the form b_V is just the condition that the map $b_V\colon \mathfrak{gl}(V) \otimes \mathfrak{gl}(V) \to \mathsf{k}$ is a map of $\mathfrak{gl}(V)$-representations, where k is viewed as the trivial representation of $\mathfrak{gl}(V)$. In fact this follows from Example I.11. To see this we use the notation of Appendix I.2. Indeed, using the notation of Remark I.12, the bilinear map is just the composition

$$\mathfrak{gl}(V) \otimes \mathfrak{gl}(V) \longrightarrow V^* \otimes V \otimes V^* \otimes V \xrightarrow{\iota_{32}} V^* \otimes \mathsf{k} \otimes V \longrightarrow V^* \otimes V \xrightarrow{\iota} \mathsf{k}.$$

where the first map is $\theta^{-1} \otimes \theta^{-1}$, and the third map is induced by the scalar multiplication isomorphism $V \otimes \mathsf{k} \cong V$. Since this description of the trace form uses only the map c and identity maps, it is clearly a $\mathfrak{g}$-homomorphism, and hence an invariant bilinear form. □

Definition 6.1.3. If $\mathfrak{g}$ is a Lie algebra, and let (V,ρ) be a representation of $\mathfrak{g}$. we may define a bilinear form $t_V\colon \mathfrak{g} \times \mathfrak{g} \to \mathsf{k}$ on $\mathfrak{g}$, known as a trace form of the representation (V,ρ), to be $\rho^*(b_V)$. Explicitly, we have

$$t_V(x,y) = \mathrm{tr}_V(\rho(x)\rho(y)), \quad \forall x,y \in \mathfrak{g}.$$

Definition 6.1.4. The *Killing form* $\kappa\colon \mathfrak{g} \times \mathfrak{g} \to \mathsf{k}$ is the trace form given by the adjoint representation, that is:

$$\kappa(x,y) = \mathrm{tr}(\mathrm{ad}(x)\mathrm{ad}(y)).$$

Note that if $\mathfrak{a} \subseteq \mathfrak{g}$ is a subalgebra, the Killing form of $\mathfrak{a}$ is not necessarily equal to the restriction of that of $\mathfrak{g}$. We will write $\kappa^{\mathfrak{g}}$ when it is not clear from context which Lie algebra is concerned.

If $\mathfrak{a}$ is an ideal in $\mathfrak{g}$, then in fact the Killing form is unambiguous, as the following Lemma shows.

Lemma 6.1.5. *Let $\mathfrak{a}$ be an ideal of $\mathfrak{g}$. The Killing form $\kappa^{\mathfrak{a}}$ of $\mathfrak{a}$ is given by the restriction of the Killing form $\kappa^{\mathfrak{g}}$ on $\mathfrak{g}$, that is:*

$$\kappa^{\mathfrak{g}}_{|\mathfrak{a}} = \kappa^{\mathfrak{a}}.$$

Moreover, the subspace orthogonal to $\mathfrak{a}$, that is, $\mathfrak{a}^{\perp} = \{x \in \mathfrak{g} : \kappa(x,y) = 0 \forall y \in \mathfrak{a}\}$ is also an ideal.

Proof. If $a \in \mathfrak{a}$ we have $\mathrm{ad}(a)(\mathfrak{g}) \subseteq \mathfrak{a}$, thus the same will be true for the composition $\mathrm{ad}(a_1)\mathrm{ad}(a_2)$ for any $a_1,a_2 \in \mathfrak{a}$. Thus if we pick a vector space complement W to $\mathfrak{a}$ in $\mathfrak{g}$, the matrix of $\mathrm{ad}(a_1)\mathrm{ad}(a_2)$ with respect to a basis compatible with the subspaces $\mathfrak{a}$ and W will be of the form

$$\begin{pmatrix} A & B \\ 0 & 0. \end{pmatrix}$$

where $A \in \mathrm{End}(\mathfrak{a})$ and $B \in \mathrm{Hom}_{\mathsf{k}}(\mathfrak{a}, W)$. Then clearly $\mathrm{tr}(\mathrm{ad}(a_1)\mathrm{ad}(a_2)) = \mathrm{tr}(A)$. Since A is clearly given by $\mathrm{ad}(a_1)_{|\mathfrak{a}}\mathrm{ad}(a_2)_{|\mathfrak{a}}$, we are done.

Since κ is invariant, it corresponds to a $\mathfrak{g}$-homomorphism $\theta\colon \mathfrak{g} \to \mathfrak{g}^*$, where $\theta(x)(y) = \kappa(x, y)$. An ideal is simply a subrepresentation of $\mathfrak{g}$ (viewed as a representation under the adjoint action) and $\mathfrak{a}^\perp = \theta^{-1}(\mathfrak{a}^0)$, where $\mathfrak{a}^0$ is the annihilator of $\mathfrak{a}$ in $\mathfrak{a}^*$. But if V is any representation and $U < V$ is a subrepresentation, then its annihilator U^0 is a subrepresentation of V^*, and clearly the preimage of a subrepresentation under a homomorphism of $\mathfrak{g}$-representations is a subrepresentation, thus $\mathfrak{a}^\perp$ is an ideal as required.

□

6.1.2 Cartan criteria for solvable Lie algebras *In this section* k *is an algebraically closed field of characteristic zero.*

We now wish to show how the Killing form yields a criterion for determining whether a Lie algebra is solvable or not. For this we need a couple of technical preliminaries.

Lemma 6.1.6. *Let* $\mathfrak{g}$ *be a Lie algebra and let* $\mathfrak{h}$ *be a Cartan subalgebra with associated Cartan decomposition* $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\lambda\in\Phi} \mathfrak{g}_\lambda$. *Let* (V, ρ) *be a finite dimensional representation of* $\mathfrak{g}$ *and let* $V = \bigoplus_{\mu\in\Psi} V_\mu$ *be the generalised weight-space decomposition of* V *as an* $\mathfrak{h}$*-representation, where* $\Phi \sqcup \{0\}$ *denotes the set of irreducible representations of* $\mathfrak{h}$ *which occur as composition factors of* $\mathfrak{g}$ *and* Ψ *denotes the set which occur as composition factors of* V. *Let* $\lambda \in \Psi$ *and* $\alpha \in \Phi$. *Then there is an* $r \in \mathbb{Q}$ *such that the restriction of* λ *to* $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ *is equal to* $r\alpha$.

Proof. The set of weights Ψ is finite, thus there are positive integers p, q such that $V_{\lambda+t\alpha} \neq 0$ only for integers t with $-p \leq t \leq q$; in particular, $\lambda - (p+1)\alpha \notin \Psi$ and $\lambda + (q+1)\alpha \notin \Psi$. Let $M = \bigoplus_{-p\leq t\leq q} V_{\lambda+t\alpha}$. If $z \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ is of the form $[x, y]$ where $x \in \mathfrak{g}_\alpha, y \in \mathfrak{g}_{-\alpha}$ then, using also Lemma 5.3.31, since

$$\rho(x)(V_{\lambda+q\alpha}) \subseteq V_{\lambda+(q+1)\alpha} = \{0\}, \quad \rho(y)(V_{\lambda-p\alpha}) \subseteq V_{\lambda-(p+1)\alpha} = \{0\}$$

we see that $\rho(x)$ and $\rho(y)$ preserve M. Thus the action of $\rho(z)$ on M is the commutator of the action of $\rho(x)$ and $\rho(y)$ on M, and so $\mathrm{tr}(\rho(z), M) = 0$. On the other hand, we may also compute the trace of $\rho(z)$ on M directly:

$$\begin{aligned} 0 &= \mathrm{tr}(\rho(z), M) \\ &= \sum_{-p\leq t\leq q} \mathrm{tr}(\rho(z), V_{\lambda+t\alpha}) \\ &= \sum_{-p\leq t\leq q} (\lambda(z) + t\alpha(z))\dim(V_{\lambda+t\alpha}). \end{aligned}$$

since any $h \in \mathfrak{h}$ acts on a generalised weight-space V_μ with unique eigenvalue $\mu(h)$. Rearranging the above equation gives $\lambda(z) = r\alpha(z)$ for some $r \in \mathbb{Q}$ as required (where the denominator is a sum of dimensions of subspaces which are not all zero, and hence is nonzero, and clearly r does not depend on z). □

Definition 6.1.7. Let $\mathfrak{g}$ be a Lie algebra over a field k. We say that $\mathfrak{g}$ is *perfect* if $\mathfrak{g} = D(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}]$. A perfect Lie algebra therefore has no nontrivial abelian quotients.

Proposition 6.1.8. *Let* $\mathfrak{g}$ *be a Lie algebra and* (V, ρ) *a* $\mathfrak{g}$*-representation for which* $\rho(\mathfrak{g}) \neq 0$. *Suppose that* $D\mathfrak{g} = \mathfrak{g}$. *Then there is an* $x \in \mathfrak{g}$ *for which* $t_V(x, x) \neq 0$.

Proof. Since $\rho(D\mathfrak{g}) = D(\rho(\mathfrak{g}))$, by replacing $\mathfrak{g}$ with its image in $\mathfrak{gl}(V)$ we may assume that $\mathfrak{g}$ is a subalgebra of $\mathfrak{gl}(V)$ (and hence we will suppress ρ the inclusion map). Suppose that $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$ so that that $\mathfrak{g} = \bigoplus_{\lambda\in\Phi\cup\{0\}} \mathfrak{g}_\lambda$ is the associated Cartan decomposition, where $\mathfrak{h} = \mathfrak{g}_0$.

If we let $V = \bigoplus_{\mu\in\Psi} V_\mu$ be the decomposition of V into generalised $\mathfrak{h}$-weight spaces as in Theorem 5.3.24, then since $\mathfrak{g} \subseteq \mathfrak{gl}(V) \cong V^* \otimes V$, it follows that $\Phi \bigcup\{0\} \subseteq \{\mu_1 - \mu_2 : \mu_1, \mu_2 \in \Psi\}$. In particular, if $V = V_0$, then $\mathfrak{gl}(V) = \mathfrak{gl}(V)_0$ and hence $\mathfrak{g} = \mathfrak{g}_0 = \mathfrak{h}$. But $\mathfrak{h}$ is nilpotent and hence solvable, whereas $D\mathfrak{g} = \mathfrak{g}$ implies $\mathfrak{g}$ is not solvable. It follows there must be some $\lambda \in \Psi\backslash\{0\}$.

Next observe that

$$\mathfrak{g} = D\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}] = [\bigoplus_{\lambda\in\Phi\cup\{0\}} \mathfrak{g}_\lambda, \bigoplus_{\mu\in\Phi\cup\{0\}} \mathfrak{g}_\mu] = \sum_{\lambda,\mu} [\mathfrak{g}_\lambda, \mathfrak{g}_\mu].$$

Since we know that $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subseteq \mathfrak{g}_{\lambda+\mu}$, and moreover $\mathfrak{h} = \mathfrak{g}_0$, it follows that we must have

$$\mathfrak{h} = [\mathfrak{h}, \mathfrak{h}] + \sum_\alpha [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}],$$

where the sum runs over those roots α such that $-\alpha \in \Phi$. But by definition, λ vanishes on $D\mathfrak{h}$, so that there must be some $\alpha \in \Phi$ with $\lambda([\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]) \neq 0$.

Picking $x \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ such that $\lambda(x) \neq 0$, we find that:

$$t_V(x,x) = \mathrm{tr}(x^2) = \sum_{\lambda\in\Psi} \dim(V_\lambda)\lambda(x)^2.$$

But now by Lemma 6.1.6 we know that for any $\mu \in \Psi$ there is an $r_{\alpha,\mu} \in \mathbb{Q}$ such that $\mu(x) = r_{\alpha,\mu}.\alpha(x)$ for all $x \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$. In particular, $0 \neq \lambda(x) = r_{\alpha,\lambda}\alpha(x)$ so that $r_{\alpha,\lambda} \neq 0$ and $\alpha(x) \neq 0$. Hence we see that

$$t_V(x,x) = \left(\sum_{\mu\in\Psi} \dim(V_\mu) r_{\alpha,\mu}^2\right) \alpha(x)^2.$$

Since the terms in the sum are nonnegative, and the term corresponding to λ is positive, we conclude $t_V(x,x) \neq 0$ are required. □

Applying the previous Proposition to the Killing form we can give a criterion for a Lie algebra to be solvable.

Theorem 6.1.9 (*Cartan's criterion for solvability*). *A Lie algebra $\mathfrak{g}$ is solvable if and only if the Killing form restricted to $D\mathfrak{g}$ is identically zero.*

Proof. First suppose that κ vanishes on $D(\mathfrak{g})$. Consider the derived series $\{D^k(\mathfrak{g})\}_{k\geq 0}$. If there is some $k \geq 1$ with $D^k(\mathfrak{g}) = D^{k+1}(\mathfrak{g}) = D(D^k\mathfrak{g}) \neq \{0\}$, then Proposition 6.1.8 applied to $D^k(\mathfrak{g})$ and $(\mathfrak{g}, \mathrm{ad}_{|D^k(\mathfrak{g})})$ the restriction to $D^k(\mathfrak{g})$ of the adjoint representation of $\mathfrak{g}$, shows that there is an $x \in D^k\mathfrak{g}$ with $\kappa^{\mathfrak{g}}(x,x) \neq 0$, and hence $\kappa^{\mathfrak{g}}$ is not identically zero on $D^k(\mathfrak{g})$, contradicting the fact that $D^k(\mathfrak{g}) \subseteq D(\mathfrak{g})$. Thus we conclude $D^{k+1}\mathfrak{g}$ is a proper subspace of $D^k\mathfrak{g}$ whenever $D^k\mathfrak{g}$ is nonzero, and hence since $\mathfrak{g}$ is finite dimensional, it must be solvable as required.

For the converse, assume $\mathfrak{g}$ is solvable. We may replace $\mathfrak{g}$ by its image $\mathfrak{g}_1 = \mathrm{ad}(\mathfrak{g})$ in $\mathfrak{gl}(\mathfrak{g})$ since $\mathfrak{g}$ is solvable if and only if $\mathfrak{g}_1 \cong \mathfrak{g}/\mathfrak{z}(\mathfrak{g})$ is. By Lie's theorem, we can find a complete flag $\mathcal{F} = (0 = F_0 < F_1 < \ldots < F_n = \mathfrak{g})$ in $\mathfrak{g}$ such that $\mathfrak{g}_1 \subseteq \mathfrak{b}_{\mathcal{F}}$. But then by Example 5.1.7 we have $D(\mathfrak{g}_1) \subseteq D(\mathfrak{b}_{\mathcal{F}}) \subseteq \mathfrak{n}_{\mathcal{F}}$. Since $\mathfrak{n}_{\mathcal{F}}$ is an associative subalgebra of $\mathrm{End}(\mathfrak{g})$ consisting of nilpotent endomorphisms, if $x, y \in D(\mathfrak{g})$ then $\mathrm{ad}(x)\mathrm{ad}(y)$ is nilpotent and hence $\kappa^{\mathfrak{g}}(x,y) = \mathrm{tr}(\mathrm{ad}(x)\mathrm{ad}(y)) = 0$ as required.

□

6.2 The radical and semisimple Lie algebras

Suppose that $\mathfrak{g}$ is a Lie algebra, and $\mathfrak{a}$ and $\mathfrak{b}$ are solvable Lie ideals of $\mathfrak{g}$. It is easy to see that $\mathfrak{a} + \mathfrak{b}$ is again solvable (for example, because $0 \subseteq \mathfrak{a} \subseteq \mathfrak{a} + \mathfrak{b}$, and $\mathfrak{a}$ and $(\mathfrak{a} + \mathfrak{b})/\mathfrak{a} \cong \mathfrak{b}/(\mathfrak{a} \cap \mathfrak{b})$ are both solvable). It follows that if $\mathfrak{g}$ is finite dimensional, then it has a largest solvable ideal $\mathfrak{r}$. Note that this is in the strong sense: every solvable ideal of $\mathfrak{g}$ is a subalgebra of $\mathfrak{r}$ (*c.f.* Definition 5.3.18 where the same strategy was used to define the subrepresentation V_S of a $\mathfrak{g}$-representation given an irreducible representation S of $\mathfrak{g}$).

Definition 6.2.1. Let $\mathfrak{g}$ be a finite dimensional Lie algebra. The largest solvable ideal $\mathfrak{r}$ of $\mathfrak{g}$ is known as the (*solvabel*) *radical* of $\mathfrak{g}$, and will be denoted $\mathrm{rad}(\mathfrak{g})$. We say that $\mathfrak{g}$ is *semisimple* if $\mathrm{rad}(\mathfrak{g}) = 0$, that is, if $\mathfrak{g}$ contains no non-zero solvable ideals.

Lemma 6.2.2. *The Lie algebra* $\mathfrak{g}/rad(\mathfrak{g})$ *is semisimple, that is, it has zero radical.*

Proof. Suppose that $\mathfrak{s}$ is a solvable ideal in $\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$. Then if $\mathfrak{s}'$ denotes the preimage of $\mathfrak{s}$ in $\mathfrak{g}$, we see that $\mathfrak{s}'$ is an ideal of $\mathfrak{g}$, and moreover it is solvable since $\mathrm{rad}(\mathfrak{g})$ and $\mathfrak{s} = \mathfrak{s}'/\mathrm{rad}(\mathfrak{g})$ as both solvable. But then by definition we have $\mathfrak{s}' \subseteq \mathrm{rad}(\mathfrak{g})$ so that $\mathfrak{s}' = \mathrm{rad}(\mathfrak{g})$ and $\mathfrak{s} = 0$ as required. □

Example 6.2.3. The Lemma shows that any Lie algebra $\mathfrak{g}$ contains a canonical solvable ideal $\mathrm{rad}(\mathfrak{g})$ such that $\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$ is a semisimple Lie algebra. Thus we have a short exact sequence:

$$0 \longrightarrow \mathrm{rad}(\mathfrak{g}) \longrightarrow \mathfrak{g} \longrightarrow \mathfrak{g}/\mathrm{rad}(\mathfrak{g}) \longrightarrow 0,$$

so that any Lie algebra is an extension of the semisimple Lie algebra $\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$ by the solvable Lie algebra $\mathrm{rad}(\mathfrak{g})$.

In characteristic zero, every Lie algebra $\mathfrak{g}$ is built out of $\mathrm{rad}(\mathfrak{g})$ and $\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$ as a semidirect product.

Theorem 6.2.4. *(Levi's theorem) Let* $\mathfrak{g}$ *be a finite dimensional Lie algebra over a field* k *of characteristic zero, and let* $\mathfrak{r}$ *be its radical. Then there exists a subalgebra* $\mathfrak{s}$ *of* $\mathfrak{g}$ *such that* $\mathfrak{g} \cong \mathfrak{r} \ltimes \mathfrak{s}$. *In particular* $\mathfrak{s} \cong \mathfrak{g}/\mathfrak{r}$ *is semisimple.*

6.2.1 Cartan's Criterion for semisimplicity The Killing form gives us a way of detecting when a Lie algebra is semisimple. Recall that, given a symmetric bilinear form $B\colon V \times V \to \mathsf{k}$, the *radical* of B is $\mathrm{rad}(B) = \{v \in V : \forall w \in V, B(v,w) = 0\} = V^{\perp}$. The form B said to be nondegenerate if $\mathrm{rad}(B) = \{0\}$. We first note the following simple result.

Lemma 6.2.5. *A finite dimensional Lie algebra* $\mathfrak{g}$ *is semisimple if and only if it does not contain any non-zero abelian ideals.*

Proof. Clearly if $\mathfrak{g}$ contains an abelian ideal, it contains a solvable ideal, so that $\mathrm{rad}(\mathfrak{g}) \neq 0$. Conversely, if $\mathfrak{s}$ is a non-zero solvable ideal in $\mathfrak{g}$, then the last term in the derived series of $\mathfrak{s}$ will be an abelian ideal of $\mathfrak{g}$ (*check this*!). □

We have the following simple characterisation of semisimple Lie algebras.

Theorem 6.2.6. *A Lie algebra* $\mathfrak{g}$ *is semisimple if and only if the Killing form is nondegenerate.*

Proof. Let $\mathfrak{g}^{\perp} = \{x \in \mathfrak{g} : \kappa(x,y) = 0, \forall y \in \mathfrak{g}\}$. Then by Lemma 6.1.5 $\mathfrak{g}^{\perp}$ is an ideal in $\mathfrak{g}$, and clearly the restriction of κ to $\mathfrak{g}^{\perp}$ is zero, so by Cartan's Criterion, and Lemma 6.1.5 the ideal $\mathfrak{g}^{\perp}$ is solvable. It follows that if $\mathfrak{g}$ is semisimple we must have $\mathfrak{g}^{\perp} = \{0\}$ and hence κ is non-degenerate.

Conversely, suppose that κ is non-degenerate. To show that $\mathfrak{g}$ is semisimple it is enough to show that any abelian ideal of $\mathfrak{g}$ is trivial, thus suppose that $\mathfrak{a}$ is an abelian ideal, and pick W a complementary subspace to $\mathfrak{a}$ so that $\mathfrak{g} = \mathfrak{a} \oplus W$. With respect to this decomposition, if $x \in \mathfrak{g}$ and $a \in \mathfrak{a}$, we have

$$\mathrm{ad}(x) = \begin{pmatrix} x_1 & x_2 \\ 0 & x_3 \end{pmatrix}, \quad \mathrm{ad}(a) = \begin{pmatrix} 0 & a_2 \\ 0 & 0 \end{pmatrix} \in \begin{pmatrix} \mathrm{Hom}_{\mathsf{k}}(\mathfrak{a},\mathfrak{a}) & \mathrm{Hom}_{\mathsf{k}}(W,\mathfrak{a}) \\ \mathrm{Hom}_{\mathsf{k}}(\mathfrak{a},W) & \mathrm{Hom}_{\mathsf{k}}(W,W) \end{pmatrix}.$$

But then we see that $\mathrm{ad}(x) \circ \mathrm{ad}(a) = \begin{pmatrix} 0 & x_1 a_2 \\ 0 & 0 \end{pmatrix}$, and hence $\mathrm{tr}(\mathrm{ad}(x)\mathrm{ad}(a)) = 0$. It follows that $\mathfrak{a} \subseteq \mathfrak{g}^{\perp} = \{0\}$ as κ is non-degenerate and hence $\mathfrak{a} = \{0\}$ as required. □

Remark 6.2.7. It is worth noting that the proof of the previous theorem establishes two facts: first, that $\mathfrak{g}^\perp$ is a solvable ideal in $\mathfrak{g}$ for any Lie algebra $\mathfrak{g}$, and secondly, that any abelian ideal of $\mathfrak{g}$ is contained in $\mathfrak{g}^\perp$. Combined with the previous Lemma this shows that $\mathfrak{g}^\perp = \{0\} \iff \mathrm{rad}(\mathfrak{g}) = \{0\}$, but in general the containment $\mathfrak{g}^\perp \subseteq \mathrm{rad}(\mathfrak{g})$ need not be an equality.

