LIE ALGEBRAS

KEVIN MCGERTY (WITH MINOR MODIFICATIONS BY BALÁZS SZENDRŐI 2013, DAN CIUBOTARU 2015)

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 (and hopefully if you don't but are willing to think imprecisely at some points) it will help to put the course in context. For those worried about such things, fear not, it is non-examinable.

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

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

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. 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 $\mathcal{C}^\infty(\mathbb{R}^2)$ 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 $\nu\colon \mathcal{C}^\infty(\mathbb{R}^2) \to \mathcal{C}^\infty(\mathbb{R}^2)$, given by

$$\begin{aligned} f \mapsto & \frac{d}{dt}\big(f(g(-t).\begin{pmatrix} x \\ y \end{pmatrix})\big)_{|t=0} \\ &= y\partial_x - x\partial_y. \end{aligned}$$

It is immediate from the product rule for differentiation, that the operator ν constructed in the above example obeys the "Leibniz rule":

$$\nu(f.g) = \nu(f).g + f.\nu(g).$$

An operator on smooth functions which satisfies this property is called a *derivation*. It's not hard to see that any such operator on $\mathcal{C}^\infty(\mathbb{R}^2)$ will be of the form $a(x,y)\partial_x + b(x,y)\partial_y$ where $a, b \in \mathcal{C}^\infty(\mathbb{R}^2)$. Thus, heuristically for now, 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 ν, but for actions of larger group we would attach a larger space of derivations). It turns out this set of

[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 $\mathbb{R}^n$.

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]. Let's set this up a little more formally.

Definition 1.2. A *vector field* on $X = \mathbb{R}^n$ (or, with a bit more work, any manifold) is a (smooth) function $\nu\colon \mathbb{R}^n \to \mathbb{R}^n$, 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 $\nu = (a_1, a_2, \ldots, a_n)$ in standard coordinates (here $a_i : \mathbb{R}^n \to \mathbb{R}$), then set

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

This formula gives an action of Θ_X on the space of smooth functions $\mathcal{C}^\infty(X)$, since if $f \in \mathcal{C}^\infty(X)$, then so is $\nu(f)$. This action is linear, and interacts with multiplication of functions according to the Leibniz rule, that is, if ν is a vector field, and $f_1, f_2 \in \mathcal{C}^\infty(X)$ then

$$\nu(f_1 f_2) = \nu(f_1).f_2 + f_1.\nu(f_2),$$

in other words, vector fields act as derivations on smooth functions in the above sense. Note we can talk about vector fields and derivations interchangeably, since the derivation given by a vector field completely determines the vector field, and any derivation comes from a vector field (this is an exercise that is worth checking).

Note that if we compose two derivations $\nu_1 \circ \nu_2$ 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 ν_1, ν_2 are derivations, then $[\nu_1, \nu_2] = \nu_1 \circ \nu_2 - \nu_2 \circ \nu_1$ 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 1.3. Consider the action of $\mathrm{SO}_3(\mathbb{R})$ on $\mathbb{R}^3$. Using the fact that every element of $\mathrm{SO}_3(\mathbb{R})$ is a rotation about some axis through the origin it is not too hard to find the space of vector fields on $\mathbb{R}^3$ which can be associated to this action, and check that it forms a Lie algebra. Indeed as we saw before, the action of the circle fixing the z-axis gives the derivation $-y\partial_x + x\partial_y$, and the derivation obtained from rotation about any other axis will be obtained by an orthogonal change of coordinates. It can be shown that these form the 3-dimensional 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. 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.

Definition 2.1. 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 2.2. It is easy to check directly from the definition that the Lie bracket 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 $\mathrm{char}(\mathsf{k}) \neq 2$, the alternating condition is equivalent to skew-symmetry. Note that a Lie algebra is an algebra with a product which is neither commutative nor associative, and moreover it does not have an identity element[4]. We will normally simply write $[.,.]$ and reserve use the decorated bracket only for emphasis or where there is the potential for confusion.

[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.

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

Example 2.3. (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) 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 a Lie algebra under the commutator product, that is:

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

Indeed the alternating property is immediate, so the only thing to check is the Jacobi identity, which is an easy computation.

(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) If A is an associative k-algebra, then if $a \in A$ let $\delta_a\colon A \to A$ be the linear map given by

$$\delta_a(b) = a.b - b.a, \quad b \in A.$$

One can check that δ_a is a derivation on A, and that this is equivalent to the statement that $(A, [.,.]_A)$ is a Lie algebra, where $[.,.]_A$ is the commutator bracket on A, that is $[a,b]_A = a.b - b.a$. Thus any associative algebra can be given the structure of a Lie algebra. (This is a generalisation of the case of $n \times n$ matrices).

Remark 2.4. One could begin to try and classify all (say finite-dimensional) Lie algebras. In very low dimension this is actually possible. For dimension 1 clearly there is a unique (up to isomorphism[5]) 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. In this course we will focus on a particular class of Lie algebras, known as *semisimple Lie algebras*, for which an elegant classification theorem is known.

3. Homomorphisms and ideals

We have already introduced the notion of a subalgebra of a Lie algebra in the examples above, but there are other standard constructions familiar from rings which make sense for Lie algebras. A *homomorphism* of Lie algebras $(\mathfrak{g}, [.,.]_{\mathfrak{g}})$ and $(\mathfrak{g}', [.,.]_{\mathfrak{g}'})$ is a k-linear map $\phi\colon \mathfrak{g} \to \mathfrak{g}'$ which respects the Lie brackets, that is:

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

An *isomorphism* of Lie algebras is a bijective homomorphism. An *ideal* in a Lie algebra $\mathfrak{g}$ is a subspace I such that for all $a \in \mathfrak{g}$ and $x \in I$ we have $[a, x]_{\mathfrak{g}} \in I$. It is easy to check that if $\phi\colon \mathfrak{g} \to \mathfrak{g}'$ is a homomorphism, then $\ker(\phi) = \{a \in \mathfrak{g} : \phi(a) = 0\}$ is an ideal of $\mathfrak{g}$. Conversely, if I is an ideal of $\mathfrak{g}$ then it is easy to check that the quotient space $\mathfrak{g}/I$ inherits the structure of a Lie algebra, and the canonical quotient map $q\colon \mathfrak{g} \to \mathfrak{g}/I$ is a Lie algebra homomorphism with kernel I.

Remark 3.1. 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 it is *simple*.

[5]Of course I haven't said what an isomorphism of Lie algebras *is* yet (see below) but you probably know...

Just as for groups and rings, we have the normal stable of isomorphism theorems, and the proofs are identical.

Theorem 3.2. (1) *Let $\phi\colon \mathfrak{g} \to \mathfrak{g}'$ be a homomorphism of Lie algebras. The subspace $\phi(\mathfrak{g}) = im(\phi)$ is a subalgebra of $\mathfrak{g}'$ and ϕ induces an isomorphism $\bar{\phi}\colon \mathfrak{g}/ker(\phi) \to im(\phi)$.*

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

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

(3) *If I, J are ideals of $\mathfrak{g}$ then we have*

$$(I + J)/J \cong I/(I \cap J).$$

4. Representations of Lie algebras

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 objects. For Lie algebras, we will use actions on linear spaces, or in other words, "representations". Formally we make the following definition.

Definition 4.1. 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 $x \in \mathfrak{g}, v \in V$.

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 4.2. (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 the fact that this is a representation is just a rephrasing[6] 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 $y \in \mathfrak{g}$.

(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.

There are a number of standard ways of constructing new representations from old, all of which have their analogue for group representations. For example, 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. Next, 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 $\rho \oplus \sigma$. More interestingly, the vector space $\mathrm{Hom}(V, W)$ of linear maps from V to W has the structure of a $\mathfrak{g}$-representation via

$$x(\phi) = \sigma(x) \circ \phi - \phi \circ \rho(x), \quad \forall x \in \mathfrak{g}, \phi \in \mathrm{Hom}(V, W). \tag{4.1}$$

---

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

It is straight-forward to check that this gives a Lie algebra homomorphism from $\mathfrak{g}$ to $\mathfrak{gl}(\text{Hom}(V,W))$. One way to do this is simply to compute directly. Another, slightly quicker way, is to notice that $\text{End}_{\mathsf{k}}(V \oplus W) = \text{Hom}_{\mathsf{k}}(V \oplus W, V \oplus W)$ contains $\text{End}_{\mathsf{k}}(V)$, $\text{End}_{\mathsf{k}}(W)$ and $\text{Hom}_{\mathsf{k}}(V,W)$ since

$$\text{Hom}_{\mathsf{k}}(V \oplus W, V \oplus W) = \text{Hom}_{\mathsf{k}}(V,V) \oplus \text{Hom}_{\mathsf{k}}(V,W) \oplus \text{Hom}_{\mathsf{k}}(W,V) \oplus \text{Hom}_{\mathsf{k}}(W,W),$$

thus we may combine ρ and σ to give a homomorphism $\tau\colon \mathfrak{g} \to \mathfrak{gl}(V \oplus W)$. Since $\text{Hom}_{\mathsf{k}}(V,W)$ is clearly preserved by $\tau(\mathfrak{g})$, it is a subrepresentation of $\mathfrak{gl}(V \oplus W)$, where the latter is a $\mathfrak{g}$-representation via the composition of $\tau\colon \mathfrak{g} \to \mathfrak{gl}(V \oplus W)$ with the adjoint representation $\text{ad}\colon \mathfrak{gl}(V \oplus W) \to \mathfrak{gl}(\mathfrak{gl}(V \oplus W))$. It is then easy to see that the equation (4.1) describes the resulting action of $\mathfrak{g}$ on $\text{Hom}_{\mathsf{k}}(V,W)$.

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 $V^* = \text{Hom}(V,\mathsf{k})$, 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[7] $-\rho(x)^t$. 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.

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

Definition 4.3. 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.

Example 4.4. 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.

4.1. **Tensor products and $\mathfrak{g}$-representations.** Another 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. In this section we try to understand why this should be the case.