6.2.2 Simple and semisimple Lie algebras Recall that we say that a Lie algebra is *simple* if it is non-Abelian and has no nontrivial proper ideal. We now show that this notion is closed related to our notion of a semisimple Lie algebra.

Proposition 6.2.8. *Let $\mathfrak{g}$ be a semisimple Lie algebra, and let I be an ideal of $\mathfrak{g}$. Then $\mathfrak{g} = I \oplus I^\perp$.*

Proof. Since $\mathfrak{g}$ is semisimple, the Killing form is nondegenerate, hence by Lemma I.17 in Appendix 1, we have

$$\dim(I) + \dim(I^\perp) = \dim(\mathfrak{g}). \tag{6.1}$$

Now consider $I \cap I^\perp$. The Killing form of $\mathfrak{g}$ vanishes identically on $I \cap I^\perp$ by definition, and since it is an ideal, the Killing form of $I \cap I^\perp$ is just the restriction of the Killing form of $\mathfrak{g}$. It follows from Cartan's Criterion that $I \cap I^\perp$ is solvable, and hence since $\mathfrak{g}$ is semisimple we must have $I \cap I^\perp = 0$. But then by Equation (6.1) we must have $\mathfrak{g} = I \oplus I^\perp$ as required (note that this is a direct sum of Lie algebras, since $[I, I^\perp] \subset I \cap I^\perp$). □

Proposition 6.2.9. *Let $\mathfrak{g}$ be a semisimple Lie algebra.*

1. *Any ideal and any quotient of $\mathfrak{g}$ is semisimple.*
2. *Then there exist ideals $\mathfrak{g}_1, \mathfrak{g}_2, \dots \mathfrak{g}_k \subseteq \mathfrak{g}$ which are simple Lie algebras and for which the natural map:*

$$\mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \dots \oplus \mathfrak{g}_k \to \mathfrak{g},$$

is an isomorphism. Moreover, any simple ideal $\mathfrak{a} \in \mathfrak{g}$ is equal to some $\mathfrak{g}_i$ ($1 \leq i \leq k$). In particular the decomposition above is unique up to reordering, and $\mathfrak{g} = D\mathfrak{g}$.

Proof. For the first part, if I is an ideal of $\mathfrak{g}$, by the previous Proposition we have $\mathfrak{g} = I \oplus I^\perp$, so that the Killing form of $\mathfrak{g}$ restricted to I is nondegenerate. Since this is just the Killing form of I, Cartan's criterion shows that I is semisimple. Moreover, clearly $\mathfrak{g}/I \cong I^\perp$ so that any quotient of $\mathfrak{g}$ is isomorphic to an ideal of $\mathfrak{g}$ and hence is also semisimple.

For the second part we use induction on the dimension of $\mathfrak{g}$. Let $\mathfrak{a}$ be a minimal non-zero ideal in $\mathfrak{g}$. If $\mathfrak{a} = \mathfrak{g}$ then $\mathfrak{g}$ is simple, so we are done. Otherwise, we have $\dim(\mathfrak{a}) < \dim(\mathfrak{g})$. Then $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$, and by induction $\mathfrak{a}^\perp$ is a direct sum of simple ideals, and hence clearly $\mathfrak{g}$ is also.

To show the moreover part, suppose that $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \dots \oplus \mathfrak{g}_k$ is a decomposition as above and $\mathfrak{a}$ is a simple ideal of $\mathfrak{g}$. Now as $\mathfrak{z}(\mathfrak{g}) = \{0\}$, we must have $0 \neq [\mathfrak{g}, \mathfrak{a}] \subset \mathfrak{a}$, and hence by simplicity of $\mathfrak{a}$ it follows that $[\mathfrak{g}, \mathfrak{a}] = \mathfrak{a}$. But then we have

$$\mathfrak{a} = [\mathfrak{g}, \mathfrak{a}] = [\bigoplus_{i=1}^{k} \mathfrak{g}_i, \mathfrak{a}] = [\mathfrak{g}_1, \mathfrak{a}] \oplus [\mathfrak{g}_2, \mathfrak{a}] \oplus \dots \oplus [\mathfrak{g}_k, \mathfrak{a}],$$

(the ideals $[\mathfrak{g}_i, \mathfrak{a}]$ are contained in $\mathfrak{g}_i$ so the last sum remains direct). But $\mathfrak{a}$ is simple, so direct sum decomposition must have exactly one nonzero summand and we have $\mathfrak{a} = [\mathfrak{g}_i, \mathfrak{a}]$ for some i ($1 \leq i \leq k$). Finally, using the simplicity of $\mathfrak{g}_i$ we see that $\mathfrak{a} = [\mathfrak{g}_i, \mathfrak{a}] = \mathfrak{g}_i$ as required. To see that $\mathfrak{g} = D\mathfrak{g}$ note that it is now enough to check it for simple Lie algebras, where it is clear[3].

□

[3]This is one reason for insisting simple Lie algebras are nonabelian.

6.3 Weyl's theorem

In this section we assume that our field is algebraically closed of characteristic zero.

In this section we study the representations of a semisimple Lie algebra. eviews the other basic representation theory that we will need in this section. Our goal is to show, just as for representations of a finite group over $\mathbb{C}$, that every representation is *completely irreducible*, that is, is a direct sum of irreducible representations. Note that, as we check Lemma II.13 in the Appendix 2, V is completely reducible if and only if V is the sum of its irreducible subrepresentations.

Definition 6.3.1. We say that a representation V of a Lie algebra is *semisimple* if any subrepresentation has a complement, that is, for any subrepresentation U of V, there is a subrepresentation W such that $V = U \oplus W$.

Lemma 6.3.2. *Let $\mathfrak{g}$ be a Lie algebra.*

i) A representation V is semisimple if and only if any short exact sequence

$$0 \longrightarrow U \xrightarrow{i} V \xrightarrow{q} W \longrightarrow 0$$

splits.

iii) If the representations of $\mathfrak{g}$ are semisimple then they are completely reducible.

Proof. For the first part, suppose that any short exact sequence with V as its middle term splits. Then if U is a subrepresentation of V, then if $i\colon U \to V$ denotes the inclusion map and $W = V/U$ with $q\colon V \to W$ the quotient map, we obtain a short exact sequence. A splitting of this short exact sequence $s\colon W \to V$ is determined by its image, because q restricts to an isomorphism on $s(W)$. Now if $v \in U \cap s(W)$ then $v = s(w)$ for some $w \in W$, then $0 = q(v) = q \circ s(w) = w$, and hence $s(w) = 0$. It follows that $V = U \oplus s(W)$ and $s(W)$ is a complement to U.

Conversely, if V is semisimple, then given any short exact sequence with V as its middle term, the subrepresentation $i(U)$ has a complement U'. Then $q_{|U'}\colon U' \to W$ is an isomorphism, hence we may define $s\colon W \to V$ to be its inverse.

For the second part, use induction on $\dim(V)$. If V is irreducible then we are clearly done, otherwise V has a proper subrepresentation U. But then U has a complement in V, say $V = U \oplus T$. But since $\dim(U), \dim(T) < \dim(V)$, they are completely reducible, hence V is completely reducible as required. □

Remark 6.3.3. hows that if a $\mathfrak{g}$-representation is semisimple, then all of its subrepresentations and quotients are also semisimple. Knowing this, the above proof shows that if a representation is semisimple then it is completely reducible (whereas in the above we showed that if *all* representations of a Lie algebra $\mathfrak{g}$ are semisimple, then they are all completely reducible.)

Theorem 6.3.4. *(Weyl's theorem.) If $\mathfrak{g}$ is a semisimple Lie algebra, then any short exact sequence of representations of $\mathfrak{g}$ splits. Equivalently any surjective map of $\mathfrak{g}$-representations has a splitting. Consequently, every representation of $\mathfrak{g}$ is semisimple, and thus completely reducible.*

Remark 6.3.5. We will establish this theorem by first showing that, for any representation of a semisimple Lie algebra $\mathfrak{g}$, the invariants $V^{\mathfrak{g}}$ form a direct summmand of V. Then, using this decomposition and the fact that, if V and W are $\mathfrak{g}$-representations then $\mathrm{Hom}(V, W)$ is also, to deduce the semisimplicity result. The argument is similar, but not identical, to the one used in proving representations of a finite group over the complex numbers are semisimple.

6.3.1 Casimir operators

Lemma 6.3.6. *Suppose that $\mathfrak{g}$ is semisimple and (V, ρ) is a representation of $\mathfrak{g}$. Then the radical of t_V is precisely the kernel of ρ. In particular if (V, ρ) is faithful then t_V is nondegenerate.*

Proof. The image $\mathfrak{g}_1 = \rho(\mathfrak{g}) \subseteq \mathfrak{gl}(V)$ of $\mathfrak{g}$ is a semisimple Lie algebra (since $\mathfrak{g}$ is) and the statement of the Lemma is exactly that t_V is nondegenerate on $\mathfrak{g}_1$. But the radical $\mathfrak{r} = \mathrm{rad}(t_V)$ is an ideal of $\rho(\mathfrak{g})$. Now Proposition 11.2 shows that if we let $(D^k\mathfrak{r})_{k\geq 0}$ be the derived series of $\mathfrak{r}$, we must have $D^{k+1}\mathfrak{r} \subsetneq D^k\mathfrak{r}$ whenever $D^k\mathfrak{r} \neq \{0\}$, thus $\mathfrak{r}$ must be solvable. Since $\mathfrak{g}_1$ is semisimple, this forces the radical to be zero as required. $\square$

This allows us to make the following definition:

Definition 6.3.7. Let $\mathfrak{g}$ be a semisimple Lie algebra and (V, ρ) a representation of $\mathfrak{g}$ with $\rho(\mathfrak{g}) \neq 0$. Then we set $\mathfrak{g}_1 = \rho(\mathfrak{g}) \subseteq \mathfrak{gl}(V)$, and, as above, write $\tau\colon \mathfrak{g}_1 \to \mathfrak{g}_1^*$ for the isomorphism induced by t_V. We have a sequence of $\mathfrak{g}_1$ (and $\mathfrak{g}$) -homomorphism

$$\mathrm{Hom}(\mathfrak{g}_1, \mathfrak{g}_1) \longrightarrow \mathfrak{g}_1^* \otimes \mathfrak{g}_1 \longrightarrow \mathfrak{g}_1 \otimes \mathfrak{g}_1 \longrightarrow \mathfrak{gl}(V)$$

where the first map is θ^{-1} (see Appendix I.2) the second is $\tau^{-1} \otimes 1$ and the third is composition of linear maps. These are all $\mathfrak{g}$-homomorphisms, and hence the image of $\mathrm{Id}_{\mathfrak{g}_1}$ under their composition is a $\mathfrak{g}$-endomorphism of V, which we denote by C_V, the *Casimir operator* of V. Note also that the composite map $\mathrm{Hom}(\mathfrak{g}_1, \mathfrak{g}_1) \to \mathfrak{gl}(V)$ is compatible with the trace map, which implies that

$$\mathrm{tr}_V(C_V) = \mathrm{tr}_{\mathfrak{g}_1}(\mathrm{Id}_{\mathfrak{g}_1}) = \dim(\mathfrak{g}_1) \neq 0.$$

Remark 6.3.8. If $\mathfrak{g}$ is simple, rather than just semisimple, then by Schur's Lemma (which holds for representations of Lie algebras just as it does for groups – see Lemma II.2) $\mathrm{Hom}_{\mathsf{k}}(\mathfrak{g}, \mathfrak{g})^{\mathfrak{g}} = \mathrm{Hom}_{\mathfrak{g}}(\mathfrak{g}, \mathfrak{g})$ is one-dimensional (the scalar multiples of the identity). Since $\mathrm{Hom}_{\mathsf{k}}(\mathfrak{g}, \mathfrak{g}) \cong \mathfrak{g} \otimes \mathfrak{g}$ as $\mathfrak{g}$-representations, the invariants $(\mathfrak{g} \otimes \mathfrak{g})^{\mathfrak{g}}$ in $\mathfrak{g} \otimes \mathfrak{g}$ must also be one-dimensional (the image of the scalar multiples of the identity under any isomorphism). If we pick a non-zero element $C \in (\mathfrak{g} \otimes \mathfrak{g})^{\mathfrak{g}}$, then, for any representation on which $\mathfrak{g}$ acts non-trivially, there is a non-zero scalar λ_V such that

$$C_V = \lambda_V . m \circ (\rho \otimes \rho)(C)$$

Thus the Casimir operators C_V, up to scaling, all come from the same element of $\mathfrak{g} \otimes \mathfrak{g}$.

Example 6.3.9. Let us take $\mathfrak{g} = \mathfrak{sl}_2$. Then the trace form $t(x, y) = \mathrm{tr}(x.y)$ is non-degenerate and invariant, with

$$t(e, f) = t(e, h) = 1, \quad t(h, h) = 2, \quad t(e, e) = t(f, f) = t(e, h) = t(f, h) = 0$$

Thus the corresponding isomorphism $\tau\colon \mathfrak{sl}_2 \to \mathfrak{sl}_2^*$ gives

$$(\tau^{-1} \otimes 1)\theta^{-1}(\mathrm{id}) = f \otimes e + \frac{1}{2} h \otimes h + e \otimes f.$$

For any $\mathfrak{sl}_2$-representation (V, ρ) we thus get a $\mathfrak{g}$-endomorphism of V by applying $m \circ (\rho \otimes \rho)$ to this element, namely $\rho(e)\rho(f) + \frac{1}{2}\rho(h)^2 + \rho(f)\rho(e)$. This is exactly the operator used in Sheet 3 of the problem set.

Definition 6.3.10. Recall that if (V, ρ) is a representation of $\mathfrak{g}$, then $V^{\mathfrak{g}} = \{v \in V : \rho(x)(v) = 0, \forall x \in \mathfrak{g}\}$ is the subrepresentation of invariants in V. We also define $\mathfrak{g}.V = \mathrm{span}\{\rho(x)(v) : x \in \mathfrak{g}, v \in V\}$. One can check directly that $\mathfrak{g}.V$ is a subrepresentation, or note that it is the image of the $\mathfrak{g}$-homomorphism $a\colon \mathfrak{g} \otimes V \to V$ given by $a(x \otimes v) = \rho(x)(v)$. See Example 3.3.8 for more details.

The next Proposition shows that the invariants in a representation of semisimple Lie algebra form a direct summand of V. This can be thought of as the analogue of Maschke's theorem: The key to that result of that, if G is a finite group, and (V, ρ) is any representation of G, then the operator $\pi_0 = |G|^{-1} \sum_{g \in G} \rho(g)$ is a G-equivariant projection onto V^G the invariants in V and hence its kernel gives a complementary subrepresentation, so that $V = V^G \oplus \ker(\pi_0)$.

Proposition 6.3.11. *Let (V, ρ) be representation of a semisimple Lie algebra $\mathfrak{g}$. Then $V = V^{\mathfrak{g}} \oplus \mathfrak{g}.V$.*

Proof. We prove the statement by induction on $\dim(V)$, the case $\dim(V) = 0$ being trivial). If $V = V^{\mathfrak{g}}$ the certainly $\mathfrak{g}.V = \{0\}$ and the statement holds. Thus we may assume that $V \neq V^{\mathfrak{g}}$, so that $\rho(\mathfrak{g}) \neq \{0\}$. Let C_V be the Casimir operator of V. Since it is a $\mathfrak{g}$-endomorphism, if $V = \bigoplus V_\lambda$ is the decomposition of V into the generalised eigenspaces of C_V, each V_λ is a subrepresentations of V. Since if the statement of the Lemma holds for representations U and W it certainly holds for their direct sum $U \oplus W$, we are done by induction unless C_V has exactly one generalised eigenspace, *i.e.* $V = V_\lambda$. But then $\dim(V).\lambda = \mathrm{tr}(C_V) = \dim(\rho(\mathfrak{g}))$, so that $\lambda \neq 0$[4], and hence C_V is invertible. Since it is clear from the definition of C_V that $V^{\mathfrak{g}} \subseteq \ker(C_V)$ we see that $V^{\mathfrak{g}} = \{0\}$, and moreover $V = C_V(V) \subseteq \rho(\mathfrak{g})(\rho(\mathfrak{g})(V))$, so that $V = \mathfrak{g}.V$, and we are done. □

Definition 6.3.12. If $\mathfrak{g}$ is a semisimple Lie algebra and (V, ρ) is a representation of $\mathfrak{g}$, then let $\pi_0^V : V \to V^{\mathfrak{g}}$ be the projection to $V^{\mathfrak{g}}$ with kernel $\mathfrak{g}.V$. To avoid cluttered notation, we will sometimes just write π_0. The following Corollary indicates shows this is unlikely to cause confusion:

Corollary 6.3.13. *Let V and W be $\mathfrak{g}$-representations and let $\theta \in \mathrm{Hom}_{\mathfrak{g}}(V, W)$. Then $\pi_0^W \circ \theta = \theta \circ \pi_0^V$.*

Proof. From the definitions it is clear that $\theta(V^{\mathfrak{g}}) \subseteq W^{\mathfrak{g}}$ and similarly $\theta(\mathfrak{g}.V) \subseteq \mathfrak{g}.W$. The result follows immediately. □

We now prove that $\mathfrak{g}$-representations are semisimple following the same strategy as in the case of finite groups:

Theorem 6.3.14. *Let V be a representation of a semisimple Lie algebra $\mathfrak{g}$. Then V is semisimple, that is, any subrepresentation has a complement.*

Proof. Let $i \colon U \to V$ denotes the inclusion map and $q \colon V \to W = V/U$ the quotient map, so that we have a short exact sequence:

$$0 \longrightarrow U \xrightarrow{i} V \xrightarrow{q} W \longrightarrow 0$$

If $s \colon W \to V$ is any linear map satisfying $q \circ s = 1_W$, that is, a linear splitting of the quotient map q, then $\mathrm{im}(s)$ is a complement to $\mathrm{im}(i) = \ker(q)$. Thus if s is $\mathfrak{g}$-invariant for the $\mathfrak{g}$-action on $\mathrm{Hom}(W, V)$ so that $s \in \mathrm{Hom}_{\mathfrak{g}}(W, V)$, then $\mathrm{im}(s)$ is a complementary subrepresentation to $\mathrm{im}(i) = \ker(q)$ as required.

Now the map $q_* \colon \mathrm{Hom}(W, V) \to \mathrm{Hom}(W, W)$ given by $q_*(\psi) = q \circ \psi$ is a $\mathfrak{g}$-homomorphism because q is. Moreover, if s is any linear splitting map, and $\phi \in \mathrm{Hom}(W, W)$, we have $q_*(s \circ \phi) = (q \circ s) \circ \phi = 1_W \circ \phi = \phi$, so that q_* is surjective. But now note that $1_W \in \mathrm{Hom}(W, W)^{\mathfrak{g}}$ and hence by Corollary 6.3.13 we have

$$q_*(\pi_0(s)) = \pi_0(q_*(s)) = \pi_0(q \circ s) = \pi_0(1_W) = 1_W.$$

It follows that $\pi_0(s)$ is the required splitting map, and $V = U \oplus \mathrm{im}(\pi_0(s))$ as required.

□

6.4 The Jordan Decomposition

Unless explicitly stated to the contrary, in this section we work over a field k which is algebraically closed of characteristic zero.

Definition 6.4.1. Let V be a k-vector space and suppose that $x \in \mathrm{End}(V)$. Then we say $x \in \mathrm{End}(V)$ is *semisimple* if V is a direct sum of eigenlines for x, that is, there are lines $L_1, \dots, L_n$ such that $V = \bigoplus_{i=1}^n L_i$ and $x(L_i) \subseteq L_i$. We say that x is *nilpotent* if $V = V_0$, that is x has 0 as its only eigenvalue, and hence $x^n = 0$ for sufficiently large n.

[4]This is where we use that the characteristic of the field is 0.

By Proposition I.1 for any $x \in \mathrm{End}(V)$, then we may write

$$V = \bigoplus_{\lambda \in \mathsf{k}} V_{\lambda,x}, \quad \text{where } V_{\lambda,x} = \{v \in V : \exists N > 0, (x-\lambda)^N(v) = 0\}.$$

That is, $V_{\lambda,x}$ is the generalized λ-eigenspace of x. This direct sum decomposition can be used to give a decomposition of the endomorphism x in a semisimple (or diagonalisable) and nilpotent part:

Proposition 6.4.2. *Let V be a finite dimensional vector space $x \in End(V)$. Then we may write $x = x_s + x_n$ where x_s is semisimple and x_n is nilpotent, and x_s and x_n commute, i.e. $[x_s, x_n] = 0$. Moreover, if U is a subspace of V preserved by x, it is also preserved by x_s, x_n.*

A proof is given in Appendix I.1. In fact, given $x = x_s + x_n$, the conditions that x_s is semisimple and x_n is nilpotent along with the fact that they commute, determines them uniquely. To see this we use the following:

Lemma 6.4.3. *Let V be a k-vector space and $x, n \in End(V)$ be such that $[x, n] = 0$ and n is nilpotent. Then we have $V_{\lambda,x} = V_{\lambda,x+n}$.*

Proof. It suffices to show that $V_{\lambda,x} \subseteq V_{\lambda,x+n}$ for all such pairs (x, n) in $\mathrm{End}(V)$. Indeed the lemma follows once one also knows the reverse inclusion, but this follows by considering the pair $(x+n, -n)$.

To prove the claim, note that since $[x, n] = 0$, we have $n(V_{\lambda,x}) \subseteq V_{\lambda,x}$. But by definition $(x - \lambda)$ is nilpotent on $V_{\lambda,x}$ and hence $(x + n) - \lambda = (x - \lambda) + n$, when restricted to $V_{\lambda,x}$, is the sum of two commuting nilpotent endomorphisms of $V_{\lambda,x}$. It follows from Lemma 5.3.14 that $(x + n) - \lambda$ acts nilpotently on $V_{\lambda,x}$, and hence $V_{\lambda,x} \subseteq V_{x+n,\lambda}$ as required. □

Corollary 6.4.4. *The (naive or concrete) Jordan decomposition is unique, that is, given $x \in End(V)$, there is a unique pair (x_s, x_n) with x_s semisimple and x_n nilpotent, such that $x = x_s + x_n$ and $[x_s, x_n] = 0$.*

Proof. The previous Lemma shows that x_s and $x_s + x_n = x$ have the same generalised eigenspaces. But as x_s is semisimple, its generalized eigenspaces are precisely its eigenspaces and hence it is completely determined by these. It follows x_s is unique, and hence $x_n = x - x_s$ is also. □

Lemma 6.4.5. *Let V be a vector space and $x \in End(V)$. If x is semisimple then*

$$ad(x) \colon End(V) \to End(V)$$

is also semisimple, and similarly if x is nilpotent.

Proof. First note that the action of $\mathrm{ad}(x)$ on $\mathfrak{gl}(V)$ is just the action of x on the tensor product $V^* \otimes V$. When x is nilpotent, the result is proved in Lemma 5.3.15.

If x is semisimple, then we may write $V = \bigoplus_{i=1}^n L_i$ and $V^* = \bigoplus_{j=1}^n L_j^*$. But then

$$V^* \otimes V = (\bigoplus_{i=1}^n L_i)^* \otimes (\bigoplus_{j=1}^n L_j) = \bigoplus_{i,j=1}^n L_i^* \otimes L_j,$$

and since $\dim(L_i^* \otimes L_j) = 1$, it follows that the action of $\mathrm{ad}(x)$ on $\mathfrak{gl}(V)$ is semisimple also. □

Corollary 6.4.6. *Let $x \in End(V)$, and suppose $x = x_s + x_n$ is the Jordan decomposition of x. Then $ad(x) = ad(x_s) + ad(x_n)$ is the Jordan decomposition of $ad(x)$.*

Proof. By the previous Lemma, $\mathrm{ad}(x_s)$ and $\mathrm{ad}(x_n)$ are semisimple and nilpotent respectively, and as ad is a representation, $[\mathrm{ad}(x_s), \mathrm{ad}(x_n)] = \mathrm{ad}([x_s, x_n]) = 0$. □

We now return to Lie algebras. The above linear algebra allows us to define an "abstract" Jordan decomposition for the elements of any Lie algebra (over an algebraically closed field).

Definition 6.4.7. Suppose that $\mathfrak{g}$ is a Lie algebra and $x \in \mathfrak{g}$. The endomorphism $\mathrm{ad}(x) \in \mathfrak{gl}(\mathfrak{g})$ has a unique Jordan decomposition $\mathrm{ad}(x) = \mathrm{ad}(x)_s + \mathrm{ad}(x)_n$ in $\mathfrak{gl}(\mathfrak{g})$. Then if $s, n \in \mathfrak{g}$ are such that $\mathrm{ad}(s) = \mathrm{ad}(x)_s$ and $\mathrm{ad}(n) = \mathrm{ad}(x)_n$, we say the Lie algebra elements s, n are an *abstract Jordan decomposition* of x. Note that if $z \in \mathfrak{z}(\mathfrak{g}) \neq \{0\}$ then if (s, n) is a Jordan decomposition of x so is $(s + z, n - z)$, thus the Jordan decomposition is unique if and only if $\mathfrak{z}(\mathfrak{g}) = \{0\}$.

Note that that if $\mathfrak{g} = \mathfrak{gl}(V)$ for some vector space V, then Lemma 6.4.5 shows that the abstract Jordan decomposition for an element $x \in \mathfrak{gl}(V)$ is just the naive one (*i.e.* the one for x thought of as a linear map from V to itself).

For a Lie algebra $\mathfrak{g}$, the space $\mathrm{Der}_{\mathsf{k}}(\mathfrak{g})$ of k-derivations of $\mathfrak{g}$ is a Lie algebra, which we may view as a subalgebra of the Lie algebra $\mathfrak{gl}(\mathfrak{g})$. The map $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is in fact a Lie algebra homomorphism from $\mathfrak{g}$ into $\mathrm{Der}_{\mathsf{k}}(\mathfrak{g})$. Its image is denoted $\mathrm{Inn}_{\mathsf{k}}(\mathfrak{g})$.

Lemma 6.4.8. *Let $\mathfrak{a}$ be a semisimple Lie algebra.*

1. *Suppose that $\mathfrak{a}$ is an ideal of a Lie algebra $\mathfrak{g}$. Then $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$.*
2. *All derivations of $\mathfrak{a}$ are inner, that is, $\mathrm{Der}_{\mathsf{k}}(\mathfrak{a}) = \mathrm{Inn}_{\mathsf{k}}(\mathfrak{a})$.*

Proof. For the first part, let κ denote the Killing form for $\mathfrak{g}$ so that $\mathfrak{a}^\perp = \{x \in \mathfrak{g} : \kappa(x, y) = 0, \forall y \in \mathfrak{a}\}$ is an ideal in $\mathfrak{g}$. Now since $\mathfrak{a}$ is an ideal in $\mathfrak{g}$, the restriction $\kappa_{|\mathfrak{a}}$ of the Killing form of $\mathfrak{g}$ to $\mathfrak{a}$ is the Killing form $\kappa^{\mathfrak{a}}$ of $\mathfrak{a}$. Since $\mathfrak{a}$ is semisimple, by Cartan's Criterion $\kappa^{\mathfrak{a}}$ is non-degenerate, hence $\kappa_{|\mathfrak{a}}$ is. But then Lemma I.17 shows that $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$ as required.

For the second part, note that the Lie algebra of derivations $D = \mathrm{Der}_{\mathsf{k}}(\mathfrak{a})$ is a subalgebra of $\mathfrak{gl}(\mathfrak{a})$ containing the image I of ad as the subalgebra of "inner derivations" which, since it is isomorphic to $\mathfrak{a}$, is semisimple. We first claim that this subalgebra is an ideal: indeed if $\mathrm{ad}(x)$ is any inner derivation, and $\delta \in D$, then

$$[\delta, \mathrm{ad}(x)](y) = \delta[x, y] - [x, \delta(y)] = [\delta(x), y] = \mathrm{ad}(\delta(x))(y)$$

thus $[\delta, \mathrm{ad}(x)] \in I$, and hence I is an ideal in D. Now since I is semisimple, by the first part we see that $D = I \oplus I^\perp$, thus it is enough to prove that $I^\perp = \{0\}$. Thus suppose that $\delta \in I^\perp$. Then since $[I, I^\perp] \subset I \cap I^\perp = \{0\}$ we see that

$$[\delta, \mathrm{ad}(x)] = \mathrm{ad}(\delta(x)) = 0, \forall x \in \mathfrak{a},$$

so that, again by the injectivity of ad, we have $\delta = 0$ and so $I^\perp = \{0\}$ as required. □

Lemma 6.4.9. *Let $\mathfrak{a}$ be a Lie algebra and $\mathrm{Der}_{\mathsf{k}}(\mathfrak{a}) \subset \mathfrak{gl}(\mathfrak{a})$ the Lie algebra of k-derivations on $\mathfrak{a}$. Let $\delta \in \mathrm{Der}_{\mathsf{k}}(\mathfrak{a})$. If $\delta = s + n$ is the Jordan decomposition of δ as an element of $\mathfrak{gl}(\mathfrak{a})$, then $s, n \in \mathrm{Der}_{\mathsf{k}}(\mathfrak{a})$.*

Proof. We may decompose $\mathfrak{a} = \bigoplus_\lambda \mathfrak{a}_\lambda$ where $\mathfrak{a}_\lambda$ is the generalized eigenspace of δ with eigenvalue $\lambda \in \mathsf{k}$ say. Now since δ is a derivation the map $\mathfrak{a}_\lambda \otimes \mathfrak{a}_\mu \to \mathfrak{a}$ given by $x \otimes y \mapsto [x, y]$ is compatible with the action of δ. But then by Lemma 5.3.22, if $x \in \mathfrak{a}_\lambda$ and $y \in \mathfrak{a}_\mu$, we have $[x, y] \in \mathfrak{a}_{\lambda+\mu}$. It follows immediately that s is a derivation on $\mathfrak{a}$, and since $n = \delta - s$ we see that n is also. □

Theorem 6.4.10. *Let $\mathfrak{g}$ be a semisimple Lie algebra. Then given any $x \in \mathfrak{g}$ has an abstract Jordan decomposition: that is, there exist unique elements $s, n \in \mathfrak{g}$ such that $x = s + n$ and $[s, n] = 0$, and $\mathrm{ad}(s)$ is semisimple, while $\mathrm{ad}(n)$ is nilpotent.*

Proof. As noted above, since $\mathfrak{g}$ is semisimple, $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is an embedding, and the conditions on s and n show that if they exist, they must satisfy $\mathrm{ad}(s) = \mathrm{ad}(x)_s$ and $\mathrm{ad}(n) = \mathrm{ad}(x)_n$, where $\mathrm{ad}(x) = \mathrm{ad}(x)_s + \mathrm{ad}(x)_n$ is the Jordan decomposition of $\mathrm{ad}(x) \in \mathfrak{gl}(\mathfrak{g})$. Thus it remains to show that $\mathrm{ad}(x)_s$ and $\mathrm{ad}(x)_n$ lie in the image of ad. But $\mathrm{ad}(x)$ acts as a derivation on $I = \mathrm{im}(\mathrm{ad})$, so by Lemma 6.4.9 so do $\mathrm{ad}(x)_s$ and $\mathrm{ad}(x)_n$. But then by Lemma 6.4.8, we see that $\mathrm{ad}(x)_s = \mathrm{ad}(s)$ for some $s \in \mathfrak{g}$ and $\mathrm{ad}(x)_n = \mathrm{ad}(n)$ for some $n \in \mathfrak{g}$. The conditions on $s, n \in \mathfrak{g}$ then follow from the injectivity of ad, and we are done. □

Remark 6.4.11. One can show that the Jordan decomposition is compatible with representations, in the sense that if (V, ρ) is a representation of a semsimple Lie algebra $\mathfrak{g}$ and $x = s + n$ is the Jordan decomposition of $x \in \mathfrak{g}$, then $\rho(x) = \rho(s) + \rho(n)$ is the (naive) Jordan decomposition of $\rho(x) \in \mathfrak{gl}(V)$. The main point, of course, is to show that $\rho(s)$ and $\rho(n)$ are, respectively, semisimple and nilpotent endomorphisms of V. The proof, which we shall not give here, is similar to the proof of the existence of the abstract Jordan form, except that one also needs to use Weyl's theorem.

Chapter 7

Root systems and the classification of semisimple Lie algebras

7.1 The Cartan decomposition of a semisimple Lie algebra

In this section we work over an algebraically closed field k *of characteristic zero.*

Although the Cartan decomposition makes sense in any Lie algebra, we will now restrict attention to semisimple Lie algebras $\mathfrak{g}$, where we can give much more precise information about the structure of the root spaces than in the general case.

Proposition 7.1.1. *Suppose that* $\mathfrak{g}$ *is a semisimple Lie algebra and* $\mathfrak{h}$ *is a Cartan subalgebra, and* $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha$ *the associated Cartan decomposition.*

1. *Let* κ *be the Killing form. Then* $\kappa(\mathfrak{g}_\lambda, \mathfrak{g}_\mu) = 0$ *unless* $\lambda + \mu = 0$.
2. *If* $\alpha \in \Phi$*, then* $-\alpha \in \Phi$.
3. *The restriction* $\kappa_{|\mathfrak{h}}$ *of* κ *to* $\mathfrak{h}$ *is nondegenerate.*

Proof. For the first part, since $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subseteq \mathfrak{g}_{\lambda+\mu}$, if $x \in \mathfrak{g}_\lambda, y \in \mathfrak{g}_\mu$, we see that $\mathrm{ad}(x)\mathrm{ad}(y)(\mathfrak{g}_\nu) \subseteq \mathfrak{g}_{\lambda+\mu+\nu}$. But then picking a basis of $\mathfrak{g}$ compatible with the Cartan decomposition it is clear the matrix of $\mathrm{ad}(x)\mathrm{ad}(y)$ will have no non-zero diagonal entry unless $\lambda+\mu = 0$, hence $\kappa(x,y) = 0$ unless $\lambda + \mu = 0$ as required.

For the second part, recall that if α is a root, then $\alpha \neq 0$ and $\mathfrak{g}_\alpha \neq 0$. If $-\alpha \notin \Phi$ then $\mathfrak{g}_{-\alpha} = 0$ and so $\mathfrak{g}_\alpha \subseteq \mathrm{rad}(\kappa) = \mathfrak{g}^\perp = \{0\}$, which is impossible since $\mathfrak{g}_\alpha$ is non-zero by assumption.

For the third part note that $\mathfrak{h}^\perp$ contains all the $\mathfrak{g}_\alpha$ for $\alpha \in \Phi$ by part (1). Since κ is nondegenerate, by dimension counting this must be equal to $\mathfrak{h}^\perp$. It follows that $\kappa_{|\mathfrak{h}}$ must be nondegenerate as claimed. □

Lemma 7.1.2. *Let* $\mathfrak{g}$ *be a semisimple Lie algebra and* $\mathfrak{h}$ *a Cartan subalgebra. Then if* $x, y \in \mathfrak{h}$ *we have*

$$\kappa(x,y) = \sum_{\alpha \in \Phi} dim(\mathfrak{g}_\alpha)\alpha(x)\alpha(y). \tag{7.1}$$

Mooreover, $\mathfrak{h}$ *is abelian.*

Proof. The formula for $\kappa(x,y)$ follows immediately from the Cartan Decomposition and the fact that $\mathfrak{g}_\alpha$ has k_α as its only composition factor. We show that $[\mathfrak{h},\mathfrak{h}] = D\mathfrak{h} = 0$. By part (3) of Proposition 7.1.1 it is enough to show that $D\mathfrak{h}$ lies in the radical of $\kappa_{|\mathfrak{h}}$. But for any $\alpha \in \Phi$, α is a one-dimensional representation of $\mathfrak{h}$, so it vanishes on $D\mathfrak{h}$. It is then immediate from Equation (7.1) that $\kappa(x,y) = 0$ for any $x \in D\mathfrak{h}$ and all $y \in \mathfrak{h}$, and so if $x \in D\mathfrak{h}$ then $x \in \mathrm{rad}(\kappa_{|\mathfrak{h}}) = \{0\}$ as claimed.

□

Remark 7.1.3. Since the restriction of κ to $\mathfrak{h}$ is non-degenerate, it yields an isomorphism $\tau\colon \mathfrak{h}^* \to \mathfrak{h}$. Indeed if $\lambda \in \mathfrak{h}^*$ then there is a unique $t_\lambda \in \mathfrak{h}$ such that $\kappa(t_\lambda, y) = \lambda(y)$ for all $y \in \mathfrak{h}$, and the assignment $\lambda \to t_\lambda$ is clearly linear. (See the notes on bilinear forms for more details.)

In the next Proposition need to use Lemma 6.1.6, applied to the adjoint representation. This shows that, if $\mathfrak{h}_\alpha = [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ then for each $\beta \in \Phi$ there is a rational number $r_\beta \in \mathbb{Q}$ such that $\beta_{|\mathfrak{h}_\alpha} = r_\beta \alpha_{|\mathfrak{h}_\alpha}$.

Proposition 7.1.4. *Let* $(\mathfrak{g}, \mathfrak{h})$ *be a pair consisting of a semisimple Lie algebra* $\mathfrak{g}$ *and a Cartan subalgebra* $\mathfrak{h} \subseteq \mathfrak{g}$*, and let* $\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha\in\Phi} \mathfrak{g}_\alpha$ *the associated Cartan decomposition. Write* $\mathfrak{g}_\alpha^s = \{x \in \mathfrak{g}_\alpha : ad(h)(x) = \alpha(h).x\} \subseteq \mathfrak{g}_\alpha$ *for the non-zero subrepresentation of* $\mathfrak{g}_\alpha$ *on which* $\mathfrak{h}$ *acts by* α.

i) If $x \in \mathfrak{g}_\alpha^s$*,* $y \in \mathfrak{g}$ *and* $h \in \mathfrak{h}$ *then*

$$\kappa(h, [x, y]) = \alpha(h)\kappa(x, y),$$

and hence $[x, y] = \kappa(x, y)t_\alpha$.

ii) The set of roots Φ *spans* $\mathfrak{h}^*$.

iii) The subspace $\mathfrak{h}_\alpha^s = [\mathfrak{g}_\alpha^s, \mathfrak{g}_{-\alpha}] \subseteq \mathfrak{h}$ *is one-dimensional, and is spanned by* t_α*. Moreover* $\alpha(\mathfrak{h}_\alpha^s) \neq 0$.

iv) If $\alpha \in \Phi$*, we may find* $e_\alpha \in \mathfrak{g}_\alpha^s$*,* $f_\alpha \in \mathfrak{g}_{-\alpha}$ *and* $h_\alpha \in \mathfrak{h}_\alpha^s$ *so that the map* $e \mapsto e_\alpha$*,* $f \mapsto f_\alpha$ *and* $h \mapsto h_\alpha$ *gives an embedding* $\mathfrak{sl}_2 \to \mathfrak{g}_\alpha \oplus \mathfrak{h}_\alpha \oplus \mathfrak{g}_{-\alpha}$.

Proof. For *i*), let x, y, h be as in the statement of the Proposition. Then by invariance we have:

$$\kappa(h, [x, y]) = \kappa([h, x], y) = \kappa(\alpha(h)x, y) = \alpha(h)\kappa(x, y),$$

where the first equality uses the invariance of κ, and the second follows from the fact that $x \in \mathfrak{g}_\alpha^s$. The identity $\kappa(x, y)t_\alpha = [x, y]$ follows immediately from Proposition 7.1.1 the definitions.

For *ii*), suppose that $\mathfrak{u} = \text{span}\{\Phi\}$. If $\mathfrak{u}$ is a proper subspace of $\mathfrak{h}^*$, then we may find a non-zero $h \in \mathfrak{u}^0 \subseteq \mathfrak{h} \cong \mathfrak{h}^{**}$ such that $\alpha(h) = 0$ for all $\alpha \in \Phi$. But then it follows from (7.1) that $\kappa(h, x) = 0$ for all $x \in \mathfrak{h}$, which contradicts the nondegeneracy of the form $\kappa_{|\mathfrak{h}}$.

For *iii*), as in Remark 7.1.3 above, since $\kappa_{|\mathfrak{h}}$ is nondegenerate it yields an isomorphism $\tau\colon \mathfrak{h}^* \to \mathfrak{h}$, where if $\tau(\lambda) = t_\lambda$ then $\kappa(t_\lambda, h) = \lambda(h)$ for all $h \in \mathfrak{h}$. Since we know that Φ spans $\mathfrak{h}^*$, it follows that $\{t_\alpha : \alpha \in \Phi\}$ spans $\mathfrak{h}$. Suppose that $x \in \mathfrak{g}_\alpha^s$, $y \in \mathfrak{g}_{-\alpha}$. Then by *i*) we see that $[x, y] = \kappa(x, y)t_\alpha$, so that $\mathfrak{h}_\alpha^s \subseteq \text{span}\{t_\alpha\}$. Since κ is nondegenerate on $\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}$, given $x \in \mathfrak{g}_\alpha^s$, we may find $y \in \mathfrak{g}_{-\alpha}$ such that $\kappa(x, y) \neq 0$, hence $\mathfrak{h}_\alpha^s = \text{span}\{t_\alpha\}$ as required. Next we wish to show that $\alpha(\mathfrak{h}_\alpha^s) \neq 0$. For this, note that if we write $\mathfrak{h}_\alpha = [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$, then by Lemma 6.1.6, if $\beta \in \Phi$, then we have $\beta_{|\mathfrak{h}_\alpha} = r_\beta.\alpha_{|\mathfrak{h}_\alpha}$ for some $r_\beta \in \mathbb{Q}$. But then if $\alpha(\mathfrak{h}_\alpha^s) = 0$ it follows $\beta(\mathfrak{h}_\alpha^s) = r_\beta.\alpha(\mathfrak{h}_\alpha^s) = 0$ for all $\beta \in \Phi$, and hence by Equation 7.1 we see that $\mathfrak{h}_\alpha^s \subseteq \text{rad}(\kappa) = \{0\}$, which is impossible, as $t_\alpha \in \mathfrak{h}_\alpha^s$.

□

Lemma 7.1.5. *Let* $\mathfrak{g}$ *be a semisimple Lie algebra and* $\mathfrak{h}$ *a Cartan subalgebra with Cartan decomposition* $\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha\in\Phi} \mathfrak{g}_\alpha$*. Then*

i) The root spaces $\mathfrak{g}_\alpha$ *are one-dimensional, in particular* $\mathfrak{g}_\alpha = \mathfrak{g}_\alpha^s$ *and thus* $\mathfrak{h}$ *acts on* $\mathfrak{g}$ *semisimply (i.e.* $ad(h)$ *acts diagonalisably on* $\mathfrak{g}$ *for all* $h \in \mathfrak{h}$*).*

ii) We have $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = \mathfrak{h}_\alpha = \mathsf{k}.t_\alpha$*, and thus* $\mathfrak{sl}_\alpha = \mathfrak{g}_\alpha \oplus \mathfrak{h}_\alpha \oplus \mathfrak{g}_{-\alpha}$ *is a subalgebra isomorphic to* $\mathfrak{sl}_2(\mathsf{k})$.

iii) If $\alpha \in \Phi$ *and* $c\alpha \in \Phi$ *for some* $c \in \mathbb{Z}$ *then* $c = \pm 1$.