It is useful first recall the case of group representations. Thus suppose that G is a group (which can be finite or infinite) and that (V,ρ) and (W,σ) are representations of G. The $(\rho,\sigma)\colon G \times G \to \text{GL}(V) \times \text{GL}(W)$ is a homomorphism of groups, hence it suffices to show that $V \otimes W$ is a representation of $\text{GL}(V) \times \text{GL}(W)$. To see why that is true, notice that if U is a representation of a product of groups $G \times H$ say, then if $\tau\colon G \times H \to \text{GL}(U)$ is a representation of $G \times H$, it is clearly determined by its restrictions to $G \times \{1\} \cong G$ and $\{1\} \times H \cong H$, which we denote by τ_G and τ_H. But since the subgroups G and H commute in $G \times H$, the action maps τ_G, τ_H cannot be arbitrary, they must have the property that $\tau(g)\tau(h) = \tau(h)\tau(g)$ for all $g \in G, h \in H$. Thus to give a representation of $\text{GL}(V) \times \text{GL}(W)$ on $V \otimes W$ we must give actions of G and H on $V \otimes W$ which commute with each other. But this is easy: if $\alpha \in \text{Hom}(V,V)$ and $\beta \in \text{Hom}(W,W)$, then the pair (α,β) defines a linear map $\alpha \otimes \beta \in \text{Hom}(V \otimes W, V \otimes W)$ where $\alpha \otimes \beta$ is given by $(\alpha \otimes \beta)(v \otimes w) = \alpha(v) \otimes \beta(w)$. (The map $(v,w) \mapsto \alpha(v) \otimes \beta(w)$ from $V \times W$ to $V \otimes W$ is bilinear, and hence by the universal property it induces a linear map from $V \otimes W$ to itself.) Moreover, if $\alpha_1, \alpha_2 \in \text{Hom}(V,V)$ and $\beta_1, \beta_2 \in \text{Hom}(W,W)$, then, using the universal property, and the fact that composition of linear maps is bilinear, it follows that

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

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

$$(\alpha \otimes 1_W) \circ (1_V \otimes \beta) = \alpha \otimes \beta = (1_V \otimes \beta) \circ (\alpha \otimes 1_W), \tag{4.2}$$

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

[7]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.

Suppose now that $\mathfrak{g}$ is a Lie algebra and (V,ρ) and (W,σ) are $\mathfrak{g}$-representations. We proceed in the same manner as for the group case: The map $\rho\oplus\sigma\colon \mathfrak{g}\to\mathfrak{gl}(V)\oplus\mathfrak{gl}(W)$ given by $(\rho\oplus\sigma)(x) = (\rho(x),\sigma(x))$ is a Lie algebra homomorphism, and the representations V and W are obtained from $\rho\oplus\sigma$ by composition with the obvious projection maps from $\mathfrak{gl}(V)\oplus\mathfrak{gl}(W)$ to $\mathfrak{gl}(V)$ and $\mathfrak{gl}(W)$ respectively, thus the question of whether, for an arbitrary Lie algebra $\mathfrak{g}$, the tensor product of two representations $V\otimes W$ is naturally a $\mathfrak{g}$-representation, it suffices to determine whether $V\otimes W$ carries the structure of a representation of $\mathfrak{gl}(V)\oplus\mathfrak{gl}(W)$ in some natural way.

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 4.5. *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 $[\alpha_1(\mathfrak{g}_1),\alpha_2(\mathfrak{g}_2)]=0$.*

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 assumption that $[\alpha_1(\mathfrak{g}_1),\alpha_2(\mathfrak{g}_2)] = [\alpha_2(\mathfrak{g}_2),\alpha_1(\mathfrak{g}_1)]=0$. 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 (4.2), 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 4.5 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 4.6. 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

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

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

$$(\rho\otimes\sigma)(x)(v\otimes w)=\rho(x)(v)\otimes w+v\otimes\sigma(x)(w),\quad\forall v\in V, w\in W. \tag{4.3}$$

Remark 4.7. 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 (4.3), 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 4.8. 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.

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 4.3 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:

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

Thus all the maps between tensor products of vector spaces discuss in Appendix 22.1 yield maps of $\mathfrak{g}$-representations.

5. Nilpotent Lie algebras

We now begin to study particular classes of Lie algebra. The first class we study are nilpotent Lie algebras, which are somewhat analogous to nilpotent groups. We need a few more definitions.

Definition 5.1. 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 if $i \in I, j \in J, x \in \mathfrak{g}$ we have:

$$[x, [i, j]] = -[i, [j, x]] - [j, [x, i]] = [i, [x, j]] + [[x, i], j] \in [I, J]$$

using the Jacobi identity in the first equality and skew-symmetry in the second.

Definition 5.2. 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 $\mathfrak{g}$, and we say $\mathfrak{g}$ is *nilpotent* if $C^N(\mathfrak{g}) = 0$ for some $N > 0$. If N is the smallest integer such that $C^N(\mathfrak{g}) = 0$ then we say that $\mathfrak{g}$ is an *N-step nilpotent* Lie algebra.

For example, a Lie algebra is 1-step nilpotent if and only if it is abelian. The definition can be rephrased as follows: there is an $N > 0$ such that for any N elements $x_1, x_2, \ldots, x_N$ of $\mathfrak{g}$ the iterated Lie bracket

$$[x_1, [x_2, [\ldots, [x_{N-1}, x_N]] \ldots] = \mathrm{ad}x_1(\mathrm{ad}x_2(\ldots \mathrm{ad}x_{N-1}(x_N)) \ldots) = 0.$$

In particular, all the elements $\mathrm{ad}(x) \in \mathfrak{gl}(\mathfrak{g})$ for $x \in \mathfrak{g}$ are nilpotent.

Remark 5.3. The ideal $C^1(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}]$ is known as the *derived subalgebra*[8] of $\mathfrak{g}$ and is also denoted[9] $D(\mathfrak{g})$ and sometimes $\mathfrak{g}'$.

Lemma 5.4. *Let $\mathfrak{g}$ be a Lie algebra. Then*

1. *If $\mathfrak{g}$ is nilpotent, so is any subalgebra or quotient of $\mathfrak{g}$.*
2. *If $\mathfrak{g}/\mathfrak{z}(\mathfrak{g})$ is nilpotent, then $\mathfrak{g}$ is nilpotent.*

Proof. The first part is immediate from the definition. Indeed if $\mathfrak{h} \subseteq \mathfrak{g}$ is a subalgebra of $\mathfrak{g}$ then clearly we have $C^i(\mathfrak{h}) \subseteq C^i(\mathfrak{g})$, so that if $C^N(\mathfrak{g}) = 0$ we also have $C^N(\mathfrak{h}) = 0$. Similarly, an easy induction say shows that if $\mathfrak{h}$ is an ideal, then $C^i(\mathfrak{g}/\mathfrak{h}) = (C^i(\mathfrak{g}) + \mathfrak{h})/\mathfrak{h}$, and so again if $C^N(\mathfrak{g}) = 0$ we must also have $C^N(\mathfrak{g}/\mathfrak{h}) = 0$.

For the second claim, taking $\mathfrak{h} = \mathfrak{z}(\mathfrak{g})$, an ideal of $\mathfrak{g}$, we see from the first part that if $\mathfrak{g}/\mathfrak{z}(\mathfrak{g})$ is nilpotent, then there is some N with $(C^N(\mathfrak{g}) + \mathfrak{z}(\mathfrak{g}))/\mathfrak{z}(\mathfrak{g}) = C^N(\mathfrak{g}/\mathfrak{z}(\mathfrak{g})) = 0$, and so $C^N(\mathfrak{g}) \subseteq \mathfrak{z}(\mathfrak{g})$. But then it is clear that $C^{N+1}(\mathfrak{g}) = 0$, and so $\mathfrak{g}$ is nilpotent as required. □

Remark 5.5. Notice that if I is an arbitrary ideal in $\mathfrak{g}$, and I and $\mathfrak{g}/I$ are nilpotent it does *not* follow that $\mathfrak{g}$ is nilpotent. Indeed recall that if $\mathfrak{g}$ is the non-abelian 2-dimensional Lie algebra, then $\mathfrak{g}$ can be given a basis x, y with $[x, y] = y$. Hence $\mathsf{k}.y$ is a 1-dimensional ideal in $\mathfrak{g}$ (which is thus abelian and so nilpotent) and the quotient is again 1-dimensional and so nilpotent. However, $\mathrm{ad}(x)$ has y as an eigenvector with eigenvalue 1, so $\mathfrak{g}$ cannot be nilpotent, and indeed $C^i(\mathfrak{g}) = \mathsf{k}.y$ for all $i \geq 1$.

Example 5.6. Let V be a vector space, and

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

a sequence of subspaces with $\dim(F_i) = i$ (such a sequence is known as a *flag* or *complete flag* in V). Let $\mathfrak{n} = \mathfrak{n}_{\mathcal{F}}$ be the subalgebra of $\mathfrak{gl}(V)$ consisting of linear maps $X \in \mathfrak{gl}(V)$ such that $X(F_i) \subseteq F_{i-1}$ for all $i \geq 1$. We claim the Lie algebra $\mathfrak{n}$ is nilpotent. To see this we show something more precise. Indeed for each positive integer k, consider 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 $\mathfrak{n}^k \subset \mathfrak{n}$, and $\mathfrak{n}^k = 0$ for any $k \geq n$. We claim that $C^k(\mathfrak{n}) \subseteq \mathfrak{n}^{k+1}$, which will therefore prove $\mathfrak{n}$ is nilpotent. The claim is immediate for $k = 0$, so suppose we know by induction that $C^k(\mathfrak{n}) \subseteq \mathfrak{n}^{k+1}$. Then 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 $C^{n-1}(\mathfrak{n}) = 0$ (note that if $\dim(V) = 1$ then $\mathfrak{n} = 0$). If we pick a basis $\{e_1, e_2, \ldots, e_n\}$ of V such that $F_i = \mathrm{span}(e_1, e_2, \ldots, e_i\}$ then the matrix A representing an element $x \in \mathfrak{n}$ with respect to this basis is strictly upper triangular, that is, $a_{ij} = 0$ for all $i \geq j$. It follows that $\dim(\mathfrak{n}) = \binom{n}{2}$, so when $n = 2$ we

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

[9] 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.

just get the 1-dimensional Lie algebra, thus the first nontrivial case is when $n = 3$ and in that case we get a 3-dimensional 2-step nilpotent Lie algebra.

Note that the subalgebra $\mathfrak{t} \subset \mathfrak{gl}_n$ of diagonal matrices is nilpotent, since it is abelian, so a nilpotent linear Lie algebra need not consist of nilpotent endomorphisms. Nevertheless we will show there is a close connection between the two notions.

Lemma 5.7. *If $x \in \mathfrak{gl}(V)$ be a nilpotent endomorphism then $ad(x) \in \mathrm{End}(\mathfrak{gl}(V))$ is also nilpotent.*

Proof. The map $\lambda_x \colon \mathfrak{gl}(V) \to \mathfrak{gl}(V)$ given by $y \mapsto xy$ is clearly nilpotent if x is nilpotent, and similarly for the map $\rho_x \colon \mathfrak{gl}(V) \to \mathfrak{gl}(V)$ given by $y \mapsto yx$. Moreover, λ_x and ρ_x clearly commute with each other, so since $\mathrm{ad}(x) = \lambda_x - \rho_x$, it is also nilpotent. Indeed for $m \geq 0$ we have

$$(\lambda_x - \rho_x)^m = \sum_{i=0}^{m} (-1)^i \binom{m}{i} \lambda_x^{m-i} \rho_x^i$$

and so if $x^n = 0$, so that $\lambda_x^n = \rho_x^n = 0$, then if $m \geq 2n$, every term on the right-hand side must be zero, and so $\mathrm{ad}(x)^m = 0$ as required. □

For the next proposition we need the notion of the normaliser of a subalgebra.

Definition 5.8. Let $\mathfrak{g}$ be a Lie algebra and let $\mathfrak{a}$ be a subalgebra. The subspace

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

is a subalgebra of $\mathfrak{g}$ (check this using the Jacobi identity) which is called the *normaliser* of $\mathfrak{a}$. It is the largest subalgebra of $\mathfrak{g}$ in which $\mathfrak{a}$ is an ideal.

Definition 5.9. If $\mathfrak{g}$ is any Lie algebra and (V, ρ) is a representation of $\mathfrak{g}$, then define

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

It is called the (possibly zero in general) subrepresentation of *invariants* in V.

Proposition 5.10. *Let $\mathfrak{n}$ be a Lie algebra, and (V, ρ) a representation of $\mathfrak{n}$ such that for every $x \in \mathfrak{n}$, the linear map $\rho(x)$ is nilpotent. Then the invariant subspace*

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

is non-zero.

Proof. We use induction on $d = \dim(\mathfrak{n})$, the case $d = 1$ being clear. Clearly the statement of the proposition is unchanged if we replace $\mathfrak{n}$ by its image in $\mathfrak{gl}(V)$, so we may assume that $\mathfrak{n}$ is a subalgebra of $\mathfrak{gl}(V)$.

Now consider $\mathfrak{a} \subsetneq \mathfrak{n}$ a proper subalgebra. Since the elements $a \in \mathfrak{a}$ are nilpotent endomorphisms, the previous lemma shows that $\mathrm{ad}(a) \in \mathfrak{gl}(\mathfrak{n})$ is also nilpotent, hence $\mathrm{ad}(a)$ also acts nilpotently on the quotient[10] $\mathfrak{n}/\mathfrak{a}$. Since $\dim(\mathfrak{a}) < \dim(\mathfrak{n})$, by induction (applied with $V = \mathfrak{n}/\mathfrak{a}$ and $\mathfrak{a}$) we can find $x \notin \mathfrak{a}$ such that $\mathrm{ad}(a)(x) = [a, x] \in \mathfrak{a}$ for all $a \in \mathfrak{a}$. It follows that for any proper subalgebra $\mathfrak{a}$, its normaliser

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

strictly contains $\mathfrak{a}$.

Thus if $\mathfrak{a}$ is a proper subalgebra of $\mathfrak{n}$ of maximal dimension, we must have $N_{\mathfrak{n}}(\mathfrak{a}) = \mathfrak{n}$, or in other words, $\mathfrak{a}$ must actually be an ideal of $\mathfrak{n}$.

Now if $\mathfrak{n}/\mathfrak{a}$ is not one-dimensional, the preimage of a one-dimensional subalgebra in $\mathfrak{n}/\mathfrak{a}$ would be a proper subalgebra of $\mathfrak{n}$ strictly containing $\mathfrak{a}$, which again contradicts the maximality of $\mathfrak{a}$. Thus $\mathfrak{n}/\mathfrak{a}$ is one-dimensional, and we may find $z \in \mathfrak{n}$ so that $\mathsf{k}.z \oplus \mathfrak{a} = \mathfrak{n}$.

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 $U = V^{\mathfrak{n}}$ we are done. □

[10]The notation here can be a bit confusing: the Lie algebra $\mathfrak{n}$ is an $\mathfrak{a}$-representation by the restriction of the adjoint representation of $\mathfrak{g}$ to $\mathfrak{a}$. Since $\mathfrak{a}$ is a sub-$\mathfrak{a}$-representation of $\mathfrak{g}$, the quotient $\mathfrak{g}/\mathfrak{a}$ is an $\mathfrak{a}$-representation. Note however that it is not itself a Lie algebra unless $\mathfrak{a}$ is an ideal of $\mathfrak{g}$.

Definition 5.11. If $\mathfrak{g}$ is a Lie algebra and $x \in \mathfrak{g}$, we say that an element x is ad-nilpotent if $\text{ad}(x) \in \mathfrak{gl}(\mathfrak{g})$ is a nilpotent endomorphism. For convenience[11] we say that $\mathfrak{g}$ is ad-nilpotent if all of its elements are.

Theorem 5.12. *(Engel's theorem) A Lie algebra $\mathfrak{g}$ is nilpotent if and only if ad(x) is nilpotent for every $x \in \mathfrak{g}$.*

Proof. In the terminology of the above definition, the theorem states that a Lie algebra is nilpotent if and only if it is ad-nilpotent. As we already noted, it is immediate from the definition that a nilpotent Lie algebra is ad-nilpotent, so our task is to show the converse. For this, note that it is clear that if $\mathfrak{g}$ is ad-nilpotent, then any quotient of it is also ad-nilpotent. Hence using induction on dimension along with Lemma 5.4 (2), it will be enough to show that $\mathfrak{z}(\mathfrak{g})$ is a non-zero. Now

$$\begin{aligned}\mathfrak{z}(\mathfrak{g}) &= \{z \in \mathfrak{g} : [z, x] = 0, \forall x \in \mathfrak{g}\}\\ &= \{z \in \mathfrak{g} : -\text{ad}(x)(z) = 0, \forall x \in \mathfrak{g}\}\\ &= \mathfrak{g}^{\text{ad}(\mathfrak{g})},\end{aligned}$$

hence applying Proposition 5.10 to the adjoint representation of $\mathfrak{g}$, it follows immediately that $\mathfrak{z}(\mathfrak{g}) \neq 0$ and we are done.

□

The following is an important consequence of the Proposition 5.10.

Corollary 5.13. *Let $\mathfrak{g}$ be a Lie algebra and (V, ρ) a representation of $\mathfrak{g}$ such that $\rho(x)$ is a nilpotent endomorphism for all $x \in \mathfrak{g}$. Then there is a complete flag $\mathcal{F} = (0 = F_0 \subset F_1 \subset \ldots \subset F_n = V)$ such that $\rho(\mathfrak{g}) \subseteq \mathfrak{n}_{\mathcal{F}}$.*

Proof. Let us say that $\mathfrak{g}$ respects a flag $\mathcal{F}$ if $\rho(\mathfrak{g}) \subseteq \mathfrak{n}_{\mathcal{F}}$. Use induction on $\dim(V)$. By Proposition 5.10, we see that the space $V^{\mathfrak{g}} \neq 0$. Thus by induction we may find a flag $\mathcal{F}'$ in $V/V^{\mathfrak{g}}$ which $\mathfrak{g}$ respects. Taking its preimage and extending arbitrarily (by picking any complete flag in $V^{\mathfrak{g}}$), we get a complete flag in V that $\mathfrak{g}$ clearly respects as required. □

6. Solvable Lie algebras

We now consider another class of Lie algebras which is slightly larger than the class of nilpotent algebras. For a Lie algebra $\mathfrak{g}$, let $D^0\mathfrak{g} = \mathfrak{g}$, and $D^{i+1}\mathfrak{g} = [D^i\mathfrak{g}, D^i\mathfrak{g}]$. $D^i\mathfrak{g}$ is the i-th derived ideal of $\mathfrak{g}$. Note that $C^1(\mathfrak{g}) = D^1\mathfrak{g}$ is $D\mathfrak{g}$ the derived subalgebra of $\mathfrak{g}$.

Definition 6.1. A Lie algebra $\mathfrak{g}$ is said to be *solvable* if $D^N\mathfrak{g} = 0$ for some $N > 0$.

Since it is clear from the definition that $D^i\mathfrak{g} \subset C^i(\mathfrak{g})$, any nilpotent Lie algebra is solvable, but as one can see by considering the non-abelian 2-dimensional Lie algebra, there are solvable Lie algebras which are not nilpotent.

Example 6.2. Let V be a finite dimensional vector space and $\mathcal{F} = (0 = F_0 < F_1 < \ldots < F_n = V)$ a complete flag in V. Let

$$\mathfrak{b}_{\mathcal{F}} = \{x \in \mathfrak{gl}(V) : x(F_i) \subseteq F_i\},$$

that is, $\mathfrak{b}_{\mathcal{F}}$ is the subspace of endomorphisms which preserve the complete flag $\mathcal{F}$. We claim that $\mathfrak{b}_{\mathcal{F}}$ is solvable. Since any nilpotent Lie algebra is solvable, and clearly $\mathfrak{b}_{\mathcal{F}}$ is solvable if and only if $D^1\mathfrak{b}_{\mathcal{F}}$ is, the solvability of $\mathfrak{g}$ will follow if we can show that $D^1\mathfrak{b}_{\mathcal{F}} \subseteq \mathfrak{n}_{\mathcal{F}}$. To see this, 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 End(F_i/F_{i-1}), which since F_i/F_{i-1} is one-dimensional, is abelian, so that all commutators are zero.

If we pick a basis $\{e_1, e_2, \ldots, e_n\}$ of V such that $F_i = \text{span}(e_1, \ldots, e_i)$, 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.

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 6.3. *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 $\mathfrak{g}$.*

[11] Though because of the following theorem this terminology won't be used much.

(3) *If im(ϕ) and ker(ϕ) 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. □

Note that, as mentioned above, the part 3) of the above Lemma is *false* for nilpotent Lie algebras.