Proof. Choose a nonzero vector $e_\alpha \in \mathfrak{g}_\alpha^s$. Then as in the proof of Proposition 7.1.4 we may find an element $e_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $[e_\alpha, e_{-\alpha}] = h_\alpha \neq 0 \in \mathfrak{h}$. Rescaling e_α if necessary we may assume that $\alpha(h_\alpha) = 2$. Consider the subspace

$$M = \mathsf{k}.e_\alpha \oplus \mathsf{k}.h_\alpha \oplus \bigoplus_{p<0} \mathfrak{g}_{p\alpha}$$

of $\mathfrak{g}$; this is a finite direct sum as $\mathfrak{g}$ is finite-dimensional. Then since $\mathrm{ad}(e_\alpha)(e_\alpha) = 0$, and by *iii)* of Proposition 7.1.4 we know that $[\mathfrak{g}_\alpha^s, \mathfrak{g}_{-\alpha}] = \mathsf{k}.h_\alpha$, and $[h_\alpha, e_\alpha] = 2e_\alpha$, it is easy to see that M is stable under $e_\alpha, e_{-\alpha}$ and h_α. We compute the trace of h_α on M in two ways: on the one hand, it is a commutator and so has trace zero. On the other hand it acts on each of the direct summands defining M so that we may compute

$$\begin{aligned} 0 = \mathrm{tr}(h_\alpha) &= \alpha(h_\alpha) + \sum_{p<0} \dim(\mathfrak{g}_{p\alpha}).p\alpha(h_\alpha) \\ &= \alpha(h_\alpha)\left(1 - \sum_{p>0} p.\dim(\mathfrak{g}_{-p\alpha})\right). \end{aligned}$$

Since we know that $\alpha(h_\alpha) \neq 0$, the only way the above equality can hold is if $\dim(\mathfrak{g}_{-p\alpha}) = 0$ for $p > 1$ and $\dim(\mathfrak{g}_{-\alpha}) = 1$. Since $-\alpha \in \Phi$ if and only if $\alpha \in \Phi$, it follows that $\dim(\mathfrak{g}_\alpha) = 1$ for all $\alpha \in \Phi$, so that in particular $\mathfrak{g}_\alpha^s = \mathfrak{g}_\alpha$. Since we already know $\mathfrak{h}$ is abelian, it follows that $\mathrm{ad}(h) \in \mathfrak{gl}(\mathfrak{g})$ is a semisimple (or diagonalisable) linear map for all $h \in \mathfrak{g}$.

For *ii*), note that by part *i*) since $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ are one-dimensional it is clear that $\mathfrak{sl}_\alpha$ is spanned by $\{e_\alpha, e_{-\alpha}, h_\alpha\}$, and as these satisfy $[h_\alpha, e_\alpha] = 2e_\alpha$, $[h_\alpha, e_{-\alpha}] = -2e_{-\alpha}$ and $[e_\alpha, e_{-\alpha}] = h_\alpha$, it follows immediately that $\mathfrak{sl}_\alpha \cong \mathfrak{sl}_2(\mathsf{k})$.

Finally, for *iii*) note first that since $\Phi \subseteq V\backslash\{0\}$, if $c\alpha \in \Phi$ then $c \neq 0$. We have shown in the proof of *i*) that if $\alpha \in \Phi$ and $p > 0$ then $-p.\alpha \in \Phi$ if and only if $p = 1$. Since $\alpha \in \Phi$ if and only if $-\alpha \in \Phi$, we may apply this to $-\alpha$ also, we find that if $c \in \mathbb{Z}$ and $c\alpha \in \Phi$ then $c = \pm 1$ as claimed. □

Remark 7.1.6. A triple of elements $\{e, f, h\}$ in a Lie algebra $\mathfrak{g}$ which obey the relations of the standard generators of $\mathfrak{sl}_2$ (that is, $[e, f] = h, [h, e] = 2e, [h, f] = 2f$) is called an $\mathfrak{sl}_2$-*triple*. Specifying an $\mathfrak{sl}_2$-triple is equivalent to giving a homomorphism $\varphi\colon \mathfrak{sl}_2 \to \mathfrak{g}$. Note that since $\mathfrak{sl}_2$ is simple, such a homomorphism is injective whenever it is non-zero, and is unique up to a scalar.

In fact the techniques we have already used can be refined somewhat to give a finer information about the set of roots associated to the Cartan decomposition of a semisimple Lie algebra. For this we need the some more terminology:

Definition 7.1.7. Suppose that α, β are two roots in $\mathfrak{g}$. Then we may consider the roots which have the form $\beta + k\alpha$. Clearly, since $\mathfrak{g}$ is finite dimensional, there are integers $p, q > 0$ such that $\beta + k\alpha \in \Phi \cup \{0\}$ for each k with $-p \leq k \leq q$, but neither $\beta - (p+1)\alpha$ nor $\beta + (q+1)\alpha$ are in $\Phi \cup \{0\}$. This set of roots is called[1] the *α-string* through β.

Proposition 7.1.8. *Let $\alpha, \beta \in \Phi$ and suppose that $\beta - p\alpha, \dots, \beta + q\alpha$ is the α-string through β. Then we have*

i)

$$\beta(h_\alpha) = \kappa(h_\alpha, t_\beta) = \frac{2\kappa(t_\alpha, t_\beta)}{\kappa(t_\alpha, t_\alpha)} = p - q.$$

In particular $\beta - \beta(h_\alpha).\alpha \in \Phi$.

ii) If $c \in \mathsf{k}$, then $c\alpha \in \Phi$ if and only if $c = \pm 1$.

[1]Some references will impose the condition that α and β are linearly independent, in which case the α-string through β will be a subset of Φ.

iii) If $S = \{k \in \mathbb{Z} : \beta + k\alpha \in \Phi \cup \{0\}\}$*, then* $S = \{k \in \mathbb{Z} : -p \leq k \leq q\}$.

Proof. We consider the subspace $M = \bigoplus_{-p \leq k \leq q} \mathfrak{g}_{\beta+k\alpha}$, where if $\beta + k\alpha = 0$, we take $\mathfrak{g}_{\beta+k\alpha}$ to be $\mathfrak{h}_\alpha$ (rather than $\mathfrak{h}$). Let $\{e_\alpha, h_\alpha, e_{-\alpha}\}$ be an $\mathfrak{sl}_2$-triple corresponding to an isomorphism $\varphi\colon \mathfrak{sl}_2 \to \mathfrak{sl}_\alpha$ as in Remark 7.1.6. Using Lemma 7.1.5 we see that $\mathfrak{sl}_\alpha$ preserves M, hence

$$\mathrm{tr}_M(\mathrm{ad}(h_\alpha)_{|M}) = \mathrm{tr}_M([\mathrm{ad}(e_\alpha)_{|M}, \mathrm{ad}(e_{-\alpha})_{|M}]) = 0.$$

Thus, using the fact root spaces are 1-dimensional, we obtain the identity:

$$0 = \sum_{-p \leq k \leq q} (\beta + k\alpha)(h_\alpha),$$

and so

$$0 = (q(q+1)/2 - p(p+1)/2)\alpha(h_\alpha) + (p+q+1)\beta(h_\alpha).$$

Hence since $p+q+1 \neq 0$ and $\alpha(h_\alpha) = 2$, it follows that $\beta(h_\alpha) = p-q$ as required. Since $\beta-(p-q)\alpha = (\beta - p\alpha) + q\alpha$ is the q-th term in a string containing $p+q+1$ terms starting with $\beta - p.\alpha$, it is certainly in the α-string through β it follows that $\beta - \beta(h_\alpha).\alpha \in \Phi$.

For *ii*), taking $\beta = c\alpha$ in part *i*) we find that $2c = p - q$. If $p - q \in 2\mathbb{Z}$ then we are done by part *iii*) of Lemma 7.1.5. On the other hand, if $p-q$ is odd, the α-string through $\beta = \frac{(p-q)}{2}\alpha$ has the form:

$$\frac{-(p+q)}{2}\alpha, \dots, \frac{(p-q)}{2}\alpha, \dots, \frac{(p+q)}{2}\alpha,$$

which clearly then contains $\frac{1}{2}\alpha$ so that $\frac{1}{2}\alpha \in \Phi$. But then we get a contradiction as $\alpha = 2(\frac{1}{2}\alpha)$.

Finally, for *iii*), we may replace β by $\beta + m\alpha$ where $m = \min(S)$, so that $S \subseteq \mathbb{Z}_{\geq 0}$ and in particular $p = 0$ and $\beta(h_\alpha) = -q$ for some integer $q \geq 0$. Let $N = \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}_{\beta+k\alpha}$ (where again we take $\mathfrak{g}_{\beta+k\alpha}$ to be $\mathfrak{h}_\alpha$ if $\beta + k\alpha = 0$). If $S \neq \{0, 1, \dots, q\}$, then there is some $t > q+1$ such that $\{t, t+1, \dots, t+r\} \subset S$ but $t + r + 1 \notin S$ (where $r \in \mathbb{Z}_{\geq 0}$). But then considering the α-string through $\beta_1 = \beta + t\alpha$ we find that $\beta_1(h_\alpha) = -r \leq 0$. On the other hand $\beta_1(h_\alpha) = (\beta + t\alpha)(h_\alpha) = 2t - q > t + 1 > 0$, which is a contradiction. □

7.2 $\mathfrak{h}$ and inner product spaces

Recall that since $\kappa_{|\mathfrak{h}}$ is non-degenerate, it gives an isomorphism $\theta\colon \mathfrak{h}^* \to \mathfrak{h}$. For $\lambda \in \mathfrak{h}^*$, we write t_λ for $\theta(\lambda)$, so that $\kappa(t_\lambda, h) = \lambda(h)$, $(\forall \lambda \in \mathfrak{h}^*, h \in \mathfrak{h})$. Given a root $\alpha \in \Phi$, we have seen that $\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ span a subalgebra isomorphic to $\mathfrak{sl}_2$. We will denote this subalgebra as $\mathfrak{sl}_\alpha$. We note the following simple lemma.

Lemma 7.2.1. *Let* $\alpha \in \Phi$ *and let* $h_\alpha \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = \mathfrak{h}_\alpha = \mathsf{k}.t_\alpha$ *be such that* $\alpha(h_\alpha) = 2$*. Then*

i) We have

$$h_\alpha = \frac{2}{\kappa(t_\alpha, t_\alpha)} t_\alpha, \quad \kappa(t_\alpha, t_\alpha).\kappa(h_\alpha, h_\alpha) = 4$$

ii) If $\alpha, \beta \in \Phi$ *then* $\kappa(h_\alpha, h_\beta) \in \mathbb{Z}$ *and* $\kappa(t_\alpha, t_\beta) \in \mathbb{Q}$*.*

Proof. For *i*) note that by *ii*) of Lemma 7.1.5 we have $h_\alpha = c.t_\alpha$ for some $c \in \mathsf{k}$, and $2 = \alpha(h_\alpha) = \kappa(t_\alpha, h_\alpha)$, it follows readily that $c = 2/\kappa(t_\alpha, t_\alpha)$. Next, note that

$$\kappa(t_\alpha, t_\alpha)\kappa(h_\alpha, h_\alpha) = 2\kappa(t_\alpha, h_\alpha) = 2\alpha(h_\alpha) = 4.$$

For part *ii*), using the Cartan decomposition to compute $\kappa(x, y)$ for $x, y \in \mathfrak{h}$ we see that (since we now know that root spaces are one-dimensional) by Proposition 7.1.8:

$$\kappa(h_\alpha, h_\beta) = \sum_{\gamma \in \Phi} \gamma(h_\alpha)\gamma(h_\beta) \in \mathbb{Z}.$$

It thus follows from the identity in part *i*) that $\kappa(t_\alpha, t_\alpha) \in \mathbb{Q}$, and hence $t_\alpha \in \mathbb{Q}.h_\alpha$. Thus $\kappa(t_\alpha, t_\beta) \in \mathbb{Q}$ as required. □

Let $(-,-)$ denote the bilinear form on $\mathfrak{h}^*$ which is obtained by identifying $\mathfrak{h}^*$ with $\mathfrak{h}$: that is

$$(\lambda, \mu) = \kappa(t_\lambda, t_\mu).$$

Clearly it is a nondegenerate symmetric bilinear form, and via the previous Lemma, for all $\alpha, \beta \in \Phi$ we have $(\alpha, \beta) = \kappa(t_\alpha, t_\beta) \in \mathbb{Q}$.

Lemma 7.2.2. *The $\mathbb{Q}$-span of the roots Φ is a $\mathbb{Q}$-vector space of dimension $dim_{\mathsf{k}}(\mathfrak{h}^*)$.*

Proof. We know that Φ spans $\mathfrak{h}^*$, so we may pick a subset $\{\alpha_1, \alpha_2, \dots, \alpha_l\}$ which forms a k-basis of $\mathfrak{h}^*$. To prove the Lemma it is enough to show that every $\beta \in \Phi$ lies in the $\mathbb{Q}$-span of the $\{\alpha_i : 1 \le i \le l\}$. But now if we write $\beta = \sum_{j=1}^{l} c_j \alpha_j$ for $c_j \in \mathsf{k}$, then we see that $(\alpha_i, \beta) = \sum_{i=1}^{l} (\alpha_i, \alpha_j) c_j$. But the matrix $C = (\alpha_i, \alpha_j)_{i,j}$ is invertible since $(-,-)$ is nondegenerate, and its entries are in $\mathbb{Q}$ hence so are those of C^{-1}. But then we have $(c_i) = C^{-1}((\alpha_i, \beta))$, and the objects on the right-hand side all have $\mathbb{Q}$-entries, so we are done. □

Let $\mathfrak{h}^*_{\mathbb{Q}}$ denote the $\mathbb{Q}$-span of the roots. Although you are perhaps more used to inner product spaces over $\mathbb{R}$ or $\mathbb{C}$, the definition of a positive definite symmetric bilinear form makes perfectly good sense over $\mathbb{Q}$. We now show that $(-,-)$ is such an inner product on $\mathfrak{h}^*_{\mathbb{Q}}$.

Proposition 7.2.3. *The form $(-,-)$ is positive definite on $\mathfrak{h}^*_{\mathbb{Q}}$.*

Proof. Using the root space decomposition to compute κ we have

$$(\lambda, \lambda) = \kappa(t_\lambda, t_\lambda) = \sum_{\alpha \in \Phi} \alpha(t_\lambda)^2 = \sum_{\alpha \in \Phi} (\alpha, \lambda)^2 \ge 0,$$

and since we may have equality if and only if $(\alpha, \lambda) = 0$ for all $\alpha \in \Phi$, and the elements of Φ span $\mathfrak{h}^*$ it follows that the form is definite as required. □

7.3 Abstract root systems

In this section we study the geometry which we are led to by the configuration of roots associated to a Cartan decomposition of a semisimple Lie algebra. These configurations will turn out to have a very special, highly symmetric, form which allows them to be completely classified.

We will work with an inner product space, that is a vector space equipped with a positive definite symmetric bilinear form $(-,-)$. Such a form makes sense over any field which has a notion of positive elements, and so in particular over $\mathbb{R}$ and $\mathbb{Q}$, but *not* over arbitrary characteristic zero fields[2]. Since the roots Φ associated to a Cartan decomposition of a semisimple Lie algebra naturally live in the $\mathbb{Q}$-inner product space $\mathfrak{h}^*_{\mathbb{Q}}$, we will assume our field is $\mathbb{Q}$ unless otherwise stated. We let $\mathrm{O}(V)$ denote the group of orthogonal linear transformations of V, that is the linear transformations which preserve the inner product, so that $g \in \mathrm{O}(V)$ precisely when $v, w \in V$ then $(v, w) = (g(v), g(w))$ for all $v, w \in V$.

Definition 7.3.1. A *reflection* is a nontrivial element of $\mathrm{O}(V)$ which fixes a subspace of codimension 1 (*i.e.* dimension $\dim(V) - 1$). If $s \in \mathrm{O}(V)$ is a reflection and $W < V$ is the $+1$-eigenspace, then $L = W^\perp$ is a line preserved by s, hence the restriction $s_{|L}$ of s to L is an element of $O(L) = \{\pm 1\}$, which since s is nontrivial must be -1. In particular s has order 2. If v is any nonzero element of L then it is easy to check that s is given by

$$s(u) = u - \frac{2(u, v)}{(v, v)} v.$$

Given $v \neq 0$ we will write s_v for the reflection given by the above formula, and refer to it as the "reflection in the hyperplane perpendicular to v".

[2] Inner product spaces have, in addition to a notion of distance, a notion of angle. Their geometry is thus pretty much that of Prelims Geometry I. Positive-definite forms do not make sense over characteristic zero fields such as $\mathbb{C}$, although in the case of $\mathbb{C}$ one can use Hermitian forms as a replacement.

We now give the definition which captures the geometry of the root of a semisimple Lie algebra.

Definition 7.3.2. Let V be a $\mathbb{Q}$-vector space equipped with an inner product $(-,-)$. A finite subset $\Phi \subset V\backslash\{0\}$ is called a *root system* if it satisfies the following properties:

i) Φ spans V;

ii) If $\alpha \in \Phi$ then $c\alpha \in \Phi$ if and only if $c = \pm 1$;

iii) If $\alpha \in \Phi$ then $s_\alpha \colon V \to V$ preserves Φ;

iv) If $\alpha, \beta \in \Phi$ and we define
$$\langle \alpha, \beta \rangle = \frac{2(\alpha, \beta)}{(\alpha, \alpha)}, \tag{7.2}$$
then $\langle \alpha, \beta \rangle \in \mathbb{Z}$. We say $\langle \alpha, \beta \rangle$ is a *Cartan integer*.

This definition is, unsurprisingly, motivated by the following result.

Lemma 7.3.3. *Let $(\mathfrak{g}, \mathfrak{h})$ be a pair consisting of a semisimple Lie algebra $\mathfrak{g}$ together with a Cartan subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$. Let $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha$ be the associated Cartan decomposition and let $\mathfrak{h}^*_{\mathbb{Q}}$ be the $\mathbb{Q}$-span of Φ in $\mathfrak{h}^*$. Then $(\mathfrak{h}^*_{\mathbb{Q}}, \Phi)$ is an abstract root system.*

Proof. Let $(-,-)$ denotes the symmetric bilinear form on $\mathfrak{h}^*$ induced by the restriction of the Killing form $\kappa_{|\mathfrak{h}}$ as in §7.2. The results of that section show that $(-,-)$ restricts to an inner product on $\mathfrak{h}^*_{\mathbb{Q}}$.

Property *i*) for an abstract root system follows immediately from the definitions, and the remaining properties follow from Proposition 7.1.8: Property *ii*) of that Proposition establishes property *ii*), while part *i*) establishes properties *iii*) and *iv*).

□

Remarkably, the finite set of vectors given by a root system has both a rich enough structure that it captures the isomorphism type of a semisimple Lie algebra, but is also explicit enough that we can completely classify them, and hence classify semisimple Lie algebras.

Definition 7.3.4. Let (V, Φ) be a root system. Then the *Weyl group* of the root system is the group $W = \langle s_\alpha : \alpha \in \Phi \rangle$. Since its generators preserve the finite set Φ and these vectors span V, it follows that it is a finite subgroup of $\mathrm{O}(V)$.

Example 7.3.5. Let $\mathfrak{g} = \mathfrak{sl}_n$, Then let $\mathfrak{d}_n$ denote the diagonal matrices in $\mathfrak{gl}_n$ and $\mathfrak{h}$ the (traceless) diagonal matrices in $\mathfrak{sl}_n$. As you saw in the problem sets, $\mathfrak{h}$ forms a Cartan subalgebra in $\mathfrak{sl}_n$. Let $\{\varepsilon_i : 1 \leq i \leq n\}$ be the basis of $\mathfrak{d}_n^*$ dual to the basis $\{E_{ii} : 1 \leq i \leq n\}$ of $\mathfrak{d}_n$ in $\mathfrak{gl}_n$. Then $\mathfrak{h}^*_{\mathbb{Q}}$ is the quotient space
$$\mathfrak{h}^*_{\mathbb{Q}} = \left\{ \sum_{i=1}^n c_i \varepsilon_i : c_i \in \mathbb{Q} \right\} \Big/ \{\mathbb{Q}.(\varepsilon_1 + \ldots + \varepsilon_n)\},$$
the roots in $\mathfrak{h}^*_{\mathbb{Q}}$ are the (images of the) vectors $\{\varepsilon_i - \varepsilon_j : 1 \leq i, j \leq n, i \neq j\}$. The Weyl group W in this case is the group generated by the reflections s_α which, for $\alpha = \varepsilon_i - \varepsilon_j$ interchange the basis vectors ε_i and ε_j, so it is easy to see that W is just the symmetric group on n letters.

The first crucial point about the geometry of root systems is that the angles between roots are highly constrained.

Lemma 7.3.6. *Let (V, Φ) be a root system and let $\alpha, \beta \in \Phi$ be such that $\alpha \neq \pm\beta$. Then $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle \in \{0, 1, 2, 3\}$. It follows that the angle between two such roots α, β lies in the set*
$$\{\pi/2, \pi/3, 2\pi/3, \pi/4, 3\pi/4, \pi/6, 5\pi/6\}.$$
Moreover, the ratios of root lengths which are not perpendicular must be $1, 2, 1/2, 3$ or $1/3$.

Proof. By assumption, we know that both $\langle \alpha, \beta \rangle$ and $\langle \beta, \alpha \rangle$ are integers. On the other hand, by the cosine formula (*i.e.* by Cauchy-Schwarz) we see that if θ denotes the angle between α and β, then:

$$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4\cos(\theta)^2 < 4. \tag{7.3}$$

Since $\cos(\theta)^2$ determines the angle between the two vectors (or rather the one which is less than π) and $\langle \beta, \alpha \rangle / \langle \alpha, \beta \rangle = ||\alpha||^2/||\beta||^2$ (where we write $||v||^2 = (v, v)$), the rest of the Lemma follows by a case-by-case check as we see from the following table:

$\langle \alpha, \beta \rangle$	$\langle \beta, \alpha \rangle$	θ	$\lVert\alpha\rVert^2/\lVert\beta\rVert^2$
0	0	$\pi/2$	undetermined
1	1	$\pi/3$	1
-1	-1	$2\pi/3$	1
1	2	$\pi/4$	2
-1	-2	$3\pi/4$	2
1	3	$\pi/6$	3
-1	-3	$5\pi/6$	3

□

7.3.1 Root systems and bases Since the set of roots Φ spans V, it certainly contains (many) subsets which form a basis of V. The key to the classification of root systems is to show that there is a special class of such bases which capture enough of the geometry of the set of roots that the entire root system can be recovered from the bases of this form. Although it is not immediate that this should be the correct definition, the distinguishing property of these bases is a kind of positivity property:

Definition 7.3.7. Given a set of vectors X in a vector space V, we will write

$$\mathbb{N}.X = \left\{ \sum_{s \in J} c_s.s : J \subseteq X \text{ finite}, c_s \in \mathbb{N} \right\}.$$

The set $\mathbb{N}.X$ is closed under vector addition and multiplication by elements of $\mathbb{N}$.

Definition 7.3.8. Let (V, Φ) be a root system, and let Δ be a subset of Φ. We say that Δ is a *base* (or a set of *simple roots*) for Φ if Δ is a basis of V and for each $\alpha \in \Phi$, exactly one of α or $-\alpha$ lies in $\mathbb{N}.\Delta$.

Given a base of Φ we may declare a root positive or negative according to the sign of the nonzero coefficients which occur when we write it in terms of the base. We write Φ^+ for the set of positive roots and Φ^- for the set of negative roots.

The obvious drawback with the above definition at first sight, is that it is not at all obvious that a root system necessarily contains a base. In order to show that this is indeed the case however, we use the following:

Definition 7.3.9. Let V be a $\mathbb{Q}$-vector space. A *positive set* $\mathcal{P}$ in V is a subset $\mathcal{P} \subseteq V \backslash \{0\}$ such that

- For each $v \in V \backslash \{0\}$, exactly one of v or $-v$ lies in $\mathcal{P}$.
- if $v_1, v_2 \in \mathcal{P}$ then $v_1 + v_2 \in \mathcal{P}$.
- If $\lambda \in \mathbb{Q}$ satisfies $\lambda > 0$ and $v \in \mathcal{P}$ then $\lambda.v \in \mathcal{P}$.

If $\mathcal{P}$ is a positive set, then we define a total order $<$ on V by $v_1 < v_2$ if and only if $v_2 - v_1 \in \mathcal{P}$.

Example 7.3.10. If V is a $\mathbb{Q}$-vector space and $B = \{e_1, \dots, e_n\}$ is an ordered basis of V, then it is easy to check that

$$\mathcal{P}_B = \left\{ \sum_{i=k}^{n} \lambda_i e_i : k \in \{1, 2, \dots, n\}, \lambda_k > 0 \right\}$$

is a positive system. Note that, with respect to the total order $\mathcal{P}$ defines, $e_1 > e_2 > \dots > e_n$.

Definition 7.3.11. Let (V, Φ) be a root system, and fix a positive set $\mathcal{P}$ in V. Let $\Phi^+_{\mathcal{P}} = \Phi \cap \mathcal{P}$ and $\Phi^-_{\mathcal{P}} = \Phi \cap (-\mathcal{P})$ (where the positive set $\mathcal{P}$ is understood from context, we will simply write Φ^+). We say that $\alpha \in \Phi^+$ is *decomposable* if $\alpha = \beta + \gamma$ for some $\beta, \gamma \in \Phi^+$. A root is *indecomposable* if it is not decomposable. Let $\Delta = \Delta_{\mathcal{P}} \subseteq \Phi^+$ be the set of indecomposable roots in Φ^+.

Remark 7.3.12. If we assume that $\Delta \subseteq \Phi$ is a base, then picking an arbitrary ordering $\Pi = \{\alpha_1, \dots, \alpha_l\}$ say, we obtain an ordered basis of V and hence a positive set $\mathcal{P}_\Delta$ as in Example 7.3.10. It is easy to see that in this situation we must have $\Delta = \Delta_{\mathcal{P}}$. Thus if bases exist, it must be the case that they arise as the set of indecomposable roots corresponding to the choice of a positive set in V.

Lemma 7.3.13. *Let $\Delta \subseteq \Phi^+$ be the set of indecomposable roots as above. Then for all $\alpha, \beta \in \Delta$ we have $(\alpha, \beta) \leq 0$.*

Proof. Suppose that $(\alpha, \beta) > 0$. Then $\langle \alpha, \beta \rangle > 0$. Since $\alpha, \beta \in \Phi^+$, they are not linearly dependent, and hence $\langle \alpha, \beta \rangle . \langle \beta, \alpha \rangle \in \{1, 2, 3\}$. It follows one of $\langle \alpha, \beta \rangle$ or $\langle \beta, \alpha \rangle = 1$. By symmetry we may assume $\langle \alpha, \beta \rangle = 1$, and hence $s_\alpha(\beta) = \beta - \alpha \in \Phi$. But then one of $\alpha - \beta$ or $\beta - \alpha$ lies in Φ^+. But as $\alpha = (\alpha - \beta) + \beta$ and $\beta = (\beta - \alpha) + \alpha$, one of α or β is decomposable, which is a contradiction. □

Lemma 7.3.14. *Suppose that V is an inner product space equipped with a positive set $\mathcal{P}$. If $S \subset \mathcal{P}$ is such that $(s_1, s_2) \leq 0$ for all $s_1 \neq s_2$ in S, then S is linearly independent.*

Proof. Suppose that $\sum_{s \in T} c_s s = 0$ is a linear dependence between elements of S, where $T \subseteq S$ is some finite subset. Then let $Q = \{s \in S : c_s > 0\}$ and set $z = \sum_{s \in Q} c_s s = \sum_{t \in T \setminus Q} |c_t| t$. Now as $T \subseteq \mathcal{P}$, the sum $\sum_{s \in Q} c_s s$ equals zero if and only if all the coefficients c_s vanish. Similarly $\sum_{t \in T \setminus Q} |c_t| t = 0$ if and only if all the c_t vanish. Thus it suffices to show $z = 0$. But we have

$$0 \leq (z, z) = \left(\sum_{s \in Q} c_s s, \sum_{t \in T \setminus Q} |c_t| t \right) = \sum_{\substack{s \in Q \\ t \in T \setminus Q}} c_s |c_t| (s, t) \leq 0.$$

since each term $c_s |c_t| (s, t) \leq 0$. By positive definiteness, it follows $z = 0$ as required. □

Proposition 7.3.15. *Let (V, Φ) be a root system and suppose that $\mathcal{P}$ is a positive system. Then the indecomposable roots Δ in Φ^+ form a base of (V, Φ). In particular, any root system has a base.*

Proof. Lemma 7.3.13 and 7.3.14 show that Δ is linear independent, so it only remains to check that $\Phi^+ \subseteq \mathbb{N}.\Delta$ (since Φ spans V and $\Phi = \Phi^+ \sqcup -\Phi^+$, this clearly implies that Δ spans V). To see this, suppose for the sake of a contradiction that it is not the case, and take $\alpha \in \Phi^+$ minimal with respect to the total order defined by $\mathcal{P}$ such that $\alpha \notin \mathbb{N}.\Delta$. Since Δ clearly lies in $\mathbb{N}.\Delta$, it follows that α is decomposable, and hence $\alpha = \beta + \gamma$. But then by the minimality of α, we must have $\beta, \gamma \in \mathbb{N}\Delta$, and hence $\alpha = \beta + \gamma \in \mathbb{N}.\Delta$, which is a contradiction. □

It turns out that we can recover the entire root system provided we know a base for it. Before we can show this, we first show that any two bases of Φ are conjugate under the action of W.

Lemma 7.3.16. *Let Δ be a base of (V, Φ) and let Φ^+ be the corresponding set of positive roots. Then if $\alpha \in \Delta$ and $\beta \in \Phi^+$ where $\beta \neq \alpha$, then $s_\alpha(\beta) \in \Phi^+$.*

Proof. Let $\Delta = \{\alpha_1, \dots, \alpha_l\}$ where $\alpha = \alpha_1$. Now if $\beta = \sum_{r=1}^{l} n_r\alpha_r$ where $n_r \geq 0$, and $\beta \neq \alpha_1$, hence there must be some $t > 1$ with $n_t > 0$. But the

$$s_\alpha(\beta) = \beta - \langle \alpha, \beta \rangle \alpha = (n_1 - \langle \alpha_1, \beta \rangle)\alpha_1 + \sum_{r\geq 2} n_r\alpha_r,$$

and since Δ is a base and $n_t > 0$, it follows that $s_\alpha(\beta) \in \Phi^+$ as required. □

With this Lemma in hand, we can now show that the Weyl group acts transitively on the set of possible bases of the root system. In fact we will show more: Let us fix $\Delta_0 = \{\gamma_1, \dots, \gamma_l\}$ a particular (ordered) base, let $\mathcal{P}_0$ be the corresponding positive set, and Φ_0^+ the associated set of positive roots. Let W_0 be the subgroup of the Weyl group W generated by the reflections $\{s_\gamma : \gamma \in \Delta_0\}$.

Proposition 7.3.17. *Suppose that Δ_1 is any base of (V, Φ) and let Φ_1^+ be the corresponding set of positive roots (and $\mathcal{P}_1$ the associated positive set). Then there is some $w \in W_0$ such that $w(\Phi_1^+) = \Phi_0^+$, and hence $w(\Delta_1) = \Delta_0$.*

Proof. We prove this by induction on $d = |\Phi_0^+ \cap \Phi_1^-|$. If this $d = 0$ then $\Phi_1^+ = \Phi_2^+$ and hence $\Delta_0 = \Delta_1$ (hence we may take $w = e$ the identity element of the Weyl group). Next suppose that $d > 1$. Now if $\Delta_0 \subseteq \Phi_1^+ = \Phi \cap \mathcal{P}_1$, then since any element of Φ_0^+ is a positive integer combination of Δ_0, it follows $\Phi_0^+ \subseteq \mathcal{P}_1^+$ and hence $\Phi_0^+ \cap \Phi_1^- = \emptyset$, which contradicts the assumption that $d > 0$. Thus there is some $\alpha \in \Delta_0$ such that $\alpha \in \Phi_2^-$. But then by the previous Lemma we have

$$s_\alpha(\Phi_0^+) \cap \Phi_1^- = (\{-\alpha\} \cup (\Phi_0^+ \backslash \{\alpha\}) \cap \Phi_1^- = (\Phi_0^+ \backslash \{\alpha\}) \cap \Phi_1^-,$$

which shows that $s_\alpha(\Phi_0^+) \cap \Phi_1^-$ has $d-1$ elements, and hence so does $\Phi_0^+ \cap s_\alpha(\Phi_1^-)$ (as s_α is an involution). But then by induction there is a $w \in W_0$ with $ws_\alpha(\Phi_1)^+ = \Phi_0^+$. Since $ws_\alpha \in W_0$ we are done. □

Proposition 7.3.18. *Suppose that $\beta \in \Phi$. Then there is a $w \in W_0$ and an $\alpha \in \Delta_0$ such that $w(\beta) = \alpha$.*

Proof. By Proposition 7.3.17, it suffices to show that there is a base of (V, Φ) which contains β. Pick a basis $B \subset \Phi$ of V which contains β, and then if we write $B = \{\alpha_1, \alpha_2, \dots, \alpha_l\}$ where $\beta = \alpha_l$, and let $\mathcal{P}_B$ denote the positive set for V associated with it as in Example 7.3.10. Let $\Phi^+ = \Phi \cap \mathcal{P}_B$ and $\Delta_B \subseteq \Phi^+$ the set of indecomposable roots in Φ^+, a base of (V, Φ).

We claim that $\beta = \alpha_l \in \Delta_B$. Indeed suppose that $\Delta_B = \{\gamma_i : 1 \leq i \leq l\}$ and let $\alpha = \sum_{i=1}^{l} n_i\gamma_i$, where $n_i \in \mathbb{Z}_{\geq 0}$. Then if $1 \leq k \leq n$ has $n_k > 0$ we must have $\gamma_k \leq \beta = \alpha_l$. But if $\gamma_i = \sum_{r=s}^{l} c_r\alpha_r$ for $1 \leq s \leq l$, since $\gamma \in \Phi^+$, $c_s > 0$, hence if $\alpha_l - \gamma_i > 0$ we must have $s = l$. But then $\gamma_i = c_l\alpha_l$, and as $c_l\alpha_l \in \Phi$ if and only if $c = \pm 1$. Since γ_i, α_l are both positive, it follows that $\gamma_i = \alpha_l \in \Delta_B$ as required. □

Corollary 7.3.19. *The Weyl group W is generated by the reflections $\{s_\gamma : \gamma \in \Delta_0\}$, that is $W = W_0$.*

Proof. If $\beta \in \Phi$ then we have just shown in the previous proposition that there is a $w \in W_1$ such that $w(\beta) = \gamma$ for some $\gamma \in \Delta_0$. But the clearly $s_\beta = w^{-1}s_\gamma w \in W_0$, and so since W is generated by the s_βs we have $W = W_0$ as required. □

Remark 7.3.20. In fact W acts *simply transitively* on the bases of (V, Φ), that is, the action is transitive and, if Δ is a base and $w \in W$ is such that $w(\Delta) = \Delta$, then $w = 1$. The proof (which we will not give) consists of examining the minimal length expression for w in terms of these generators $\{s_\alpha : \alpha \in \Delta_0\}$.

7.3.2 Cartan matrices and Dynkin diagrams In this section we describe the data which is used in the classification of semisimple Lie algebras.

Definition 7.3.21. Let (V, Φ) be a root system. The *Cartan matrix* associated to (V, Φ) is the matrix

$$C = (\langle \alpha_i, \alpha_j \rangle)_{i,j=1}^{l}.$$

where $\{\alpha_1, \alpha_2, \ldots, \alpha_\ell\} = \Delta$ is a base of (V, Φ). Since the elements of W are isometries, and W acts transitively on bases of Φ, the Cartan matrix is independent of the choice of base (though clearly determined only up to reordering the base Δ).

Definition 7.3.22. The entries c_{ij} of the Cartan matrix are all integer with diagonal entries equal to 2, and off-diagonal entries $c_{ij} \in \{0, -1, -2, -3\}$ (where $i \neq j$) such that if $c_{ij} < -1$ then $c_{ji} = -1$ so that the pair $\{c_{ij}, c_{ji}\}$ is determined by the product $c_{ij}.c_{ji}$ and the relative lengths of the two roots (*e.g.* see the table in the Lemma about angles between roots). As a result, the matrix can be recorded as a kind of graph: the vertex set of the graph is labelled by the base $\{\alpha_1, \ldots, \alpha_l\}$, and one puts $\langle \alpha_i, \alpha_j \rangle . \langle \alpha_j, \alpha_i \rangle$ edges between α_i and α_j, directing the edges so that they go from the larger root to the smaller root. Thus for example if $\langle \alpha_i, \alpha_j \rangle = -2$ and $\langle \alpha_j, \alpha_i \rangle = -1$ so that $\|\alpha_j\|^2 > \|\alpha_i\|^2$, that is, α_j is longer than α_i, we record this in the graph as:

$$\overset{\alpha_i}{\bullet} \Longleftarrow \overset{\alpha_j}{\bullet}$$

The resulting graph is called the *Dynkin diagram*.

For the next theorem we need to formulate what it means to have an isomorphism of root systems. This is given in the natural way: if (V, Φ) and (V', Φ') are root systems, a linear map $\phi\colon V \to V'$ is an isomorphism of root systems if

1. The map ϕ is an isomorphism of vector spaces.
2. $\phi(\Phi) = \Phi'$, and $\langle \alpha, \beta \rangle = \langle \phi(\alpha), \phi(\beta) \rangle$ for all $\alpha, \beta \in \Phi$.

Note that ϕ need not be an isometry (*e.g.* we could scale V by a nonzero constant $c \in \mathbb{Q}$ to obtain $(V, c\Phi)$ a distinct, but isomorphic root system to (V, Φ).

Theorem 7.3.23. *Let (V, Φ) be a root system. Then (V, Φ) is determined up to isomorphism by the Cartan matrix, or Dynkin diagram associated to it.*

Proof. Given root systems (V, Φ) and (V', Φ') with the same Cartan matrix, we may certainly pick a base $\Delta = \{\alpha_1, \ldots, \alpha_\ell\}$ of (V, Φ) and a base $\Delta' = \{\beta_1, \ldots, \beta_\ell\}$ of (V', Φ') such that $\langle \alpha_i, \alpha_j \rangle = \langle \beta_i, \beta_j \rangle$ for all i, j, $(1 \leq i, j \leq \ell)$. We claim the map $\phi\colon \Delta \to \Delta'$ given by $\phi(\alpha_i) = \beta_i$ extends to an isomorphism of root systems. Clearly, since Δ and Δ' are bases of V and V' respectively, ϕ extends uniquely to an isomorphism of vector spaces $\phi\colon V \to V'$, so we must show that $\phi(\Phi) = \Phi'$, and $\langle \phi(\alpha), \phi(\beta) \rangle = \langle \alpha, \beta \rangle$ for each $\alpha, \beta \in \Phi$.

Let $s_i = s_{\alpha_i} \in \mathrm{O}(V)$ and $s_i' = s_{\beta_i} \in \mathrm{O}(V')$ be the reflections in the Weyl groups W and W' respectively. Then from the formula for the action of s_i it is clear that $\phi(s_i(\alpha_j)) = s_i'(\beta_j) = s_i'(\phi(\alpha_i))$, so since Δ is a basis it follows $\phi(s_i(v)) = s_i'(\phi(v))$ for all $v \in V$. But then since the s_is and s_i's generate W and W' respectively, ϕ induces an isomorphism $W \to W'$, given by $w \mapsto w' = \phi \circ w \circ \phi^{-1}$. But then given any $\alpha \in \Phi$ we know there is a $w \in W$ such that $\alpha = w(\alpha_j)$ for some j, $(1 \leq j \leq \ell)$. Thus we have $\phi(\alpha) = \phi(w(\alpha_j)) = w'(\phi(\alpha_j)) = w'(\beta_j) \in \Phi'$, so that $\phi(\Phi) \subseteq \Phi'$. Clearly the same argument applied to ϕ^{-1} shows that $\phi^{-1}(\Phi') \subseteq \Phi$ so that $\phi(\Phi) = \Phi'$.

Finally, note that it is clear from the linearity of ϕ and of $\langle \alpha, \gamma \rangle$ in the second variable, that $\langle \alpha, \gamma \rangle = \langle \phi(\alpha), \phi(\gamma) \rangle$ for all $\alpha \in \Delta$, $\gamma \in \Phi$. In the same fashion as above however, if $\alpha \in \Phi$ is arbitrary, then we may find $w \in W$ such that $\alpha = w(\alpha_j) \in \Delta$, and thus $\phi(\alpha) = w'(\beta_j)$, whence we have

$$\begin{aligned} \langle \phi(\alpha), \phi(\gamma) \rangle &= \langle w'(\beta_j), \phi(\gamma) \rangle = \langle \beta_j, (w')^{-1}\phi(\gamma) \rangle \\ &= \langle \alpha_j, w^{-1}(\gamma) \rangle = \langle w(\alpha_j), \gamma \rangle = \langle \alpha, \gamma \rangle. \end{aligned}$$

as required. $\square$

Thus to classify root systems up to isomorphism it is enough to classify Cartan matrices (or Dynkin diagrams).

Definition 7.3.24. We say that a root system (V, Φ) is *reducible* if there is a partition of the roots into two non-empty subsets $\Phi_1 \sqcup \Phi_2$ such that $(\alpha, \beta) = 0$ for all $\alpha \in \Phi_1, \beta \in \Phi_2$. Then if we set $V_1 = \text{span}(\Phi_1)$ and $V_2 = \text{span}(\Phi_2)$, clearly $V = V_1 \oplus V_2$ and we say (V, Φ) is the sum of the root systems (V_1, Φ_1) and (V_2, Φ_2). This allows one to reduce the classification of root systems to the classification of *irreducible* root systems, *i,e.* root systems which are not reducible. It is straightforward to check that a root system is irreducible if and only if its associated Dynkin diagram is connected.

Definition 7.3.25. (*Not examinable.*) The notion of a root system makes sense over the real, as well as rational, numbers. Let (V, Φ) be a real root system, and let $\Delta = \{\alpha_1, \alpha_2, \ldots, \alpha_l\}$ be a base of Φ. If $v_i = \alpha_i/||\alpha_i||$ $(1 \leq i \leq l)$ are the unit vectors in V corresponding to Δ, then they satisfy the conditions:

1. $(v_i, v_i) = 1$ for all i and $(v_i, v_j) \leq 0$ if $i \neq j$,
2. If $i \neq j$ then $4(v_i, v_j)^2 \in \{0, 1, 2, 3\}$. (This is the reason we need to extend scalars to the real numbers – if you want you could just extend scalars to $\mathbb{Q}(\sqrt{2}, \sqrt{3})$, but it makes no difference to the classification problem).

Such a set of vectors is called an *admissible set*.