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

Lemma 6.4. *(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 $W_m = \mathrm{span}\{v, x(v), \dots, x^m(v)\}$. 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 $\mathrm{tr}([x, h]) = (n + 1)\lambda([x, h])$. Since the trace of a commutator is zero[12], it follows that $(n + 1)\lambda([x, h]) = 0$, and so since $\mathrm{char}(\mathsf{k}) = 0$ we conclude that $\lambda([x, h]) = 0$. □

Theorem 6.5. *(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 $x \in \mathfrak{g}$.*

Proof. 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 $\dim(\mathfrak{g}) > 1$, consider the derived subalgebra $D^1\mathfrak{g}$. Since $\mathfrak{g}$ is solvable, $D^1\mathfrak{g}$ is a proper ideal of $\mathfrak{g}$. The quotient $\mathfrak{g}/D^1\mathfrak{g}$ is abelian, and taking the preimage of any codimension one subspace of it gives a codimension 1 ideal I of $\mathfrak{g}$. By induction we may pick a homomorphism $\lambda\colon I \to \mathfrak{gl}_1(\mathsf{k})$ such that the subspace $U = \{w \in V : h(w) = \lambda(h).w, \forall h \in I\}$ is nonzero. Now if $x \in \mathfrak{g}$, then

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

where in the second equality we used Lie's Lemma. Thus $\mathfrak{g}$ preserves U. Now since I is codimension one in $\mathfrak{g}$, we may write $\mathfrak{g} = \mathsf{k}x \oplus I$ for some $x \in \mathfrak{g}$. Taking an eigenvector $v \in U$ of x completes the proof. □

The analogue of Corollary 5.13 for solvable Lie algebras is the following:

Corollary 6.6. *Let $\mathfrak{b} \subset \mathfrak{gl}(V)$ be a solvable subalgebra of $\mathfrak{gl}(V)$. Then there is a complete flag $\mathcal{F}$ of V such that $\mathfrak{b} \subseteq \mathfrak{b}_{\mathcal{F}}$.*

Proof. By induction on $\dim(V)$, where Lie's theorem provides the induction step. □

Note that this theorem shows that if $\mathfrak{g}$ is a solvable Lie algebra, then any irreducible representation of $\mathfrak{g}$ is one-dimensional. Since conversely any one-dimensional representation of $\mathfrak{g}$ is clearly irreducible, we can rephrase Lie's theorem as the statement that any irreducible representation of a solvable Lie algebra is one-dimensional.

More generally, it is natural to note the following simple Lemma.

[12]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.

Lemma 6.7. *Let $\mathfrak{g}$ be a Lie algebra. The one-dimensional representations of $\mathfrak{g}$ are parametrized by the vector space $(\mathfrak{g}/D\mathfrak{g})^*$ as follows: if (V, ρ) is a a representation of $\mathfrak{g}$ and $\dim(V) = 1$, then $\rho\colon \mathfrak{g} \to End(V) \cong \mathsf{k}$, and $\rho(D\mathfrak{g}) = 0$, so that ρ induces a linear functional $\alpha\colon \mathfrak{g}/D\mathfrak{g} \to \mathsf{k}$.*

Proof. If $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$, then since $\dim(V) = 1$, the associative algebra End(V) is commutative, so that $\mathfrak{gl}(V)$ is abelian. Thus

$$\rho([x, y]) = \rho(x)\rho(y) - \rho(y)\rho(x) = 0, \quad (\forall x, y \in \mathfrak{g}),$$

hence $\rho(D\mathfrak{g}) = 0$, and ρ induces $\alpha \in (\mathfrak{g}/D\mathfrak{g})^*$ as claimed. Conversely, given $\alpha\colon (\mathfrak{g}/D\mathfrak{g})^*$, the map $\rho\colon \mathfrak{g} \to \mathfrak{gl}_1(\mathsf{k})$ given by $x \mapsto \alpha(x + D\mathfrak{g})$ is a Lie algebra homomorphism, because

$$\rho([x, y]) = 0 = \alpha(x + D\mathfrak{g})\alpha(y + D\mathfrak{g}) - \alpha(y + D\mathfrak{g})\alpha(x + D\mathfrak{g}).$$

□

An isomorphism classes of one-dimensional representations of $\mathfrak{g}$ are thus given by the space of linear functionals $(\mathfrak{g}/D\mathfrak{g})^*$. We will refer to an element $\alpha \in (\mathfrak{g}/D\mathfrak{g})^*$ as a *weight* of $\mathfrak{g}$. They should be thought of as the generalization of the notion of an eigenvalue of a linear map. Lie's theorem shows that if $\mathfrak{g}$ is a solvable Lie algebra (hence in particular if $\mathfrak{g}$ is nilpotent), then any irreducible representation is one-dimensional, so that $(\mathfrak{g}/D\mathfrak{g})^*$ parametrizes the irreducible representations of $\mathfrak{g}$. For $\alpha \in (\mathfrak{g}/D\mathfrak{g})^*$ let us write k_α for the representation (k, α) of $\mathfrak{g}$ on the field k given by α.

7. Representations of nilpotent Lie algebras

In this section we assume that k *is an algebraically closed field of characteristic zero.*

A representation (ρ, V) of the one-dimensional Lie algebra $\mathfrak{gl}_1(\mathsf{k})$ on a k-vector space V is given, once we chose a basis vector e of $\mathfrak{gl}_1(\mathsf{k})$, by a single linear map $\phi\colon V \to V$ via the correspondence $\phi = \rho(e)$. Thus the classification of representations of $\mathfrak{gl}_1(\mathsf{k})$ is equivalent to the classification of linear endomorphisms.[13] This classification is of course given by the Jordan normal form (at least over an algebraically closed field). In this section we will see that a (slightly weaker) version of this classification holds for representations of any nilpotent Lie algebra.

We begin by reviewing some linear algebra. Let $x\colon V \to V$ be a linear map. Let V_λ be the *generalized eigenspaces* for x with eigenvalue λ:

$$V_\lambda = \{v \in V : \exists N > 0, (x - \lambda)^N(v) = 0\}.$$

(thus V_λ is zero unless λ is an eigenvalue of x).

Lemma 7.1. *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 $\phi\colon \mathsf{k}[t] \to$ End(V) given by $t \mapsto x$ denotes the natural map, we have $\mathsf{k}[t]/(m_x) \cong \text{im}(\phi) \subseteq \text{End}(V)$. 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

$$\text{im}(\phi) \cong \mathsf{k}[t]/(m_x) \cong \bigoplus_{i=1}^k \mathsf{k}[t]/(t - \lambda_i)^{n_i},$$

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 $U_i = \text{im}(e_i)$, 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 im(ϕ).

It thus remains to check that $U_i = V_{\lambda_i}$. Since $(t - \lambda_i)^{n_i} e_i = 0 \in \mathsf{k}[t]/(m_x)$, it is clear that $U_i \subseteq V_{\lambda_i}$. To see the reverse inclusion, suppose that $v \in V_{\lambda_i}$ so that, say, $(x - \lambda_i)^n(v) = 0$ for some $n > 0$. Write $v = v_1 + \ldots + v_k$, where $v_j \in U_j$. Now since $(x - \lambda_i)^n(v) = 0$, and each U_j is stable under im(ϕ) it follows that $(x - \lambda_i)^n(v_j) = 0$ for each j, $(1 \leq j \leq k)$. If $1 \leq l \leq k$ and $l \neq i$, then since $(t - \lambda_i)^n$ and $(t - \lambda_l)^{n_l}$ are coprime, we may find $a, b \in \mathsf{k}[t]$ such that $a.(t - \lambda_i)^n + b.(t - \lambda_l)^{n_l} = 1$.

Setting $t = x$ in this equation, and applying the result to the vector v_l we find:

[13]Essentially the same is true for representations of the abelian group $\mathbb{Z}$.

$$\begin{aligned} v_l = 1.v_l &= \big(a(x).(x-\lambda_i)^n + b(x).(x-\lambda_l)^{n_l}\big)(v_l) \\ &= a(x)(t-\lambda_i)^n(v_l) + b(x)(t-\lambda_l)^{n_l}(v_l) \\ &= 0, \end{aligned} \tag{7.1}$$

where the first term is zero by the above, and the second term is zero since $v_l \in U_l$. It follows that $v = v_i \in U_i$ and so $V_{\lambda_i} \subseteq U_i$ as required.

□

Lemma 7.2. i) *Let V be a vector space and write $A = Hom_{\mathsf{k}}(V,V)$ for the associative algebra of linear maps from V to itself. Then we have:*

$$(x-\lambda-\mu)^n y = \sum_{i=0}^{n} \binom{n}{i} (ad(x)-\lambda)^i(y)(x-\mu)^{n-i}, \quad (\forall x, y \in A),$$

where as usual $ad(x)(y) = [x,y] = xy - yx$.

ii) *Let $\mathfrak{g}$ be a Lie algebra and (V,ρ) a representation of $\mathfrak{g}$. Then for all $x, y \in \mathfrak{g}$ we have*

$$(\rho(x)-\lambda-\mu)^n \rho(y) = \sum_{i=0}^{n} \binom{n}{i} \rho((ad(x)-\lambda)^i(y))(\rho(x)-\mu)^{n-i}. \tag{7.2}$$

Proof. Note first that in any associative algebra A, if $a, b \in A$ commute, *i.e.* $ab = ba$, then the binomial theorem shows that $(a+b)^n = \sum_{i=0}^n \binom{n}{i} a^i b^{n-i}$.

Let $r_x \colon A \to A$ be the map given by $r_x(y) = yx$, right-multiplication by x, and similarly let $l_x \colon A \to A$ be the map $l_x(y) = xy$ be the map given by left-multiplication by x. The $\text{ad}(x) = l_x - r_x$. Then clearly $l_x, r_x, \lambda.\text{id}_V, \mu.\text{id}_V$ all commute, and we have:

$$\begin{aligned} (x-\lambda-\mu)^n y &= (l_x - \lambda - \mu)^n(y) \\ &= (\text{ad}(x) + r_x - \lambda - \mu)(y) \\ &= ((\text{ad}(x)-\lambda) + (r_x - \mu))(y) \\ &= \sum_{i=0}^{n} \binom{n}{i} (\text{ad}(x)-\lambda)^i (r_x-\mu)^{n-i}(y) \\ &= \sum_{i=0}^{n} \binom{n}{i} (\text{ad}(x)-\lambda)^i(y)(x-\mu)^{n-i}. \end{aligned}$$

where we apply the binomial theorem to the (commuting) pair $(\text{ad}(x)-\lambda, r_x - \mu)$ to the fourth line. For the second part, simply apply the first part to $\rho(x), \rho(y)$, using the fact that $\text{ad}(\rho(x))(\rho(y)) = \rho(\text{ad}(x)(y))$. □

Theorem 7.3. *Let $\mathfrak{h}$ be a nilpotent Lie algebra and (V,ρ) a finite dimensional representation of $\mathfrak{h}$. For each $\lambda \in (\mathfrak{h}/D\mathfrak{h})^*$, set*

$$V_\lambda = \{v \in V : \textit{for all } x \in \mathfrak{h}, \exists n > 0 \textit{ such that } (\rho(x) - \lambda(x))^n(v) = 0\}.$$

Then each V_λ is a subrepresentation and

$$V = \bigoplus_\lambda V_\lambda$$

that is, V is the direct sum of subrepresentations indexed by the one-dimensional representations occuring in V.