It is straightforward to see that classifying $\mathbb{Q}$-vector spaces with a basis which forms an admissible set is equivalent to classifying Cartan matrices, and using elementary techniques it is possible to show that that the following are the only possibilities (we list the Dynkin diagram, a description of the roots, and a choice of a base):

- Type A_ℓ $(\ell \geq 1)$:

$$\bullet \!-\!\!-\!\!-\! \bullet \cdots \bullet \!-\!\!-\!\!-\! \bullet$$

$$V = \{v = \sum_{i=1}^{\ell} c_i e_i \in \mathbb{Q}^\ell : \sum c_i = 0\}, \Phi = \{\varepsilon_i - e_j : 1 \leq i \neq j \leq \ell\}$$
$$\Delta = \{\varepsilon_{i+1} - \varepsilon_i : 1 \leq i \leq \ell - 1\}$$

- Type B_ℓ $(\ell \geq 2)$:

$$\bullet \!-\!\!-\!\!-\! \bullet \cdots \bullet \Longrightarrow \bullet$$

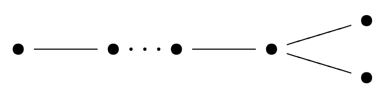
$$V = \mathbb{Q}^\ell, \Phi = \{\pm\varepsilon_i \pm \varepsilon_j : 1 \leq i, j \leq \ell, i \neq j\} \cup \{\varepsilon_i : 1 \leq i \leq \ell\},$$
$$\Delta = \{\varepsilon_1, \varepsilon_{i+1} - \varepsilon_i : 1 \leq i \leq \ell - 1\}$$

- Type C_ℓ $(\ell \geq 3)$:

$$\bullet \!-\!\!-\!\!-\! \bullet \cdots \bullet \Longleftarrow \bullet$$

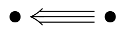
$$V = \mathbb{Q}^\ell, \Phi = \{\pm\varepsilon_i \pm \varepsilon_j : 1 \leq i, j \leq \ell, i \neq j\} \cup \{2\varepsilon_i : 1 \leq i \leq \ell\},$$
$$\Delta = \{2\varepsilon_1, \varepsilon_{i+1} - \varepsilon_i : 1 \leq i \leq \ell - 1\}$$

- Type D_ℓ $(\ell \geq 4)$:



$$V = \mathbb{Q}^\ell, \Phi = \{\pm\varepsilon_i \pm \varepsilon_j : 1 \leq i, j \leq \ell, i \neq j\},$$
$$\Delta = \{\varepsilon_1 + \varepsilon_2, \varepsilon_{i+1} - \varepsilon_i : 1 \leq i \leq \ell - 1\}$$

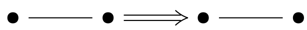
- Type G_2.



Let $e = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 \in \mathbb{Q}^3$, then:

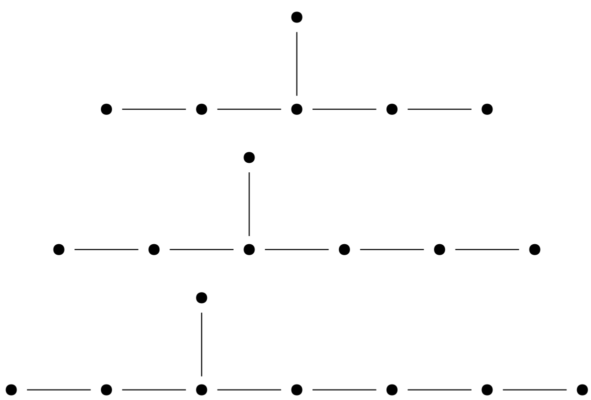
$$V = \{v \in \mathbb{Q}^3 : (v, e) = 0\}, \Phi = \{\varepsilon_i - \varepsilon_j : i \neq j\} \cup \{\pm(3\varepsilon_i - e) : 1 \leq i \leq 3\}$$
$$\Delta = \{\varepsilon_1 - \varepsilon_2, e - 3\varepsilon_1\}$$

- Type F_4:



$$V = \mathbb{Q}^4,$$
$$\Phi = \{\pm\varepsilon_i : 1 \leq i \leq 4\} \cup \{\pm\varepsilon_i \pm e_j : i \neq j\} \cup \{\frac{1}{2}(\pm\varepsilon_1 \pm \varepsilon_2 \pm \varepsilon_3 \pm \varepsilon_4)\}$$
$$\Delta = \{\varepsilon_2 - \varepsilon_3, \varepsilon_3 - \varepsilon_4, \varepsilon_4, \frac{1}{2}(\varepsilon_1 - \varepsilon_2 - \varepsilon_3 - \varepsilon_4)\}.$$

- Type E_n ($n = 6, 7, 8$).



These can all be constructed inside E_8 by taking the span of the appropriate subset of a base, so we just give the root system for E_8.

$$V = \mathbb{Q}^8, \Phi = \{\pm\varepsilon_i \pm \varepsilon_j : i \neq j\} \cup \{\frac{1}{2}\sum_{i=1}^{8}(-1)^{a_i}\varepsilon_i : \sum_{i=1}^{8} a_i \in 2\mathbb{Z}\},$$
$$\Delta = \{\varepsilon_1 + \varepsilon_2, \varepsilon_{i+1} - \varepsilon_i, \frac{1}{2}(\varepsilon_1 + \varepsilon_8 - (\varepsilon_2 + \varepsilon_3 + \ldots + \varepsilon_7)) : 1 \leq i \leq 6\}.$$

Note that the Weyl groups of type B_ℓ and C_ℓ are equal. The reason for the restriction on ℓ in the types B,C,D is to avoid repetition, *e.g.* B_2 and C_2 are the same up to relabelling the vertices.

Remark 7.3.26. I certainly don't expect you to remember the root systems of the exceptional types, but you should be familiar with the ones for type A, B, C and D. The ones of rank two (*i.e.* A_2, B_2 and G_2) are also worth knowing (because for example you can draw them!)

7.4 The Classification of Semisimple Lie algebras

Only the statements of the theorems in this section are examinable, but it is important to know these statements!

Remarkably, the classification of semisimple Lie algebras is identical to the classification of root systems: each semisimple Lie algebra decomposes into a direct sum of simple Lie algebras, and it is not hard to show that the root system of a simple Lie algebra is irreducible. Thus to any simple Lie algebra we may attach an irreducible root system.

A first problem with this as a classification strategy is that we don't know our association of a root system to a semisimple Lie algebra is canonical. The difficulty is that, because our procedure for attaching a root system to a semisimple Lie algebra involves a choice of Cartan subalgebra, we don't currently know it is a bijective correspondence – possibly the same Lie algebra has two different Cartan subalgebras which lead to different root systems. The theorem which ensures this is not the case is the following, where the first part is the more substantial result (though both require some work):

Theorem 7.4.1. *Let $\mathfrak{g}$ be a Lie algebra over any algebraically closed field* k.

1. *Let $\mathfrak{h}, \mathfrak{h}'$ be Cartan subalgebras of $\mathfrak{g}$. There is an automorphism $\phi\colon \mathfrak{g} \to \mathfrak{g}$ such that $\phi(\mathfrak{h}) = \mathfrak{h}'$.*
2. *Let $\mathfrak{g}_1, \mathfrak{g}_2$ be semisimple Lie algebras with Cartan subalgebras $\mathfrak{h}_1, \mathfrak{h}_2$ respectively, and suppose now k is of characteristic zero. Then if the root systems attached to $(\mathfrak{g}_1, \mathfrak{h}_1)$ and $(\mathfrak{g}_2, \mathfrak{h}_2)$ are isomorphic, there is an isomorphism $\phi\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ taking $\mathfrak{h}_1$ to $\mathfrak{h}_2$.*

Once you know that the assignment of a Dynkin diagram captures a simple Lie algebra up to isomorphism, we still need to show all the root systems we construct arise as the root system of a simple Lie algebra. That is exactly the content of the next theorem.

Theorem 7.4.2. *There exists a simple Lie algebra corresponding to each irreducible root system.*

There are a number of approaches to this existence theorem. A concrete strategy goes as follows: one can show that the first four infinite families A, B, C, D correspond to the classical Lie algebras, $\mathfrak{sl}_{\ell+1}$, $\mathfrak{so}_{2\ell+1}$, $\mathfrak{sp}_{2\ell}$, $\mathfrak{so}_{2\ell}$, whose root systems can be computed directly (indeed you did a number of these calculations in the problem sets). This of course also requires checking that these Lie algebras are simple (or at least semisimple) but this is also straight-forward with the theory we have developed. It then only remains to construct the five "exceptional" simple Lie algebras. This can be done in a variety of ways – given a root system where all the roots are of the same length there is an explicit construction of the associated Lie algebra by forming a basis from the Cartan decomposition (and a choice of base of the root system) and explicitly constructing the Lie bracket by giving the structure constants with respect to this basis (which, remarkably, can be chosen for the basis vectors corresponding to the root subspaces to lie in $\{0, \pm 1\}$). This gives in particular a construction of the Lie algebras of type E_6, E_7, E_8 (and also A_ℓ and D_ℓ though we already had a construction of these). The remaining Lie algebras can be found by a technique called "folding" which studies automorphisms of simple Lie algebras, and realises the Lie algebras G_2 and F_4 as fixed-points of an automorphism of D_4 and E_6 respectively.

There is also an alternative, more *a posteriori* approach to the uniqueness result which avoids showing Cartan subalgebras are all conjugate for a general Lie algebra: one can check that for a classical Lie algebra $\mathfrak{g} \subset \mathfrak{gl}_n$ as above, the Cartan subalgebras are all conjugate by an element of $\mathrm{Aut}(\mathfrak{g})$ (in fact you can show the automorphism is induced by conjugating with a matrix in $\mathrm{GL}_n(\mathsf{k})$) using the fact that a Cartan subalgebra of a semisimple Lie algebra is abelian and consists of semisimple elements. This then shows the assignment of a root system to a classical Lie algebra is unique, so it only remains to check the exceptional Lie algebras. But these all have different dimensions, and the dimension of the Lie algebra is captured by the root system, so we are done.[3]

We conclude by mentioning another, quite different, approach to the existence result, using the *Serre's presentation*: just as one can describe a group by generators and relations, one can also describe Lie algebras in a similar fashion. If $\mathfrak{g}$ is a semisimple Lie algebra and $\Delta = \{\alpha_1, \dots, \alpha_\ell\}$ is a base of the corresponding root system with Cartan matrix $C = (a_{ij})$ then picking bases for the $\mathfrak{sl}_{\alpha_i}$-subalgebras corresponding to them, it is not too hard to show that $\mathfrak{g}$ is generated by the set $\{e_\alpha, f_\alpha, h_\alpha : \alpha \in \Delta\}$.

The Serre presentation gives an explicit realisation, given an arbitrary root system, of the relations which one needs to impose on a set of generators for a Lie algebra labelled $\{e_\alpha, f_\alpha, h_\alpha : \alpha \in \Phi\}$

[3]This is completely rigorous, but feels like cheating (to me).

as above obtain a semisimple Lie algebra whose associated root system is the one we started with. This approach has the advantage of giving a uniform approach, though it takes some time to develop the required machinery.

Appendices

I (Multi)-linear algebra

I.1 Primary Decomposition

Definition I.1. Let k be an algebraically closed field. If V is a k-vector space and $x \in \mathrm{End}_{\mathsf{k}}(V)$ and $\lambda \in \mathsf{k}$, the *generalized eigenspace* for x with eigenvalue λ is

$$V_\lambda = \{v \in V : \exists n \geq 0, (x-\lambda)^n(v) = 0\},$$

Thus $V_\lambda \neq \{0\}$ if and only if x has an eigenvector $v \in V \backslash \{0\}$ with eigenvalue λ. The subspace V_λ are clearly invariant under the action of x, that is $x(V_\lambda) \subseteq V_\lambda$.

We say that x is *nilpotent* is $V = V_0$, or equivalently, if $x^n = 0$ for some $n \geq 0$. We say that x is *semisimple* (or sometimes *diagonalisable* if each V_λ is in fact an eigenspace, that is, for all $v \in V_\lambda$, we have $(x-\lambda)(v) = 0$.

The following basic result is proved in Part A Linear Algebra. We provide a proof for the sake of convenience.

Proposition I.2. *Let $x \colon V \to V$ be a linear map. There is a canonical direct sum decomposition*

$$V = \bigoplus_{\lambda \in \mathsf{k}} V_\lambda,$$

of V into the generalized eigenspaces of x. Moreover, for each λ, the projection to $a_\lambda \colon V \to V_\lambda$ (with kernel the remaining generalized eigenspace of x) can be written as a polynomial in x.

Proof. Let $m_x \in \mathsf{k}[t]$ be the minimal polynomial of x. Then if $\phi \colon \mathsf{k}[t] \to \mathrm{End}(V)$ given by $t \mapsto x$ denotes the natural map, we have $\mathsf{k}[t]/(m_x) \cong \mathrm{im}(\phi) \subseteq \mathrm{End}(V)$. If $m_x = \prod_{i=1}^k (t-\lambda_i)^{n_i}$ where the λ_i are the distinct eigenvalues of x, then the Chinese Remainder Theorem and the first isomorphism theorem shows that

$$\mathrm{im}(\phi) \cong \mathsf{k}[t]/(m_x) \cong \bigoplus_{i=1}^{k} \mathsf{k}[t]/(t-\lambda_i)^{n_i},$$

It follows that we may write $1 \in \mathsf{k}[t]/(m_x)$ as $1 = e_1 + \ldots + e_k$ according to the above decomposition. Now clearly $e_i e_j = 0$ if $i \neq j$ and $e_i^2 = e_i$, so that if $U_i = \mathrm{im}(e_i)$, then we have $V = \bigoplus_{1 \leq i \leq k} U_i$. Moreover, each e_i can be written as polynomials in x by picking any representative in $\mathsf{k}[t]$ of e_i (thought of as an element of $\mathsf{k}[t]/(m_x)$). Note in particular this means that each U_i is invariant under $\mathrm{im}(\phi)$.

It thus remains to check that $U_i = V_{\lambda_i}$. Now the characteristic polynomial of $x_{|V_{\lambda_i}}$ is clearly just $(t-\lambda_i)^{d_i}$ where $d_i = \dim(V_{\lambda_i})$, and evidently this divides $\chi_x(t)$ the characteristic polynomial of $x \in \mathrm{End}(V)$. But if $n_i = \dim(U_i)$ it is immediate from the definitions that $\chi_x(t) = \prod_{i=1}^k (t-\lambda_i)^{n_i}$, and hence $d_i \leq n_i$. Since $U_i \subseteq V_{\lambda_i}$ we also have $n_i \leq d_i$, and hence they must be equal, so $V_{\lambda_i} = U_i$ as required.

□

Corollary I.3. *Let $x\colon V \to V$ be a linear map. Then there exists a diagonalisable linear map x_s and a nilpotent linear map x_n such that $x = x_s + x_n$ and $[x_s, x_n] = 0$.*

Proof. Let $V = \bigoplus_{\lambda \in \mathsf{k}} V_\lambda$ be the generalised eigenspace decomposition of V given by the action of x. Suppose that $\{\lambda_1, \dots, \lambda_k\}$ are the distinct eigenvalues of x, and let $(e_i)_{i=1}^k$ be the projection maps to V_{λ_i}. Then if $x_s = \sum_{i=1}^k \lambda_i.e_i$, clearly x_s is semisimple, and $[x, x_s] = 0$ (since this is evident on each V_{λ_k}). Setting $x_n = x - x_s$, and noting that on each V_{λ_i} the map $x - x_s$ is equal to $x - \lambda_i$, which is nilpotent, we conclude that x_n is nilpotent as required. □

I.2 Tensor Products

I.2.1 Definition and construction : Tensor products were studied in Part B, Introduction to Representation Theory. We review their basic properties here.

Definition I.4. If $V_1, V_2, \dots, V_k$ and U are vector spaces over a field k, let

$$\mathcal{M}(V_1, \dots V_k, U) = \{\theta\colon V_1 \times \dots \times V_k \to U : \theta \text{ is } k\text{-linear}\}$$

be the vector space of all k-(multi-)linear maps on $V_1 \times \dots \times V_k$ taking values in a vector space U. Here we say that a function $\theta\colon V_1 \times V_2 \times \dots \times V_k \to U$ is a k-linear if it is linear in each component separately, that is, if for any k-tuples of vectors $(v_i)_{1 \le i \le k}, (u_j)_{1 \le j \le k} \in V_1 \times \dots \times V_k$ and any $\lambda \in \mathsf{k}$, we have for each $i \in \{1, 2, \dots, k\}$,

$$\theta(v_1, \dots, \lambda.v_i + u_i, \dots v_k) = \lambda.\theta(v_1, \dots, v_i, \dots v_k) + \theta(v_1, \dots, u_i, \dots v_k),$$

Pick a basis B_i of V_i for each i ($1 \le i \le k$), and let B_i^* denote the corresponding dual basis of V_i^*. If $b \in B_i$, let δ_b denote the corresponding element of the dual basis B_i^*, so that $B_i^* = \{\delta_b : b \in B_i\}$. Let $\mathbf{B} = B_1 \times B_2 \times \dots B_k$

Proposition I.5. *In the notation given above, the restriction to $\mathbf{B}$ gives an isomorphism*

$$r_{\mathbf{B}}\colon \mathcal{M}(V_1, \dots, V_k; U) \to U^{\mathbf{B}} = \{f : \mathbf{B} \to U\}$$

from the space of all k-multilinear maps taking values in U to the space of all U-valued functions on $\mathbf{B}$. Indeed $r_{\mathbf{B}}$ has inverse given explicitly by

$$\mathcal{F}_{\mathbf{B}}(f)(v_1, \dots, v_k) = \sum_{\mathbf{b} = (b_1, \dots, b_k) \in \mathbf{B}} \delta_{b_1}(v_1) \dots \delta_{b_k}(v_k) f(\mathbf{b}).$$

Proof. Note that if we pick $\mathbf{b} = (b_1, \dots, b_k) \in \mathbf{B}$ then the product $\delta_{\mathbf{b}} = \delta_{b_1}.\delta_{b_2} \dots \delta_{b_k}$ is a k-linear map (since multiplication distributes over addition). Since it is easy to see that $\delta_{\mathbf{b}}(\mathbf{b}') = \delta_{\mathbf{b},\mathbf{b}'}$ (that is, is zero unless $\mathbf{b} = \mathbf{b}'$ in which case it is equal to 1), it is immediate that $r_{\mathbf{B}}(\mathcal{F}_{\mathbf{B}}(f)) = f$, so we must show that $\mathcal{F}_{\mathbf{B}}(r_{\mathbf{B}}(\theta)) = \theta$ for any $\theta \in \mathcal{M}(V_1, \dots, V_k; U)$. Explicitly, we must show that

$$\theta = \sum_{\mathbf{b} \in B_1 \times \dots \times B_k} \delta_{\mathbf{b}} \theta(\mathbf{b}) \tag{I.4}$$

Indeed applying θ to a k-tuple $\mathbf{b} \in B_1 \times \dots B_k$, we see that the coefficients on the right-hand side are uniquely determined, so it remains to show the products $\delta_{\mathbf{b}}$ of dual basis vectors do indeed span.

The case $k = 1$ is simply the standard argument that the functions $\{\delta_{b_1}\}_{b_1 \in B_1}$ are indeed a basis of V_1^*: if $v_1 \in V_1$ then we may write $v_1 = \sum_{b_1 \in B_1} \lambda_{b_1} b_1$ for unique scalars $\lambda_{b_1} \in \mathsf{k}$. By the definition of the functions δ_{b_1}, it then follows that $\delta_{b_1}(v_1) = \lambda_{b_1}$, so that $v_1 = \sum_{b_1 \in B_1} \delta_{b_1}(v_1).b_1$. Applying θ gives $\theta(v_1) = \sum_{b_1 \in B_1} \delta_{b_1}(v_1).\theta(b_1)$. But as this holds for all $v_1 \in V_1$, it follows that $\theta = \sum_{b_1 \in B_1} \theta(b_1).\delta_{b_1}$, as required.

The general case then follows by an easy induction: Indeed for any k-tuple of vectors $(v_i)_{1\leq i\leq k}$ with $v_i \in V_i$, using the case $k = 1$, we may write $v_1 = \sum_{b_1\in B_1} \delta_{b_1}(v_1).b_1$. But then if θ is k-linear we have

$$\theta(v_1,\dots,v_k) = \theta(\sum_{b_1\in B_1} \delta_{b_1}(v_1).b_1, v_2,\dots,v_k) = \sum_{b_1\in B_1} \delta_{b_1}(v_1).\theta(b_1,v_2,\dots,v_k).$$

But for each $b_1 \in B_1$, the map $(v_i)_{2\leq i\leq k} \mapsto \theta(b_1, v_2,\dots,v_k)$ is a $(k-1)$-linear map from $V_2 \times \dots V_k$ to k, hence the result follows by induction. □

Remark I.6. Note that, for $k = 1$, this says that a linear map is uniquely determined by its values on a basis of V_1, and the statement should be thought of as saying that a k-linear map is similarly determined "by its values on bases" where the statement of the question gives the precise meaning to the vague phrase in quotation marks.

The previous Proposition gives one way of constructing the tensor product: If V and W are k-vector spaces and we pick bases B_V and B_W of V and W respectively, then by the Proposition, if we set $B = B_V \times B_W$, then for any vector space U, we have

$$\mathcal{M}(V,W;U) \cong U^B \cong \mathrm{Hom}_{\mathsf{k}}(S(B),U), \tag{I.5}$$

where $S(B)$ is the vector space with basis B, that is, the space of finite formal linear combinations of elements of B. The first isomorphism above is a direct consequence of the Proposition where we take $k = 2$ and $V_1 = V, V_2 = W$, while the second is the case $k = 1$ of the proposition with $V_1 = S(B)$. Now taking $U = S(B)$ in (I.5), the identity linear map from $S(B)$ to itself corresponds to a bilinear map $t\colon V \times W \to S(B)$. It is easy to see that if $\theta\colon V \times W \to U$ is bilinear, then since $\theta_{|B}\colon B \to U$ extends to a linear map $\tilde{\theta}\colon S(B) \to U$, and $\tilde{\theta} = \tilde{\theta} circ \mathrm{id}_{S(B)}$, so that $\theta = \tilde{\theta} \circ t$ as required.

Remark I.7. Note that there is a natural isomorphism $\sigma\colon V \otimes W \cong W \otimes V$ given by $v \otimes w \mapsto w \otimes v$, thus at least if $V \neq W$, we will normally abuse notation and identify these two spaces and thus write $V \otimes W = W \otimes V$. If $V = W$ however, $\sigma\colon V \otimes V \to V \otimes V$ is an involution on $V \otimes V$, and more generally, $V^{\otimes n} = V \otimes \dots \otimes V$, the tensor product of V with itself n times, has an action of S_n the symmetric group, which permutes the tensor factors: if $\tau \in S_n$ then $\tau(v_1 \otimes \dots \otimes v_n) := v_{\tau(1)} \otimes \dots \otimes \tau_{\tau(n)}$.

Example I.8. If $V = \mathsf{k}$ and W is an arbitrary k-vector space , then if $s\colon \mathsf{k} \times W \to W$ is scalar multiplication map given by $s(\lambda, w) = \lambda.w$, it is clearly bilinear and it is straight-forward to check that it has the universal property so that $\mathsf{k} \otimes W \cong W$.

Lemma I.9. *Let V and W be vector spaces. There is a natural injective map* $\theta\colon V^* \otimes W \to Hom_{\mathsf{k}}(V,W)$ *which is an isomorphism when V is finite-dimensional. Moreover, if* $\iota\colon V^* \otimes V \to \mathsf{k}$ *is the contraction map given by* $f \otimes v \mapsto f(v)$*, then if V is finite dimensional, and* $\alpha \in Hom_{\mathsf{k}}(V,V)$*, then* $(\iota \circ \theta^{-1})(\alpha) = tr(\alpha)$.

Proof. The map $(\alpha, w) \mapsto [v \mapsto \alpha(v).w]$ is bilinear[4], and so yields a linear map $\theta\colon V^* \otimes W \to \mathrm{Hom}_{\mathsf{k}}(V,W)$. To see that it is injective, let $\{\delta_i : i \in I\}$ be a basis of V^*, and $\{f_k : k \in K\}$ be a basis of W. Then if $\gamma \in V^* \otimes W$, by definition we may write $\gamma = \sum_{(i,k)\in S} \lambda_{i,k}\delta_i \otimes f_k$, where the pairs (i,k) run over a finite subset S of $I \times K$. Now if we fix $k \in K$ we have

$$\sum_{i\in I:(i,k)\in S} \lambda_{i,k}\delta_i \otimes f_k = \left(\sum_{i\in I:(i,k)\in S} \lambda_{i,k}\delta_i\right) \otimes f_k,$$

thus setting $\phi_k = \sum_{i\in I:(i,k)\in S} \lambda_{i,k}\delta_i$ it follows $\gamma = \sum_{k\in S_K} \phi_k \otimes f_k$, where $S_K = \{k \in K : \exists i \in I, (i,k) \in S\}$. But then for any $v \in V$

$$0 = \theta(\gamma)(v) = \sum_{k\in S_K} \phi_k(v).f_k,$$

[4]There is a lot of linearity going on here! The map $(\alpha, w, v) \mapsto \alpha(v).w$ is linear in all of α, v and w. For fixed α, w, this shows that the map $v \mapsto \alpha(v).w$ is a linear map from V to W, while the linearity in α and w show the map which sends a pair (α, w) to the corresponding map from V to W is bilinear in α and w.

and so by the linear independence of the f_ks we must have $\phi_k(v) = 0$ for each k. Since this is true for all $v \in V$, it follows that $\phi_k = 0$, for each k, and hence $\gamma = 0$ as required.

To see that θ is an isomorphism when V is finite dimensional, note that in that case we can assume our basis of V^* is dual to a basis $\{e_i : i \in I\}$ of V. But then if $\alpha \in \mathrm{Hom}_\mathsf{k}(V, W)$ it follows that $\alpha = \theta(\sum_{i\in I} \delta_i \otimes \alpha(e_i))$, as the two sides agree on the basis $\{e_i : i \in I\}$.

Finally we consider the contraction map $\iota\colon V^* \times V \to \mathsf{k}$. Since the map $V^* \times V \to \mathsf{k}$ given by $(f, v) \mapsto \mathsf{k}$ is clearly bilinear, it induces the linear map ι, so ι is certainly well-defined. To compute $\iota \circ \theta^{-1}$, note that when $V = W$ is finite dimensional, we can chose the basis $\{\delta_1, \dots, \delta_n\}$ of V^* to be dual to the basis $\{e_1, \dots, e_n\}$ of V, and since, as before $\theta^{-1}(\alpha) = \sum_{i=1}^n \delta_i.\alpha(e_i)$, it follows that $\iota(\theta^{-1}(\alpha)) = \sum_{i=1}^n \delta_i(\alpha(e_i)) = \mathrm{tr}(\alpha)$. as required. □

Remark I.10. Since we only use the cases where V and W are finite dimensional, the reader is welcome to ignore the generality the result is stated in and assume throughout that all vector spaces are finite dimensional. Here one can be a bit more concrete: if $\{e_1, \dots, e_n\}$ is a basis of V and $\{f_1, \dots, f_m\}$ is a basis of W, then taking the dual basis $\{\delta_1, \dots, \delta_n\}$ of V^* it is easy to see that the images of $\delta_i \otimes f_j$ under θ correspond to the elementary matrices E_{ij} under the identification of $\mathrm{Hom}_\mathsf{k}(V, W)$ given by the choice of bases for V and W, hence θ is an isomorphism. In general the image of θ is precisely the linear maps from V to W which have finite rank (as you can readily deduce from the proof of Lemma I.9). Indeed when V is infinite-dimensional, the trace map on $\mathrm{Hom}(V, V)$ is only defined for linear maps of finite rank, thus in a sense, then contraction map ι is more natural than the trace map. (We will return to this point when discussing bilinear forms.)

I.2.2 Linear maps between tensor products. Let $\alpha\colon V_1 \to V_2$ and $\beta\colon W_1 \to W_2$ be linear maps. If $v \in V_1, w \in W_1$, the map $(v, w) \mapsto \alpha(v) \otimes \beta(w)$ from $V_1 \times W_2 \to V_2 \otimes W_2$ is bilinear, and so induces a linear map $\mathrm{Hom}(V_1 \otimes W_1, V_2 \otimes W_2)$, which we denote by $\alpha \otimes \beta$. In fact, the map $(\alpha, \beta) \mapsto \alpha \otimes \beta$ is itself bilinear, and so we even obtain a map

$$\mathrm{Hom}(V_1, W_1) \otimes \mathrm{Hom}(V_2, W_2) \to \mathrm{Hom}(V_1 \otimes V_2, W_1 \otimes W_2). \tag{I.6}$$

Moreover, it follows immediately from the definitions that (I.6) also respects composition. In more detail, if $\alpha_2\colon V_2 \to V_3$ and $\beta_2\colon W_2 \to W_3$ are linear maps to any vector spaces V_3 and W_3, then $(\alpha_2 \otimes \beta_2) \circ (\alpha_1 \otimes \beta_1) = (\alpha_2 \circ \alpha_1) \otimes (\beta_2 \circ \beta_1)$. Indeed, if $v \in V_1, w \in W_1$, then

$$\begin{aligned}(\alpha_2 \otimes \beta_2) \circ (\alpha_1 \otimes \beta_1)(v \otimes w) &= (\alpha_2 \otimes \beta_2)(\alpha_1(v) \otimes \beta_1(w))\\ &= (\alpha_2 \circ \alpha_1)(v) \otimes (\beta_2 \circ \beta_1)(w)\\ &= (\alpha_2 \circ \alpha_1) \otimes (\beta_2 \circ \beta_1)(v \otimes w).\end{aligned}$$

When all the vector spaces V_1, V_2, W_1, W_2 are finite dimensional, the map (I.6) is actually an isomorphism, indeed using Lemma I.9 you can check that

$$\begin{aligned}\mathrm{Hom}(V_1, W_1) \otimes \mathrm{Hom}(V_2, W_2) &\cong (V_1^* \otimes W_1) \otimes (V_2^* \otimes W_2)\\ &\cong (V_1^* \otimes V_2^*) \otimes (W_1 \otimes W_2)\\ &\cong (V_1 \otimes V_2)^* \otimes (W_1 \otimes W_2)\\ &\cong \mathrm{Hom}(V_1 \otimes V_2, W_1 \otimes W_2),\end{aligned}$$

where the second isomorphism simply permutes the second and third tensor factors.

Example I.11. The map $\iota\colon V^* \otimes V \to \mathsf{k}$ also describes the composition of linear maps: Suppose we have three vector spaces V, W and U. The composition gives a bilinear map from $\mathrm{Hom}(U, V) \times \mathrm{Hom}(V, W)$ to $\mathrm{Hom}(U, W)$, thus it is equivalent to a linear map $\tilde{m}\colon \mathrm{Hom}(U, V) \otimes \mathrm{Hom}(V, W) \to \mathrm{Hom}(U, W)$.

$$\begin{tikzcd}
\mathrm{Hom}(U,V)\otimes\mathrm{Hom}(V,W) \arrow[r] & (U^*\otimes V)\otimes(V^*\otimes W) \arrow[dl] & \\
U^*\otimes(V\otimes V^*)\otimes W \arrow[r] & U^*\otimes\mathsf{k}\otimes W \arrow[r] & U^*\otimes W
\end{tikzcd}$$

where the first arrow is the induced by the isomorphism between $\mathrm{Hom}(U,V)$ (resp. Hom(V,W)) and $U^*\otimes V$ (resp. $V^*\otimes W$), the second from the associativity of tensor products, and the third arrow is $1_V\otimes\iota\otimes 1_{V^*}$. By Example I.8 scalar multiplication gives a natural isomorphism $s\colon \mathsf{k}\otimes W\to W$, and the final arrow is just $\mathrm{id}_{U^*}\otimes s$. Identifying the term $U^*\otimes W$ with $\mathrm{Hom}(U,W)$ this becomes the composition of linear maps.

Remark I.12. It is sometimes useful to have the following notational convention: Given a tensor product of more than two vector spaces, such as $U^*\otimes V\otimes V^*\otimes W$, then it can be convenient to write ι_{32} for the map which acts via ι on the third and second factors (that is swapping the second and third factors, applying ι and the repeating the swap) and by the identity on the remaining tensor factors.

I.2.3 Tensor products and duality Suppose that V and W are finite dimensional vector spaces. We wish to understand the relationship between the tensor product of the dual spaces $V^*\otimes W^*$ and the dual space of the tensor product $(V\otimes W)^*$. The map ι gives one way to do this: Let d be the composition

$$(V\otimes W)\otimes(V^*\otimes W^*)\cong(V\otimes V^*)\otimes(W^*\otimes W)\to\mathsf{k}\otimes\mathsf{k}\to\mathsf{k} \tag{I.7}$$

where the first isomorphism permutes the middle two factors, the second is, in the notation of Remark I.12, the map $\iota_{21}\otimes\iota_{34}$ and the final isomorphism follows from the fact that the map $v\otimes 1\to v$ gives a natural isomorphism from $V\otimes\mathsf{k}$ to V for any k-vector space V. Now the linear map d can be views as a k-valued bilinear pairing between $V\otimes W$ and $V^*\otimes W^*$, which in turn can be viewed as a linear map from $d\colon V^*\otimes W^*\to(V\otimes W)^*$. Viewed this way, d is just the multiplication map: if $\delta\in V^*, \eta\in W^*$ and $v\in V$, $w\in W$, we have

$$d(\delta\otimes\eta)(v\otimes w)=\delta(v).\eta(w),$$

that is, $d(\delta\otimes\eta)$ is the element of $(V\otimes W)^*$ which corresponds to the bilinear map $V\times W\to\mathsf{k}$ given by the product $\delta.\eta$. Moreover, again by permuting the factors, one can view d as a map $\mathrm{Hom}(V,W)\otimes\mathrm{Hom}(W,V)\to\mathsf{k}$. Viewed as a bilinear pairing, this is just the trace form.

$$(a,b)\mapsto\mathrm{tr}(ab),\quad \forall a\in\mathrm{Hom}(V,W), b\in\mathrm{Hom}(W,V)$$

This description of the trace form also makes the symmetry property $\mathrm{tr}(ab)=\mathrm{tr}(ba)$ is evident.

Remark I.13. Note that, if V is finite dimensional, then we have $\mathrm{Hom}(V,V)=V^*\otimes V$, and

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

so that $\mathrm{Hom}(V,V)$ is canonically isomorphic to its dual. This is equivalent to the fact that the trace pairing $b_V\colon\mathfrak{gl}(V)\otimes\mathfrak{gl}(V)\to\mathfrak{gl}(V)$ is non-degenerate. Viewed in this way it is easy to see that the duality map interchanges the identity $I_V\in\mathfrak{gl}(V)$ with the trace map $\mathrm{tr}\in\mathfrak{gl}(V)^*$.

I.3 Symmetric bilinear forms In this section we review the basics of symmetric bilinear forms over a field k. It is all material that is almost in Part A Algebra, but perhaps not quite phrased there the way we use it. We shall work to begin with over an arbitrary field k.

Definition I.14. Let V be a k-vector space. A function $B\colon V\times V\to\mathsf{k}$ is said to be bilinear if it is linear in each factor, that is if

$$B(\lambda_1v_2+\lambda_2v_2,w)=\lambda_1B(v_1,w)+\lambda_2B(v_2,w),$$

and

$$B(v, \lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 B(v, w_1) + \lambda_2 B(v, w_2),$$

for all $v, v_1, v_2, w, w_1, w_2 \in V$, $\lambda_1, \lambda_2 \in \mathsf{k}$. We say that it is *symmetric* if $B(v,w) = B(w,v)$. Let $\mathrm{Bil}(V)$ denote the vector space of bilinear forms on V, and[5] $\mathrm{SBil}(V)$ for the space of symmetric bilinear forms on V.

Remark I.15. If V is finite dimensional, then we have canonical isomorphisms

$$V^* \otimes V^* \cong (V \otimes V)^* \cong \mathrm{Bil}(V).$$

Combining with the isomorphism $\theta \colon V^* \otimes V^* \cong \mathrm{Hom}(V, V^*)$ from Lemma I.9 we obtain an isomorphism $\Theta \colon \mathrm{Bil}(V) \to \mathrm{Hom}(V, V^*)$. It follows that giving a bilinear form on V is equivalent to giving a linear map from V to V^*. Note that the map $\sigma \colon V^* \otimes V^* \to V^* \otimes V^*$ given by $f_1 \otimes f_2 \mapsto f_2 \otimes f_1$ gives a second isomorphism $\Theta_1 \colon \mathrm{Bil}(V) \to \mathrm{Hom}(V^*, V)$, where $\Theta_1 = \Theta \circ \sigma$. For symmetric bilinear forms the two maps agree, but for arbitrary bilinear forms they yield different isomorphisms.

Definition I.16. Given a bilinear form B, we set

$$\mathrm{rad}(B) = \mathrm{rad}_L(B) = \{v \in V : \Theta(B)(v) = 0\} = \{v \in V : B(v,w) = 0, \forall w \in V\}$$

(here the subscript "L" denotes "left"). Similarly, we set

$$\mathrm{rad}_R(B) = \ker(\Theta_1(B)) = \{v \in V : B(w,v) = 0, \forall w \in V\}.$$

If B is symmetric $\mathrm{rad}_L(B) = \mathrm{rad}_R(B)$, but this need not be true otherwise. We say that B is *nondegenerate* if $\mathrm{rad}_L(B) = \{0\}$. Note that, even though in general $\mathrm{rad}_L(B) \neq \mathrm{rad}_R(B)$, it is still the case that $\mathrm{rad}_L(B) = \{0\}$ if and only if $\mathrm{rad}_R(B) = \{0\}$.

From now on we will only work with symmetric bilinear forms. Fix $B \in \mathrm{SBil}(V)$. Then if U is a subspace of V, we define

$$U^\perp = \{v \in V : B(v,w) = 0, \forall w \in U\} = \{v \in V : \Theta(B)(v) \in U^0\}.$$

When B is nondegenerate, so that $\Theta(B)$ is an isomorphism, this shows that $\dim(U^\perp) = \dim(U^0) = \dim(V) - \dim(U)$. The next Lemma shows that this can be refined slightly.

Lemma I.17. *Let V be a finite-dimensional k-vector space equipped with a symmetric bilinear form B. Then for any subspace U of V we have the following:*

i) $\dim(U) + \dim(U^\perp) \geq \dim(V)$.

ii) The restriction of B to U is nondegenerate if and only if $V = U \oplus U^\perp$.

Proof. Let $\phi \colon V \to U^*$ be given by $\phi(v)(u) = B(v,u)$, that is $\phi(v) = (\Theta(B)(v))_{|U}$. Clearly $\ker(\phi) = U^\perp$, while $\mathrm{im}(\phi) \leq U^*$ and hence $\dim(\mathrm{im}(\phi)) \leq \dim(U)$. The inequality in *i*) now follows from rank-nullity.

For the second part, note that B is non-degenerate on U if and only if $U \cap U^\perp = \{0\}$. But then the inequality in *i*) shows that we must have $U \oplus U^\perp = V$ for dimension reasons. □

[5] This is not standard notation – it would be more normal to write something like $\mathrm{Sym}^2(V^*)$ but then I'd have to explain why...

I.3.1 *Classification of symmetric bilinear forms This subsection is not needed for the course[6] but might be clarifying. There is a natual linear action of $\mathrm{GL}(V)$ on the space $\mathrm{Bil}(V)$: if $g \in \mathrm{GL}(V)$ and $B \in \mathrm{Bil}(V)$ then we set $g(B)$ to be the bilinear form given by

$$g(B)(v,w) = B(g^{-1}(v), g^{-1}(w)), \quad (v,w \in V),$$

where the inverses ensure that the above equation defines a left action. It is clear the action preserves the subspace of symmetric bilinear forms.

Since we can find a invertible map taking any basis of a vector space to any other basis, the next lemma says that over an algebraically closed field there is only one nondegenerate symmetric bilinear form up to the action of $\mathrm{GL}(V)$, that is, when k is algebraically closed the nondegenerate symmetric bilinear forms are a single orbit for the action of $\mathrm{GL}(V)$.

Lemma I.18. *Let V be a k-vector space equipped with a nondegenerate symmetric bilinear form B. Then if $char(\mathsf{k}) \neq 2$, there is an orthonormal basis of V, i.e a basis $\{v_1, \dots, v_n\}$ of V such that $B(v_i, v_j) = \delta_{ij}$.*

Proof. We use induction on $\dim(V)$. The identity[7]

$$B(v,w) = \frac{1}{2}(B(v+w, v+w) - B(v,v) - B(w,w)),$$

shows that if $B \neq 0$ we may find a vector $v \in V$ such that $B(v,v) \neq 0$. Rescaling by a choice of square root of $B(v,v)$ (which is possible since k is algebraically closed) we may assume that $B(v,v) = 1$. But if $L = \mathsf{k}.v$ then since $B_{|L}$ is nondegenerate, the previous lemma shows that $V = L \oplus L^{\perp}$, and if B is nondegenerate on V it must also be so on $L^{\perp}$. But $\dim(L^{\perp}) = \dim(V) - 1$, and so $L^{\perp}$ has an orthonormal basis $\{v_1, \dots, v_{n-1}$. Setting $v = v_n$, it then follows $\{v_1, \dots, v_n\}$ is an orthonormal basis of V as required. □

Remark I.19. Over the real numbers, for example, there is more than one orbit of nondegenerate symmetric bilinear form, but the above proof can be modified to give a classification and it turns out that there are $\dim(V) + 1$ orbits ("Sylvester's law of inertia").

[6]So in particular you don't need to know it for any exam...

[7]Note that this identity holds unless $\mathrm{char}(\mathsf{k}) = 2$. It might be useful to remember this identity when understanding the Proposition which is the key to the proof of the Cartan Criterion: it claims that if $\mathfrak{g} = D\mathfrak{g}$ then there is an element $x \in \mathfrak{g}$ with $\kappa(x,x) \neq 0$. Noting the above identity, we see this is equivalent to asserting that κ is nonzero.

II Reminder on Representation theory

We recall here some basics of representation theory used in the course, all of which is covered (in much more detail than we need) in the Part B course on Representation theory. Let $\mathfrak{g}$ be a Lie algebra. The main body of the notes proves all that is needed in the course, but the material here might help clarify some arguments. We will always assume our representations are finite dimensional unless we explicitly say otherwise.

II.1 Basic notions

Definition II.1. A representation is *irreducible* if it has no proper nonzero subrepresentations. A representation (V, ρ) is said to be *indecomposable* if it cannot be written as a direct sum of two proper subrepresentations. A representation is said to be *completely reducible* if is a direct sum of irreducible representations.

Clearly an irreducible representation is indecomposable, but the converse is not in general true. For example k^2 is naturally a representation for the nilpotent Lie algebra of strictly upper triangular matrices $\mathfrak{n}_2 \subset \mathfrak{gl}_2(\mathsf{k})$ and it is not hard to see that it has a unique 1-dimensional sub representation, hence it is indecomposable, but not irreducible.

A basic observation about irreducible representations is Schur's Lemma:

Lemma II.2. *Let $\mathfrak{g}$ be a Lie algebra and let (V, ρ), (W, σ) be irreducible representations of $\mathfrak{g}$. Then any $\mathfrak{g}$-homomorphism $\phi\colon V \to W$ is either zero or an isomorphism. In particular, if k is algebraically closed, then $Hom_{\mathfrak{g}}(V, W)$ is one-dimensional.*

Proof. The proof is exactly the same as the proof for finite groups. If ϕ is nonzero, then $\ker(\phi)$ is a proper subrepresentation of V, hence as V is irreducible it must be zero. It follows V is isomorphic to $\phi(V)$, which is thus a nonzero subrepresentation of W. But then since W is irreducible we must have $W = \phi(V)$ and ϕ is an isomorphism as claimed.

Thus if $\mathrm{Hom}_{\mathsf{k}}(V, W)$ is nonzero, we may fix some $\phi\colon V \to W$ an isomorphism from V to W. Then given any $\mathfrak{g}$-homomorphism $\alpha\colon V \to W$, composing with ϕ^{-1} gives a $\mathfrak{g}$-homomorphism from V to V, thus it is enough to assume $W = V$. But then if $\alpha\colon V \to V$ is a $\mathfrak{g}$-endomorphism of V, since k is algebraically closed, it has an eigenvalue λ and so $\ker(\alpha - \lambda)$ is a nonzero subrepresentation, which must therefore be all of V, that is $\alpha = \lambda.\mathrm{id}_V$, so that $\mathrm{Hom}_{\mathfrak{g}}(V, V)$ is one-dimensional as claimed. □

II.2 Exact sequences of representations Parallel to the notion for Lie algebras, there is also a notion for representations. Let $\mathfrak{g}$ be a Lie algebra.