Proof. We use induction on $\dim(V)$. If $\dim(V) = 1$ the result is trivial. Next note that if $V = U \oplus W$ is a direct sum of proper subrepresentations, then applying induction we know the result holds for U and W respectively, and clearly $V_\lambda = U_\lambda \oplus W_\lambda$, so that we may conclude the result holds for V.

Now suppose that $x \in \mathfrak{h}$ and let $V = \bigoplus V_{\lambda(x)}$ denote the generalised eigenspace decomposition of V for $\rho(x)$. We claim each subspace $V_{\lambda(x)}$ is a subrepresentation of $\mathfrak{h}$. Indeed since $\mathfrak{h}$ is a nilpotent Lie algebra, the map $\text{ad}(x)$ is nilpotent, hence if $k > \dim(\mathfrak{h})$ then $\text{ad}(x)^k(y) = 0$ for all $y \in \mathfrak{h}$. It follows that if $v \in V_{\lambda(x)}$, so that $(\rho(x) - \lambda(x))^k(v) = 0$ for all $k \geq N$ say, then if $k \geq N + \dim(\mathfrak{h})$, we see from equation (7.2) (with $\mu = 0$) that $(\rho(x) - \lambda(x))^k(\rho(y)(v)) = 0$, since we must have either $i \geq \dim(\mathfrak{h})$ or $k - i \geq N$ if $k \geq N + \dim(\mathfrak{h})$, and so $\rho(y)(v) \in V_{\lambda(x)}$ as required.

It follows that if there is some $x \in \mathfrak{h}$ which has more than one eigenvalue, then the decomposition $V = \bigoplus V_{\lambda(x)}$ shows that V is a direct sum of proper subrepresentations, and we are done by induction. Thus we are reduced to the case where for all $x \in \mathfrak{h}$ the linear map $\rho(x)$ has a single eigenvalue $\lambda(x)$

say. But then $x \mapsto \lambda(x)$ must be an element of $(\mathfrak{h}/D\mathfrak{h})^*$ by Lie's theorem: Pick a one-dimensional subrepresentation L of V, then $\rho(x)$ acts by a scalar $\mu(x)$ on L, where $\mu \in (\mathfrak{h}/D\mathfrak{h})^*$ and since $\lambda(x)$ is the only eigenvalue of $\rho(x)$ by assumption we must have $\lambda(x) = \mu(x)$. Now clearly $V = V_\lambda$ and we are done.

□

Remark 7.4. The (nonzero) subrepresentations V_λ are known as the *weight spaces* of V, and the set $\Psi(V)$, sometimes written simply Ψ, of $\lambda \in (\mathfrak{h}/D\mathfrak{h})^*$ for which $V_\lambda \neq 0$ are known as the *weights* of V.

8. 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 8.1. We say that a subalgebra $\mathfrak{h}$ is 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 8.2. If $x \in \mathfrak{g}$, let $\mathfrak{g}_{0,x}$ be the generalized 0-eigenspace of $\mathrm{ad}(x)$, that is

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

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.

Lemma 8.3. (1) *If $x \in \mathfrak{g}$ is a 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. To see that $\mathfrak{h} = \mathfrak{g}_{0,x}$ is a subalgebra of $\mathfrak{g}$, use the formula

$$\mathrm{ad}(x)^n[y,z] = \sum_{k=0}^{n} \binom{n}{i} [\mathrm{ad}(x)^k(y), \mathrm{ad}(x)^{n-k}(z)]. \tag{8.1}$$

The formula can be established, for example, by an easy induction, or by applying (7.2) with $\lambda = \mu = 0$ to the adjoint representation.

Next we show 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 $\mathrm{ad}(x)^n([x,z]) = 0$, and hence $\mathrm{ad}(x)^{n+1}(z) = 0$ and $z \in \mathfrak{h}$ as required.

Assume now that x is regular. To show that the corresponding $\mathfrak{h}$ is nilpotent, we use Engel's theorem: we will show that for each $y \in \mathfrak{h}$ the map $\mathrm{ad}(y)$ is nilpotent as an endomorphism of $\mathfrak{h}$. To show $\mathrm{ad}(y)$ is nilpotent on $\mathfrak{h}$, we consider the characteristic polynomial of $\mathrm{ad}(y)$ on $\mathfrak{g}$ and $\mathfrak{h}$. Since $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, the characteristic polynomial $\chi^y(t) \in \mathsf{k}[t]$ of $\mathrm{ad}(y)$ on $\mathfrak{g}$ is the product of the characteristic polynomials of $\mathrm{ad}(y)$ 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 $\{h_1, h_2, \ldots, h_r\}$ a basis of $\mathfrak{h}$, so that we may write $y = \sum_{i=1}^{r} x_i h_i$. Then we may view the coefficients c_k as polynomials in the coordinates $\{x_i : 1 \leq i \leq r\}$. Similarly, we have $\chi_1^y = \sum_{k=0}^{r} d_k(y)t^k$and $\chi_2^y = \sum_{i=0}^{n-r} e_i(y)t^i$ where the d_i, e_j are polynomials in the $\{x_i : 1 \leq i \leq r\}$. Now we know that if $y = x$ then $\mathrm{ad}(x)$ is invertible on $\mathfrak{g}/\mathfrak{h}$, since it has no 0-eigenspace there, so that χ_2^x has $e_0(x) \neq 0$, and thus the polynomial e_0 is nonzero.

Now for each $y \in \mathfrak{h}$ the number $\min\{i : c_i(y) \neq 0\}$ is clearly $\dim(\mathfrak{g}_{0,y})$. However, suppose we write $\chi_1^y(t) = t^s \sum_{k=0}^{r-s} d_{k+s} t^k$, where $d_s \neq 0$ is a nonzero polynomial in the $\{x_i : 1 \leq i \leq r\}$. Then we would have

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

hence if we pick $z \in \mathfrak{h}$ such that $d_s(z)e_0(z)$ is nonzero, then $\mathfrak{g}_{0,z}$ has dimension s, so by the assumption that x is regular, we must have $s \geq r$. On the other hand $s \leq r$ since χ_1 has degree r, thus we conclude that $s = r$ and $\chi_1^y(t) = t^r$ 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 following fact about the coefficients of the characteristic polynomial. It 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[14]. (To apply it to the above, take $V = \mathfrak{g}$, $A = \mathfrak{h}$ and $\varphi = \text{ad}$).

Lemma 8.4. *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 $\text{End}(V)$ with $\text{Mat}_n(\mathsf{k})$ 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 8.5. 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 8.6. Although we will not prove it in this course, any two Cartan subalgebras of $\mathfrak{g}$ are conjugate by an automorphism[15] 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 $\alpha(\mathfrak{h}_1) = \mathfrak{h}_2$.

9. 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

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

The next Lemma establishes some basic properties of this decomposition.

Lemma 9.1. *Let $\mathfrak{g}, \mathfrak{h}$ be as above. Then $\mathfrak{h} = \mathfrak{g}_0$. Moreover, if λ, μ are one-dimensional representations, then*

$$[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subseteq \mathfrak{g}_{\lambda+\mu}.$$

Proof. Clearly $\mathfrak{h} \subseteq \mathfrak{g}_0$, since $\mathfrak{h}$ is nilpotent. Consider $\mathfrak{g}_0/\mathfrak{h}$ as an $\mathfrak{h}$-representation. If nonzero, then by Lie's theorem it contains a one-dimensional submodule L, and since $\mathfrak{h}$ acts nilpotently on $\mathfrak{g}_0$ be definition, $\mathfrak{h}$ acts as zero on L. But then the preimage of L in $\mathfrak{g}_0$ normalizes $\mathfrak{h}$ which contradicts the assumption that $\mathfrak{h}$ is a Cartan subalgebra.

Applying (7.2) with $\rho = \text{ad}$ we see that:

$$(\text{ad}(x) - (\lambda(x) + \mu(x))1)^n[y, z] = \sum_{i=0}^{n} \binom{n}{i} [(\text{ad}(x) - \lambda(x).1)^i(y), (\text{ad}(x) - \mu(x).1)^{n-i}(z)].$$

The containment $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subseteq \mathfrak{g}_{\lambda+\mu}$ then follows using the same argument as in the first paragraph of Theorem 7.3. □

---

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

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

Definition 9.2. 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[16] 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 8.6 above, the Cartan decomposition of $\mathfrak{g}$ is unique up to automorphism.

More generally, using Equation 7.2 in the same way as in the proof of Lemma 9.1 we can show:

Lemma 9.3. *Let $\mathfrak{g}, \mathfrak{h}$ be as above, and let V be a $\mathfrak{g}$-representation. As an $\mathfrak{h}$-representation, it then decomposes*

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

into generalised $\mathfrak{h}$-weight spaces as in Theorem 7.3, for a subset $\Psi \subset (\mathfrak{h}/D\mathfrak{h})^*$. Then for $\alpha \in \Phi$ a root, we have

$$[\mathfrak{g}_\alpha, V_\mu] \subseteq V_{\alpha+\mu}.$$

10. 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[17] is given in Appendix 1 of these notes.

10.1. Bilinear forms.

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

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

From the definition of tensor products it follows that $\text{Bil}(V)$ can be identified with $(V \otimes V)^*$. If V is a $\mathfrak{g}$-representation, this means $\text{Bil}(V)$ also has the structure of $\mathfrak{g}$-representation: explicitly, if $B \in \text{Bil}(V)$, 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 $y \in \mathfrak{g}$.

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

$$\begin{aligned} 0 &= y(B)(x,z) \\ &= -B(\text{ad}(y)(x), z) - B(x, \text{ad}(y)(z)), \\ &= -B([y,x], z) - B(x, [y,z]) \\ &= B([x,y], z) - B(x, [y,z]), \end{aligned}$$

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

Definition 10.1. We say that a bilinear form B is *invariant* if it is an invariant vector for the action of $\mathfrak{g}$ on $\text{Bil}(\mathfrak{g}) \cong (\mathfrak{g} \otimes \mathfrak{g})^*$, that is, if

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

[16] *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.

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

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 10.2. *Let V be a k-vector space. The trace form $b_V : \mathfrak{gl}(V) \otimes \mathfrak{gl}(V) \to \mathsf{k}$ given by $b_V(a.b) = tr(a \circ b)$ $(a, b \in \mathfrak{gl}(V))$ is an invariant symmetric bilinear form on $\mathfrak{gl}(V)$.*

Proof. It is clear that b_V is bilinear and it is symmetric because $\mathrm{tr}(a.b) = \mathrm{tr}(b.a)$. To see that it is invariant, note that, for $a, b, c \in \mathfrak{gl}(V)$ we have

$$\begin{aligned} \mathrm{tr}([a,b].c) &= \mathrm{tr}((ab-ba).c) = \mathrm{tr}(a.(bc)) - \mathrm{tr}(b.(ac)) \\ &= \mathrm{tr}(a.(bc)) - \mathrm{tr}((ac).b) \\ &= \mathrm{tr}(a.(bc-cb) \\ &= \mathrm{tr}(a,[b,c]). \end{aligned}$$

where going from the first to the second line we used the symmetry property of tr to replace $\mathrm{tr}(b.(ac))$ with $\mathrm{tr}((ac).b)$. □

Remark 10.3. The invariance of the form b_V is just the condition that the map $b_V : \mathfrak{gl}(V) \otimes \mathfrak{gl}(V) \to \mathsf{k}$ is a map of $\mathfrak{gl}(V)$-representations, where k is viewed as the trivial representation of $\mathfrak{gl}(V)$. Using this, one can also prove the previous Lemma using the identification of $\mathfrak{gl}(V) \cong V^* \otimes V$ as a $\mathfrak{g}$-representation, along with the description of the trace map in terms of the contraction map ι. Let $d\colon V^* \otimes V \to (V^* \otimes V)^*$ be the map given by $m \circ \sigma \circ (1 \otimes e_V)$, where $e_V : V \to V^{**}$ is the natural map, $\sigma\colon V^* \otimes V^{**} \to V^{**} \otimes V^*$ is the natural isomorphism, and m is the multiplication map (see § 22.1). Then, writing $\iota_U : U \otimes U^* \to \mathsf{k}$ for the contraction map we have

$$\mathrm{tr}(a.b) = \iota_{V^* \otimes V}(a \otimes d(b))$$

Since all of the maps m, σ, e_V and ι are maps of $\mathfrak{gl}(V)$-representations, it follows that tr is also.

Definition 10.4. If $\mathfrak{g}$ is a Lie algebra, and let (V, ρ) be a representation of $\mathfrak{g}$. we may define a bilinear form $t_V : \mathfrak{g} \times \mathfrak{g} \to \mathsf{k}$ 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 10.5. 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 10.6. *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:*

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

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. □

The Killing form also allows us to produce ideals: If $\mathfrak{a}$ denotes a subspace of $\mathfrak{g}$, then we will write $\mathfrak{a}^\perp$ for the subspace

$$\{x \in \mathfrak{g} : \kappa(x,y) = 0, \forall y \in \mathfrak{a}\}.$$

Lemma 10.7. *Let $\mathfrak{g}$ be Lie algebra and let $\mathfrak{a}$ be an ideal of $\mathfrak{g}$. Then $\mathfrak{a}^\perp$ is also an ideal of $\mathfrak{g}$.*

Proof. Suppose that $x \in \mathfrak{g}$ and $z \in \mathfrak{a}^\perp$. We need to show that $[x, z] \in \mathfrak{a}^\perp$. But if $y \in \mathfrak{a}$ we have

$$\kappa([x, z], y) = -\kappa([z, x], y) = -\kappa(z, [x, y]) = 0,$$

since $[x, y] \in \mathfrak{a}$ since $\mathfrak{a}$ is an ideal. Hence $[x, z] \in \mathfrak{a}^\perp$ as required. □

11. 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 11.1. *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. 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 9.3, 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 $\text{tr}(\rho(z), M) = 0$. On the other hand, we may also compute the trace of $\rho(z)$ on M directly:

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

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). □

Proposition 11.2. *Let $\mathfrak{g}$ be a Lie algebra. If $\mathfrak{g} = D\mathfrak{g}$ then there is an $x \in \mathfrak{g}$ such that $\kappa(x, x) \neq 0$.*

Proof. Pick $\mathfrak{h}$ a Cartan subalgebra of $\mathfrak{g}$ and let $\mathfrak{g} = \mathfrak{h} \oplus_{\alpha in \Phi} \mathfrak{g}_\alpha \mathfrak{g}_\alpha$, where Φ denotes the roots of $\mathfrak{g}$ and $\mathfrak{h} = \mathfrak{g}_0$. Now if $\mathfrak{g} = \mathfrak{h}$, then $\mathfrak{g}$ is nilpotent, hence solvable, which is impossible since $\mathfrak{g} = D\mathfrak{g}$ implies the derived series of $\mathfrak{g}$ has $D^k\mathfrak{g} = \mathfrak{g}$ for all k. Thus Φ is non-empty.

Next observe that

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

Since $[\mathfrak{g}_{,\alpha}, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$, and $\mathfrak{h} = \mathfrak{g}_0$, we must have

$$\mathfrak{h} = [\mathfrak{h}, \mathfrak{h}] + \sum_{\beta \in \Phi_\pm} [\mathfrak{g}_\beta, \mathfrak{g}_{-\beta}] = D\mathfrak{h} + \sum_{\beta \in \Phi_\pm} [\mathfrak{g}_\beta, \mathfrak{g}_{-\beta}]$$

where $\Phi_\pm$ is set of all $\beta \in \Phi$ for which $\{\pm\beta\} \subseteq \Phi$. Now by definition, each $\gamma \in \Phi$ has $\gamma(D\mathfrak{h}) = 0$, it follows that if $\alpha \in \Phi$, as $\alpha \neq 0$ there must be some $\beta_0 \in \Phi_\pm$ on which $\alpha([\mathfrak{g}_{\beta_0}, \mathfrak{g}_{-\beta_0}]) \neq \{0\}$ (in particular, $\Phi_\pm \neq \emptyset$). Then

$$\kappa(x, x) = \text{tr}(\text{ad}(x)^2) = \text{tr}_\mathfrak{h}(\text{ad}(x)^2) + \sum_{\beta \in \Phi} \text{tr}_{\mathfrak{g}_\beta}(\text{ad}(x)^2).$$

Now since $\mathfrak{h}$ is nilpotent, $\text{ad}(x)$ is nilpotent on $\mathfrak{h}$ and so $\text{tr}_\mathfrak{h}(\text{ad}(x)^2) = 0$. On each $\mathfrak{g}_\beta$, we can choose a basis such that the matrix associated to $\text{ad}(x)$ acting $\mathfrak{g}_\beta$ is upper triangular with $\beta(x)$ on the diagonal.

It follows that $\text{tr}_{\mathfrak{g}_\beta}(\text{ad}(x)^2) = \dim(\mathfrak{g}_\beta).\beta(x)^2$. Now by Lemma 11.1, the restriction of β to $[\mathfrak{g}_{\beta_0}, \mathfrak{g}_{-\beta_0}]$ is equal to $r_\beta.\beta_0$ for some rational number $r_\beta \in \mathbb{Q}$. But then it follows that

$$\kappa(x,x) = \left(\sum_{\beta\in\Phi} \dim(\mathfrak{g}_\beta) r_\beta^2\right) \beta_0(x)^2 \tag{11.1}$$

Since by assumption, $0 \neq \alpha(x) = r_\alpha.\beta_0(x)$, so that both r_α and $\beta_0(x)$ are nonzero. Now every term in the sum on the right-hand side of (11.1) is a non-negative rational number, and $r_\alpha^2 > 0$ (as we are working over a field of characteristic zero, it contains $\mathbb{Q}$) and hence $\kappa(x,x) \neq 0$ as required. □

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

Theorem 11.3 (*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. Consider the derived series $D^k\mathfrak{g}$, $(k \geq 1)$. If there is some k with $D^k\mathfrak{g} = D^{k+1}\mathfrak{g} = D(D^k\mathfrak{g}) \neq \{0\}$, then by Lemma 10.6 and Proposition 11.2 applied to $D^k(\mathfrak{g})$,, there is an $x \in D^k\mathfrak{g}$ with

$$\kappa^{\mathfrak{g}}(x,x) = \kappa^{D^k\mathfrak{g}}(x,x) \neq 0,$$

and hence κ is not identically zero. 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, if $\mathfrak{g}$ is solvable, then by Lie's theorem we can find a filtration $0 \subset F_1 \subset F_2 \subset \ldots \subset F_n = \mathfrak{g}$ of ad-stable subspaces with $\dim(F_i) = i$. Now if $x, y \in \mathfrak{g}$ the maps induced by $\text{ad}(x)\text{ad}(y)$ and $\text{ad}(y)\text{ad}(x)$ on F_i/F_{i-1} are equal, since $\mathfrak{gl}_1$ is commutative. Thus it follows that if $z = [x,y] \in D\mathfrak{g}$, then $\text{ad}(z) = [\text{ad}(x), \text{ad}(y)]$ maps F_i into F_{i-1} and hence for all $z \in D\mathfrak{g}$ we have $\text{ad}(z)(F_i) \subset F_{i-1}$. But now if $z_1, z_2 \in D\mathfrak{g}$, then $\text{ad}(z_1)\text{ad}(z_2)(F_i) \subseteq F_{i-2}$, and hence $\text{ad}(z_1)\text{ad}(z_2)$ is nilpotent, and so $\kappa(z_1, z_2) = \text{tr}(\text{ad}(z_1)\text{ad}(z_2)) = 0$ as required.

□

12. Semisimple Lie algebras and semi-direct products

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 $\mathfrak{r}$ (in the strong sense: every solvable ideal of $\mathfrak{g}$ is a subalgebra of $\mathfrak{r}$).