Definition II.3. A sequence of maps of $\mathfrak{g}$-representations

$$U \xrightarrow{\alpha} V \xrightarrow{\beta} W$$

is said to be *exact at* V if $\mathrm{im}(\alpha) = \ker(\beta)$. A sequence of maps

$$0 \longrightarrow U \xrightarrow{\alpha} V \xrightarrow{\beta} W \longrightarrow 0$$

is called a *short exact sequence* if it is exact at each of U, V and W, so that α is injective and β is surjective and $\mathrm{im}(\alpha) = \ker(\beta)$. If V is the middle term of such a short exact sequence, it contains a subrepresentation isomorphic to U, such that the corresponding quotient representation is isomorphic to W, and hence, roughly speaking, V is built by gluing together U and W. Just as for Lie algebras, an exact sequence

$$0 \longrightarrow U \xrightarrow{\alpha} V \xrightarrow{\beta} W \longrightarrow 0$$

is said to be *split* if β admits a right inverse $s\colon W \to V$, that is, a $\mathfrak{g}$-homomorphism s such that $\beta \circ s = \mathrm{id}_W$.

The next Lemma shows that the situation for representations is simpler than it is for Lie algebras[8]:

Lemma II.4. *Suppose that $\mathfrak{g}$ is a Lie algebra and*

$$0 \longrightarrow U \xrightarrow{\alpha} V \xrightarrow{\beta} W \longrightarrow 0$$

is a short exact sequence of $\mathfrak{g}$-representations. Then the sequence is split if and only if V is isomorphic to $U \oplus W$.

Proof. f the short exact sequence is split, then $s(W) \cap \alpha(U) = \{0\}$, since β is injective on $\mathrm{im}(s)$ and $\mathrm{im}(\alpha) = \ker(\beta)$, as if $v \in V$, then for any $v \in V$ we have $v = (v - s \circ \beta(v)) + s \circ \beta(v))$, where $s \circ \beta(v)$ lies in the image of s, and since $\beta \circ s = \mathrm{id}_W$, $v - s \circ \beta(v) \in \ker(\beta) = \mathrm{im}(\alpha)$. It follows that $V = \alpha(U) \oplus s(W)$, and since α and s are injective, the result follows. □

The notion of a composition series has an analogue for representations of a given Lie algebra $\mathfrak{g}$.

Definition II.5. Let V be a $\mathfrak{g}$-representation. A nested sequence $\mathcal{C} = (V = F_0 \supset F_1 \supset \dots \supset F_d = 0)$ is said to be a *composition series* for V if each F_i is a subrepresentation, and the subquotients F_{i-1}/F_i are irreducible (for each $i \in \{1, \dots, d\}$). The isomorphism classes of irreducibles in which occur as one off these irreducible subquotients are called the composition factors, and the multiplicity with which such an irreducible, say S, occurs is known as its composition multiplicity. We write $[S : V]$ for the multiplicity of S as a composition factor of V.

Remark II.6. A composition series can also be viewed as the vestige of how the representation V was built up from its composition factor $S_i = F_{i-1}/F_i$. Indeed for each $k \in \{1, \dots, d\}$ we have

$$0 \longrightarrow F_k \xrightarrow{\alpha} F_{k-1} \xrightarrow{\beta} S_k \longrightarrow 0$$

Thus starting with S_1 one constructs F_0/F_2 by extending it by S_2. One obtains F_0/F_3 by extending F_0/F_2 by S_3 and so on, until finally we get V by extending F_0/F_1 by $F_1 = S_1$ to obtain V itself!

Example II.7. To see a non-split extension, let $\mathfrak{g} = \mathfrak{n}_2$ be the one-dimensional Lie algebra, thought of as the (nilpotent) Lie algebra of 2×2 strictly upper triangular matrices. Then its natural 2-dimensional representation on k^2 given by the inclusion $\mathfrak{n}_2 \to \mathfrak{gl}_2(\mathsf{k})$ gives a non-split extension

$$0 \longrightarrow \mathsf{k}_0 \xrightarrow{i} \mathsf{k}^2 \longrightarrow \mathsf{k}_0 \longrightarrow 0$$

where k_0 is the trivial representation, and $i \colon \mathsf{k}_0 \to \mathsf{k}^2$ is the inclusion $t \mapsto (t, 0)$. The extension cannot be trivial, because the image of $\mathfrak{n}_2$ is non-zero. It is fact it's easy to see using linear algebra that for $\mathfrak{gl}_1(\mathsf{k}) = \mathfrak{n}_2$, an extension of one-dimensional representations k_α and k_β automatically splits if $\alpha \neq \beta$ while there is, up to isomorphism, one non-split extension of k_α with itself ($\alpha, \beta \in (\mathfrak{gl}_1(\mathsf{k}))^*$). The splitting statement is a special case of the following more general result, a special case of Theorem 5.3.24.

Lemma II.8. *Let $\mathfrak{g}$ be a nilpotent Lie algebra, and let $\alpha, \beta \in (\mathfrak{g}/D\mathfrak{g})^*$ be distinct. Any exact sequence of $\mathfrak{g}$-representations*

$$0 \longrightarrow \mathsf{k}_\alpha \longrightarrow V \longrightarrow \mathsf{k}_\beta \longrightarrow 0$$

splits, that is, $V \cong \mathsf{k}_\alpha \oplus \mathsf{k}_\beta$.

Thus non-isomorphic one-dimensional representations U and V of a nilpotent Lie algebra cannot be "glued together" in any way other than by taking their direct sum.

The following result is shows that the composition factors are actually independent of the filtration: If L is a simple representation then write $[L, V]$ for the number of times L occurs in the composition series (*i.e.* $[L, V] = \#\{i : 1 \leq i \leq k, V_i/V_{i-1} \cong L\}$.

Lemma II.9. *(Jordan-Hölder). The numbers $[L, V]$ are independent of the composition series.*

[8]In the sense that there are no non-trivial semi-direct products.

Proof. Use induction on the dimension of V. If $\dim(V) = 1$, then V is irreducible and the result is clear. Now suppose that $(V_i)_{i=1}^k$ and $(W_i)_{i=1}^l$ are two composition series for V. There is a smallest j such that $W_j \cap V_1$ is nonzero, and then since V_1 is irreducible we must have $V_1 \cap W_j = V_1$, that is, $V_1 \subseteq W_j$. But then the induced map $V_1 \to W_j/W_{j-1}$ must be an isomorphism by Schur's Lemma. Thus setting

$$W_i' = \begin{cases} (W_i \oplus V_1)/V_1 & \text{if } i < j, \\ W_{i+1}/V_1 & \text{if } j \leq i - 1. \end{cases}$$

we obtain a composition series of V/V_1, whose composition factors are those of the composition series $(W_i)_{i=1}^l$ for V, with one fewer copy of the isomorphism class of $V_1 \cong W_j/W_{j-1}$. By induction it has the same composition factors as the filtration $\{V_i/V_1 : 1 < i \leq k\}$, and we are done. □

Remark II.10. Note that the proof of the Jordan-Hölder theorem in the case of representations is a little easier than in the case of groups or Lie algebras. The reason is that, for representations, you can quotient by any sub-representations, whereas for Lie algebras, a subalgebra which occurs in a composition series is only necessarily an ideal in the smallest term of the series which strictly contains it. Thus we cannot start "at the bottom" of the composition series as we do in the above proof. Instead one must try to compare the terms "at the top" of the two composition series. The resulting argument (described in Theorem 4.2.3) is similar, but somewhat more elaborate, hence it seemed worthwhile to also give the simpler proof above for the case of representations.

II.3 Semisimplicity and complete reducibility

Definition II.11. A representation (V, ρ) is said to be *semisimple* if any subrepresentation U has a complement, that is, there is a subrepresentation W such that $V = U \oplus W$. A representation is said to be completely reducible if it is a direct sum of irreducible representations. Note that Lemma II.4 shows that V is semisimple if and only if every short exact sequence

$$0 \longrightarrow U \longrightarrow V \longrightarrow W \longrightarrow 0$$

splits. Indeed this follows from Lemma II.4: the image of a splitting map $s \colon W \to V$ gives a complement to the image of U, and s is determined by its image.

Lemma II.12. *If V is a semisimple representation, then any subrepresentation or quotient representation of V is semisimple.*

Proof. Supose that $q \colon V \to W$ is a surjective map, and that V is semisimple. We claim that W is semisimple. Indeed if W_1 is a subrepresentation of W, the $q^{-1}(W_1) = V_1$ is a subrepresentation of V, which has a complement V_2. Then it follows easily that $q(V_2)$ is a complement to W_1 in W: indeed it is clear that $W = W_1 + q(V_2)$ since $V = q^{-1}(V_1) \oplus V_2$, and if $w \in q(V_2) \cap W_1$, we may write $w = q(v)$ for some $v \in V_2$, but then $v \in q^{-1}(W_1)$, hence $w \in q^{-1}(V_1) \cap V_2 = \{0\}$.

Next, if U is a subrepresentation of V, then picking a complement U' to U, so that $V = U \oplus U'$, the corresponding projection map $\pi \colon V \to U$ with kernel U' shows that U is isomorphic to a quotient of V, and hence is also semisimple. □

Lemma II.13. *Let (V, ρ) be a representation. Then the following are equivalent:*

i) V is semisimple,

ii) V is completely reducible,

iii) V is the sum of its irreducible subrepresentations.

Proof. Once we know that any subrepresentation of a semisimple representation is again semisimple, the proof of part *ii*) of Lemma 6.3.2 shows that *i*) implies *ii*). Certainly *ii*) implies *iii*) so it is enough to show that *iii*) implies *i*). For this, suppose that V is the sum of its irreducible subrepresentations and that U is a subrepresentation of V. Let W be a subrepresentation of V which is maximal (with respect to containment) subject to the condition that $U \cap W = \{0\}$. We claim that $V = U \oplus W$. To see this, suppose that $U \oplus W \neq V$. Then by our assumption on V there must be some irreducible subrepresentation X with X not contained in $W \oplus U$, and hence $X \cap (W \oplus U) = \{0\}$. But then we certainly have[9] $(X \oplus W) \cap U = \{0\}$, which contradicts the maximality of W, so we are done. □

Remark II.14. If We write $\mathfrak{g}$ for the set of isomorphism classes of irreducible $\mathfrak{g}$-representations, then any completely reducible representation can be written as

$$V = \bigoplus_{\chi \in \mathfrak{g}} V_\chi, \text{ where } V_\chi = \sum_{\substack{S \leq V \\ S \in \chi}} S$$

Note that V_χ coincides with the isotypic subrepresentation of V defined in Definition 5.3.18, since if V is completely reducible, so is any subrepresentation, thus V_χ is semisimple, which by Lemma II.13 implies that it is the direct sum of its subrepresentations . This is called the *isotypic* decomposition of V, and any $\mathfrak{g}$-homomorphism $\theta \colon V \to W$ between $\mathfrak{g}$-representations V and W is compatible with this decomposition, in the sense that $\theta(V_\chi) \subseteq W_\chi$.

In particular, the isotypic summand of V corresponding to the trivial representation χ_0 is $V_{\chi_0} = V^{\mathfrak{g}}$, the invariants of V. A consequence of the complete reducibility is that $V^{\mathfrak{g}}$ should be a direct summand of V. In fact in the proof of Weyl's theorem, we first established this by showing any representation V of a semisimple Lie algebra $\mathfrak{g}$ decomposes as $V = V^{\mathfrak{g}} \oplus \mathfrak{g}.V$, and deduced semisimplicity from this.

[9]Since both are just expressing the fact that the sum $X + W + U$ is direct.

III *On the construction of simple Lie algebras

The classification of semisimple Lie algebras, as discussed in §7.4, relies on two key results: an *Isomorphism theorem*, and an *Existence theorem*: the former ensures that the root system captures enough information to determine the Lie algebra up to isomorphism, while the latter ensures that every abstract root system arises as the root system of some semisimple Lie algebra.

This section outlines one approach to the existence theorem. Clearly it is enough to construct a simple Lie algebra for each indecomposable root system, so we will assume throughout the remainder of this section that (V, Φ) is indecomposable. We will establish the existence theorem in two steps. In the first step we consider the case where all the roots in Φ have the same length, and in the second step deduce from this the general case. An alternative elementary approach is described in [Gec17].

III.1 The simply-laced case

Definition III.1. Let (V, Φ) be an (indecomposable) root system. We say that (V, Φ) is *simply-laced* if all the roots in Φ have the same length.

If Δ is a set of simple roots for Φ, since $\Phi = W.\Delta$ (where W is the Weyl group) it is equivalent to the condition that all the roots in Δ have the same length. Since (V, Φ) is indecomposable, this in turn is equivalent to the condition that $\langle \alpha, \beta \rangle = \langle \beta, \alpha \rangle$ for all $\alpha, \beta \in \Delta$, that is, the Cartan matrix is symmetric. By Lemma 7.3.6, this is equivalent to the condition that for all $\alpha, \beta \in \Phi$ the Cartan integer $\langle \alpha, \beta \rangle \in \{0, -1\}$. If we normalize the inner product on V so that the roots have length $\sqrt{2}$, then the Cartan integers are precisely the values of the inner product on pairs of simple roots.

From the classification of abstract root systems, one can check that the simply-laced indecomposable root systems are those of types A,D and E.

To construct a Lie algebra from such a root system, we need one additional ingredient: Let $\epsilon \colon Q \times Q \to \{\pm 1\}$ be a *bimultiplicative function*, that is, for all $\alpha, \beta, \gamma \in Q$,

$$\begin{aligned} \epsilon(\alpha + \beta, \gamma) &= \epsilon(\alpha, \gamma).\epsilon(\beta, \gamma), \\ \epsilon(\alpha, \beta + \gamma) &= \epsilon(\alpha, \beta).\epsilon(\alpha, \gamma). \end{aligned}$$

and suppose also that it satisfies

$$\epsilon(\alpha, \alpha) = (-1)^{(\alpha,\alpha)/2}, \quad \forall \alpha \in Q \tag{I.8}$$

(note that since $(\alpha, \alpha) = 2$ for all roots $\alpha \in \Phi$, we must have $(\beta, \beta) \in 2\mathbb{Z}$ for any $\beta \in Q$). Such a function is called an *asymmetric function*. Since $(\alpha, \beta) \in \mathbb{Z}$ for any $\alpha, \beta \in Q$ we can replace α by $\alpha + \beta$ in the second condition (I.8) for an asymmetric function to obtain:

$$\epsilon(\alpha, \beta)\epsilon(\beta, \alpha) = (-1)^{(\alpha,\beta)}. \tag{I.9}$$

Note that the bimultiplicativity property means it is determined by its values on a base Δ and moreover the second condition (I.8) requires $\epsilon(\alpha_i, \alpha_i) = -1$ for any $\alpha_i \in \Delta$. To construct such a function on the rest of $\Delta \times \Delta$, orient the edges of the Dynkin diagram, whose vertices are labelled by the base $\Delta = \{\alpha_1, \dots, \alpha_l\}$, arbitrarily, and then define for $\alpha_i \neq \alpha_j$

$$\epsilon(\alpha_i, \alpha_j) = \begin{cases} -1 & \text{if there is an edge going from } \alpha_i \text{ to } \alpha_j, \\ +1 & \text{otherwise.} \end{cases}$$

It the follows from this definition that Equation (I.9) holds for all roots in our base, and thus extending this ϵ bimultiplicatively, we obtain an asymmetric function on all of Q.

We can now give a construction of the Lie algebra $\mathfrak{g}_Q$ associated to our root system: Let $\mathfrak{h}^*$ denote the extension of scalars from $\mathbb{Q}$ to our field k of V, and similarly we can extend our inner product to a symmetric bilinear form on $\mathfrak{h}^*$. Let $\mathfrak{h}$ be the dual of $\mathfrak{h}$.

Definition III.2. Let $\mathfrak{g}_Q = \mathfrak{h} \oplus \bigoplus_{\alpha\in\Phi} \mathsf{k}.e_\alpha$ as a vector space, and let h_α be the image of α under the isomorphism between $\mathfrak{h}$ and $\mathfrak{h}^*$ given by the nondegenerate symmetric bilinear form on $\mathfrak{h}^*$ induced from the inner product on V. We define

$$\begin{aligned}
&[h,h'] = 0, \forall h,h' \in \mathfrak{h};\\
&[h,e_\alpha] = \alpha(h)e_\alpha;\\
&[e_\alpha,e_\beta] = \begin{cases} -h_\alpha, & \text{if } \alpha+\beta=0;\\ \epsilon(\alpha,\beta).e_{\alpha+\beta} & \text{if } \alpha+\beta\in\Phi;\\ 0 & \text{otherwise.}\end{cases}
\end{aligned}$$

We also extend the symmetric bilinear form on $\mathfrak{h}$ (obtained by identifying it with $\mathfrak{h}^*$) to all of $\mathfrak{g}_Q$ by setting $(e_\alpha,e_\beta) = -\delta_{\alpha,-\beta}$, and making $\mathfrak{h}$ orthogonal to $\bigoplus_{\alpha\in\Phi} \mathsf{k}.e_\alpha$. (Note the minus signs in the definition of the invariant form and in the bracket $[e_\alpha,e_{-\alpha}]$ are consistent.)

Proposition III.3. *The definition above gives a Lie algebra which has $\mathfrak{h}$ as a Cartan subalgebra and root system Φ, and the form on $\mathfrak{g}_Q$ is invariant.*

Proof. (*Sketch)*: We must show that $\mathfrak{g}_Q$ is a Lie algebra, that is, we must check that the bilinear map $[,]$ defined above is a Lie bracket. To see that it is alternating, note that if $\{\alpha,\beta,\alpha+\beta\}\subseteq\Phi$ then, since the root system is simply-laced, $(\alpha,\beta) = -1$, and hence (I.9) shows that $\epsilon(\alpha,\beta) = -\epsilon(\beta,\alpha)$. It remains to check that $[.]$ satisfies the Jacobi identity. It is enough to check this on three basis elements, x,y and z. If at least one of our basis elements is in $\mathfrak{h}$ this is easy (the properties of the bimultiplicative function beyond the one already used for the alternating property are not involved). For example, if $x = h\in\mathfrak{h}$, $y = e_\alpha$, $z = e_\beta$ then (setting $e_{\alpha+\beta} = 0$ if $\alpha+\beta\notin\Phi$)

$$\begin{aligned}
&[h,[e_\alpha,e_\beta]] + [e_\alpha,[e_\beta,h]] + [e_\beta,[h,e_\alpha]]\\
&\qquad = \epsilon(\alpha,\beta)\left((\alpha+\beta)(h)e_{\alpha+\beta} - \beta(h)e_{\alpha+\beta} - \beta(h)e_{\alpha+\beta}\right)\\
&\qquad = 0.
\end{aligned}$$

If x,y,z are of the form $e_\alpha,e_\beta,e_\gamma$ then there are a number of cases to check. Firstly, if none of $\alpha+\beta,\alpha+\gamma,\beta+\gamma$ lie in $\Phi\cup\{0\}$, then the Jacobi identity holds trivially. Thus let us suppose that $\alpha+\beta\in\Phi\cup\{0\}$. Note that $\alpha\pm\beta\in\Phi$ if and only if $(\alpha,\beta) = \mp 1$. Moreover, it follows that $\epsilon(\alpha,\alpha) = -1$ and $\epsilon(\alpha,\beta)\epsilon(\beta,\alpha) = -1$.

There are four cases: 1) $\alpha\pm\gamma\notin\Phi\cup\{0\}$; 2) either $\alpha+\gamma$ or $\alpha-\gamma = 0$; 3) $\alpha+\gamma\in\Phi$ and; 4) $\alpha-\gamma\in\Phi$. Cases 1) and 2) are easy to check, case 3) follows from

$$\epsilon(\gamma,\alpha)\epsilon(\gamma+\alpha,-\alpha) = (\alpha,\alpha).$$

In this fashion one can reduce to the case where $\alpha+\beta$, $\alpha+\gamma$ and $\beta+\gamma$ all lie in Φ. But then $(\alpha,\beta) = (\alpha,\gamma) = (\beta,\gamma) = -1$ and so $(\alpha+\beta+\gamma,\alpha+\beta+\gamma) = 0$ so that $\alpha+\beta+\gamma = 0$. In this case the Jacobi identity

$$[e_\alpha,[e_\beta,e_\gamma]] + [e_\beta,[e_\gamma,e_\alpha]] + [e_\gamma,[e_\alpha,e_\beta]] = 0$$

reduces to

$$\epsilon(\beta,\gamma)\epsilon(\alpha,\beta+\gamma) + \epsilon(\gamma,\alpha)\epsilon(\beta,\alpha+\gamma) + \epsilon(\alpha,\beta)\epsilon(\gamma,\alpha+\beta) = 0$$

which can be checked using the properties of ϵ.

It is similar, though more straight-forward, to check that the symmetric bilinear form we have defined is invariant. □

III.2 The non-simply-laced cases One can also use the construction of the simply-laced Lie simple Lie algebras to give a construction of *all* simple Lie algebras: We do this as follows: Given a simply-laced Dynkin diagram D, a *admissible diagram automorphism* is a graph automorphism $\sigma\colon D \to D$ with the property that the orbit of a vertex is discrete, that is, there is no edge between a vertex i and $\sigma^k(i)$ for any $k \in \mathbb{Z}$.

Given such an automorphism, we claim that σ induces an automorphism of $\mathfrak{g}_Q$ the associated simple Lie algebra. To see this, note that we can pick the orientation of our Dynkin diagram so that it is invariant under the diagram automorphism (we will check this shortly for the automorphisms we need). Clearly σ induced an isometry of V to itself preserving the roots Φ (it clearly preserves Q and hence Φ since Φ is the set of norm 2 vectors in Q). Moreover, it preserves the bimultiplicative function ϵ because it preserved the orientation of our Dynkin diagram (by our choice of orientation).

Defining σ on $\mathfrak{g}_Q$ by letting $\sigma(e_\alpha) = e_{\sigma(\alpha)}$ and letting it act on $\mathfrak{h}$ by extension of scalars of its action on V, it is then clear that σ is a Lie algebra homomorphism. It follows that its fixed point set is a sub-Lie algebra.

Theorem III.4. *The Lie algebra $\mathfrak{g}_Q^\sigma$ is a simple Lie algebra with Dynkin diagram D^σ given as follows: the vertices of D^σ are the orbits of σ on the vertex set of D, and, for any two orbits, they are joined if there were edges joining a vertex in one orbit to a vertex in the other, etc..*

Bibliography

[Gec17] Meinolf Geck. On the construction of semisimple lie algebras and chevalley groups. *Proc. Amer. Math. Soc.*, 145(8):3233–3247, 2017.