Definition 12.1. Let $\mathfrak{g}$ be a finite dimensional Lie algebra. The largest solvable ideal $\mathfrak{r}$ of $\mathfrak{g}$ is known as the *radical* of $\mathfrak{g}$, and will be denoted $\text{rad}(\mathfrak{g})$. We say that $\mathfrak{g}$ is *semisimple* if $\text{rad}(\mathfrak{g}) = 0$, that is, if $\mathfrak{g}$ contains no non-zero solvable ideals.

Lemma 12.2. *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 $\mathfrak{g}/\text{rad}(\mathfrak{g})$. 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 $\text{rad}(\mathfrak{g})$ and $\mathfrak{s} = \mathfrak{s}'/\text{rad}(\mathfrak{g})$ as both solvable. But then by definition we have $\mathfrak{s}' \subseteq \text{rad}(\mathfrak{g})$ so that $\mathfrak{s}' = \text{rad}(\mathfrak{g})$ and $\mathfrak{s} = 0$ as required. □

Thus we have shown that any Lie algebra $\mathfrak{g}$ contains a canonical solvable ideal $\text{rad}(\mathfrak{g})$ such that $\mathfrak{g}/\text{rad}(\mathfrak{g})$ is a semisimple Lie algebra. So, in some sense at least, every finite dimensional Lie algebra is "built up" out of a semisimple Lie algebra and a solvable one.

Definition 12.3. Suppose that $\mathfrak{g}$, $\mathfrak{h}$ are Lie algebras, and we have a homomorphism $\phi\colon \mathfrak{g} \to \text{Der}_{\mathsf{k}}(\mathfrak{h})$, the Lie algebra of derivations[18] 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*[19] of $\mathfrak{g}$ and $\mathfrak{h}$ by ϕ which as a vector space is just $\mathfrak{h} \oplus \mathfrak{g}$, 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 $\mathfrak{h} \rtimes \mathfrak{g}$.

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

[19]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.

In fact, semidirect products correspond to a relatively simple kind of extension; see Appendix 3.

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

Theorem 12.4. *(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 semisimple subalgebra $\mathfrak{s}$ of $\mathfrak{g}$ such that $\mathfrak{g} \cong \mathfrak{r} \rtimes \mathfrak{s}$.*

Note that in particular $\mathfrak{s}$ is isomorphic to $\mathfrak{g}/\text{rad}(\mathfrak{g})$. We will not prove this theorem in this course.

13. Cartan's Criterion for semisimplicity

The Killing form gives us a way of detecting when a Lie algebra is semisimple. Recall that a bilinear form $B\colon V \times V \to \mathsf{k}$ is said to be nondegenerate if $\{v \in V : \forall w \in V, B(v,w) = 0\} = \{0\}$. We first note the following simple result.

Lemma 13.1. *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 $\text{rad}(\mathfrak{g}) \neq 0$. 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 13.2. *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 10.7 $\mathfrak{g}^\perp$ 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 10.6 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. Then if $x \in \mathfrak{a}$ and $y \in \mathfrak{g}$ is arbitrary, the composition $\text{ad}(x)\text{ad}(y)\text{ad}(x)$ must be zero, since $\text{ad}(y)\text{ad}(x)(z) \in \mathfrak{a}$ for any $z \in \mathfrak{g}$, as $\mathfrak{a}$ is an ideal, and since $\mathfrak{a}$ is abelian $\text{ad}(a)(b) = 0$ for all $a, b \in \mathfrak{a}$. But then clearly $(\text{ad}(y)\text{ad}(x))^2 = 0$, so that $\text{ad}(y)\text{ad}(x)$ is nilpotent and $\kappa(x,y) = 0$ for all $y \in \mathfrak{g}$. But then $\mathfrak{a} \subseteq \mathfrak{g}^\perp = \{0\}$ and $\mathfrak{a} = \{0\}$ as required. □

14. Simple and semisimple Lie algebras

Definition 14.1. 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 14.2. *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 22.11 in Appendix 1, we have

$$\dim(I) + \dim(I^\perp) = \dim(\mathfrak{g}). \tag{14.1}$$

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 (14.1) 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 14.3. *Let $\mathfrak{g}$ be a semisimple Lie algebra.*

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

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

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 $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \ldots \oplus \mathfrak{g}_k$ 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

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

(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[20].

□

15. 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. Recall from Section 4 the definition of an irreducible representation. Appendix 1 in Section 23 reviews 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 a direct sum of irreducibles.

If (V, ρ) and (W, σ) are representations of $\mathfrak{g}$, then a linear map $\phi \colon V \to W$ is a *homomorphism* of representations if

$$\phi(\rho(x)(v)) = \sigma(x)(\phi(v)) \quad \forall x \in \mathfrak{g}, v \in V.$$

We will write $\mathrm{Hom}_{\mathfrak{g}}(V, W)$ for the space of homomorphisms from V to W. To make the notation less cluttered, where there is no danger for confusion we will often suppress the notation for the map $\rho \colon \mathfrak{g} \to \mathfrak{gl}(V)$, so that, for example, the condition for ϕ to be a homomorphism will be written simply as $\phi(x(v)) = x(\phi(v))$.

The following Lemma was on a problem sheet, and is proved in Appendix 1 (the proof is identical to the corresponding results for representations of groups). For it we suppose that our field k is algebraically closed (but not necessarily of characteristic zero).

Lemma 15.1. *(Schur's Lemma). Let (V, ρ) and (W, σ) be irreducible representations of $\mathfrak{g}$. Then if $\phi \colon V \to W$ is a homomorphism of $\mathfrak{g}$-representations, ϕ is either zero or an isomorphism. Moreover, if k is algebraically closed, then $\mathrm{Hom}_{\mathfrak{g}}(V, W)$ is one-dimensional.*

Definition 15.2. We say that a representation V of a Lie algebra is *semisimple* if any subrepresentation has a complement, that is, if U is a subrepresentation of V then there is a subrepresentation W such that $V = U \oplus W$.

Lemma 15.3. *Let V be a representation of $\mathfrak{g}$.*

i) *V is semisimple if any surjection of representations $q \colon V \to W$ has a right inverse.*
ii) *If the representations of $\mathfrak{g}$ are semisimple then they are completely reducible.*

Proof. For the first part, if U is a subrepresentation of V, then consider the quotient map $q \colon V \to V/U$. If s is a right inverse for q, then we claim $T = \mathrm{im}(s)$ is a complement to U. Certainly if $w \in U \cap T$ the $q(w) = 0$, since U is the kernel of q. But then as $w \in T$ we have $w = s(u)$ for some $u \in V/U$, and hence $u = q(s(u)) = q(w) = 0$, so that $w = s(0) = 0$. Hence U and T form a direct sum, and since s is clearly injective, by dimension it $U \oplus T = V$ as required.

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 15.4. In the problem sets it is checked that if a representation is semisimple then any subrepresentation and any quotient representation of it is 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.)

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

Remark 15.5. Given a surjective map of $\mathfrak{g}$-representations $q\colon V \to W$, then a right inverse $s\colon W \to V$ as above is also often called a *section* of q, and we say that s *splits* the map q.

Recall from Section 10 the definition of the trace form t_V associated to a representation (V, ρ).

Lemma 15.6. *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 $\rho(\mathfrak{g}) \subseteq \mathfrak{gl}(V)$ 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 $\rho(\mathfrak{g})$. 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 $\rho(\mathfrak{g})$ is semisimple, this forces the radical to be zero as required. □

15.1. **Casimir elements.** In this section we construct, for representations V of a Lie algebra $\mathfrak{g}$ whose trace form is nondegenerate, a non-zero $\mathfrak{g}$-homomorphism from V to itself. We then show that if $\mathfrak{g}$ is semisimple, this construction can be applied whenever $\ker(\rho) \neq \mathfrak{g}$. We being with a basic fact which we essentially saw in our examples of Lie algebras at the start of the course.

Lemma 15.7. *If V is a vector space then the map $ad\colon \mathfrak{gl}(V) \to \mathfrak{gl}(\mathfrak{gl}(V))$ given by $ad(a)(x) = ax - xa$ has image in $Der_\mathsf{k}(End_\mathsf{k}(V))$, that is $ad(a)(x.y) = ad(a).x + x.ad(a)(y)$. It follows that the composition map $\mathfrak{gl}(V) \times \mathfrak{gl}(V) \to \mathfrak{gl}(V)$ induces a map of $\mathfrak{gl}(V)$-representations*

$$m\colon \mathfrak{gl}(V) \otimes \mathfrak{gl}(V) \to \mathfrak{gl}(V), \quad x \otimes y \mapsto x \circ y$$

Proof. The first statement is an easy calculation:

$$\begin{aligned} \mathrm{ad}(a)(x).y + x.\mathrm{ad}(a)(y) &= (ax - xa).y + x(ay - ya) \\ &= a.xy - xay + xay - xy.a \\ &= \mathrm{ad}(a)(x.y). \end{aligned}$$

For the second part of the Lemma, one simply notes that for $a, x, y \in \mathfrak{gl}(V)$ we have

$$\begin{aligned} m(a.(x \otimes y)) &= m(\mathrm{ad}(a)(x) \otimes y + x \otimes \mathrm{ad}(a)(y)) \\ &= \mathrm{ad}(a)(x).y + x.\mathrm{ad}(a)(y) \\ &= \mathrm{ad}(a)(x.y) = \mathrm{ad}(a)(m(x \otimes y)). \end{aligned}$$

so that m is a homomorphism of $\mathfrak{gl}(V)$-representations. □

Definition 15.8. Suppose that (V, ρ) is a $\mathfrak{g}$-representation and that its trace form $\beta = t_V$ is non-degenerate. Then β induces an isomorphism $\tau\colon \mathfrak{g} \to \mathfrak{g}^*$, and hence we have a sequence of homomorphisms of $\mathfrak{g}$-representations

$$\mathrm{Hom}_\mathsf{k}(\mathfrak{g}, \mathfrak{g}) \xrightarrow{\theta^{-1}} \mathfrak{g}^* \otimes \mathfrak{g} \xrightarrow{\tau^{-1}\otimes 1} \mathfrak{g} \otimes \mathfrak{g} \xrightarrow{\rho\otimes\rho} \mathfrak{gl}(V) \otimes \mathfrak{gl}(V) \xrightarrow{m} \mathfrak{gl}(V).$$

where the first map is the inverse of the map θ from Lemma 22.4, and the map $m\colon \mathfrak{gl}(V)\otimes\mathfrak{gl}(V) \to \mathfrak{gl}(V)$ is given by the composition of the map of the previous Lemma. (Note that the first two maps are isomorphisms.)

As the identity element $\mathrm{id}_\mathfrak{g}$ is clearly an invariant vector in $\mathrm{Hom}_\mathsf{k}(\mathfrak{g}, \mathfrak{g})$, applying to it the above sequence of $\mathfrak{g}$-homomorphisms yields an invariant vector C_V in $\mathrm{Hom}_\mathsf{k}(V, V)$, that is, a $\mathfrak{g}$-homomorphism from V to itself. We call this the *Casimir operator*.

We can make the Casimir operator explicit as follows: Pick a basis $\{x_1), \ldots, x_n)\}$ of $\mathfrak{g}$. The nondegeneracy of the form t_V on $\mathfrak{g}$ implies that there is a unique basis $\{y_1, \ldots, y_n\}$ which is dual to the basis $\{x_1, \ldots, x_n\}$ in the sense that $t_V(x_i), y_j)) = 1$ if $i = j$ and is 0 otherwise. If $\tau\colon \mathfrak{g} \to \mathfrak{g}^*$ the it is clear that $y_) = \tau^{-1}(\delta_i)$, where $\{\delta_i : 1 \leq i \leq n\}$ denotes the basis of $\mathfrak{g}^*$ dual to the basis $\{x_1, \ldots, x_n\}$.

Now the identity element of $\mathrm{Hom}_\mathsf{k}(\mathfrak{g}), \mathfrak{g})$ corresponds to the element $\sum_{i=1}^n \delta_i \otimes x_i$, where $\{\delta_i : 1 \leq i \leq n\}$ is the dual basis to $\{x_1, \ldots, x_n\}$, as can be checked by computing on the basis $\{x_1, \ldots, x_n\}$. Then it follows immediately from the definitions that

$$C_V = \sum_{i=1}^{n} \rho(y_i)\rho(x_i),$$

Lemma 15.9. *Let (V, ρ) be a representation of a Lie algebra $\mathfrak{g}$ such that t_V is non-degenerate on $\mathfrak{g}$. Then the Casimir operator C_V satisfies $tr(C_V) = \dim(\mathfrak{g})$. In particular it is non-zero.*

Proof. This is immediate from the above formula for C_V.

$$\text{tr}(C_V) = \text{tr}(\sum_{i=1}^{n} \rho(y_i), \rho(x_i)) = \sum_{i=1}^{n} \text{tr}(\rho(y_i)\rho(x_i)) = \sum_{i=1}^{n} t_V(y_i, x_i) = \sum_{i=1}^{n} 1 = \dim(\mathfrak{g}).$$

□

Now suppose that $\mathfrak{g}$ is semisimple, and (V, ρ) an arbitrary representation of $\mathfrak{g}$. While it is not the case that the trace form t_V must be non-degenerate, the following Lemma identifies its radical:

Lemma 15.10. *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 $\rho(\mathfrak{g}) \subseteq \mathfrak{gl}(V)$ 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 $\rho(\mathfrak{g})$. But the radical $\mathfrak{r} = \text{rad}(t_V)$ 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 $\rho(\mathfrak{g})$ is semisimple, this forces the radical to be zero as required. □

Definition 15.11. If $\mathfrak{g}$ is a semisimple Lie algebra and (V, ρ) is a $\mathfrak{g}$-representation on which $\mathfrak{g}$ acts non-trivially (so that $\rho(\mathfrak{g}) \neq \{0\}$) then, by Lemma 15.10, the above construction can be applied to V as a representation[21] of the semisimple Lie algebra $\rho(\mathfrak{g})$, yielding a $\mathfrak{g}$-endomorphism[22] of V, with trace $\text{tr}(C_V) = \dim(\rho(\mathfrak{g}))$. Thus for any representation (V, ρ) of a semisimple Lie algebra $\mathfrak{g}$ with $\rho(\mathfrak{g}) \neq \{0\}$ we obtain a non-zero $\mathfrak{g}$-endomorphism C_V of V which we will call the Casimir of V.

Remark 15.12. If $\mathfrak{g}$ is simple, rather than just semisimple, then by Schur's Lemma $\text{Hom}_{\mathsf{k}}(\mathfrak{g}, \mathfrak{g})^{\mathfrak{g}} = \text{Hom}_{\mathfrak{g}}(\mathfrak{g}, \mathfrak{g})$ is one-dimensional (the scalar multiples of the identity). Since $\text{Hom}_{\mathsf{k}}(\mathfrak{g}, \mathfrak{g}) \cong \mathfrak{g} \otimes \mathfrak{g}$ 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 15.13. Let us take $\mathfrak{g} = \mathfrak{sl}_2$. Then the trace form $t(x, y) = \text{tr}(x.y)$ 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

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

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.

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 will need $\mathfrak{g}.V = \text{span}\{\rho(x)(v) : x \in \mathfrak{g}, v \in V\}$. It is clearly a subrepresentation of V.

Lemma 15.14. *Let (V, ρ) be representation of a semisimple Lie algebra $\mathfrak{g}$. Then $V = V^{\mathfrak{g}} \oplus \mathfrak{g}.V$. Moreover, if $q \colon V \to W$ is a surjective homomorphism, then $q(V^{\mathfrak{g}}) = W^{\mathfrak{g}}$.*

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 $\dim(V).\lambda = \text{tr}(C_V) = \dim(\rho(\mathfrak{g}))$, so that $\lambda \neq 0$[23], and hence C_V is invertible. Since it is clear from the definition

[21] with action map the inclusion map from $\rho(\mathfrak{g})$ into $\mathfrak{gl}(V)$.

[22] Since by construction C_V commutes with every element of $\rho(\mathfrak{g})$.

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

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.

For the last part, note that if $q\colon V \to W$ is a homomorphism of representations, it is clear that $q(V^{\mathfrak{g}}) \subseteq W^{\mathfrak{g}}$ and $q(\mathfrak{g}.V) \subseteq \mathfrak{g}.W$. It follows that if q is surjective we must have both containments being equalities. □

Corollary 15.15. *Let $\mathfrak{g}$ be a semisimple Lie algebra, and let $q\colon V \to W$ be a surjective $\mathfrak{g}$-homomorphism. Then q has a right inverse.*

Proof. Consider the natural map of vector spaces

$$q_*\colon \mathrm{Hom}(W,V) \to \mathrm{Hom}(W,W),$$

given by $s \mapsto q \circ s$. It is easy to see that q_* is a surjective map of representations of $\mathfrak{g}$. Now clearly $\mathrm{Hom}(W,W)$ contains a copy of the trivial representation of $\mathfrak{g}$, given by the scalar multiples of the identity. It follows from Lemma 15.14 that $\mathrm{Hom}(W,V)$ must have a copy of the trivial representation mapping to this subrepresentation. But this means there is a $\mathfrak{g}$-invariant linear map $s\colon W \to V$ such that $q_*(s) = q \circ s = \mathrm{id}_W$, i.e. the map q has a splitting. The fact that V is semisimple now follows from Lemma 15.3. □

Theorem 15.16. *(Weyl's theorem) Let V be a finite dimensional representation of a semisimple Lie algebra over an algebraically closed field k of characteristic zero. Then V is semisimple, and hence completely reducible.*

Proof. Lemma 15.3 shows that it is enough to show that if $q\colon V \to W$ is a surjective map of $\mathfrak{g}$-representations, then q has a right inverse $s\colon W \to V$. Corollary 15.15 produces the required right inverse. □

16. 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.

If V is a vector space and $x \in \mathrm{End}(V)$, then we have the natural direct sum decomposition of V into the generalized eigenspaces of x. This can be viewed as giving a decomposition of the endomorphism x in a semisimple (or diagonalisable) and nilpotent part, as the next Lemmas show.

Lemma 16.1. *If $x, y \in End(V)$ are commuting linear maps then if both are nilpotent, so is $x+y$, and similarly if both are semisimple, so is $x+y$.*

Proof. For semisimple linear maps this follows from the fact that if s is a semisimple linear map, its restriction to any invariant subspace is again semisimple. For nilpotent linear maps it follows because

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{i} x^k y^{n-k},$$

so that if n is large enough, *e.g.* $n \geq 2\dim(V)$, each of these terms will be zero (since x and y are nilpotent). □

Proposition 16.2. *Let V be a finite dimensional vector space $x \in End(V)$. Then we may write $x = x_s + x_n$ where x_s 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, this decomposition is unique, and if U is a subspace of V preserved by x, it is also preserved by x_s, x_n*

Proof. Let $V = \bigoplus_{\lambda \in \mathsf{k}} V_\lambda$ be the generalised eigenspace decomposition of V, and let $p_\lambda\colon V \to V_\lambda$ be the projection with with kernel $\bigoplus_{\mu \neq \lambda} V_\mu$. If we set x_s to be $\sum_\lambda \lambda.p_\lambda$, clearly x_s and x commute, and their difference $x_n = x - x_s$ is nilpotent. This establishes the existence of the Jordan decomposition.

To see the uniqueness, suppose that $x = s + n$ is another such decomposition. Now since s commutes with x, it must preserve the generalised eigenspaces of x, and so, since x_s is just a scalar on each V_λ, clearly s commutes with x_s. It follows s and n both commute with x_s and x_n. But then by Lemma 16.1 $x_s - s$ and $n - x_n$ are semisimple and nilpotent respectively. Since $s + n = x_s + x_n$ they are equal, and the only endomorphism which is both semisimple and nilpotent is zero, thus $s = x_s$ and $n = x_n$ as required.

Finally, to see that x_s and x_n preserve any subspace U which is preserved by x, note that if $U = \bigoplus_{\lambda \in \mathsf{k}} U_\lambda$ is the decomposition of U into generalised eigenspaces of x, then clearly $U_\lambda \subseteq V_\lambda$, $(\forall \lambda \in \mathsf{k})$ and since x_s is a scalar on V_λ it certainly preserves U_λ, and hence all of U. As $x_n = x - x_s$ clearly x_n also preserves U.

□

Lemma 16.3. *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. In particular, if $x = x_s + x_n$ is the Jordan decomposition of x, then $ad(x) = ad(x_s) + ad(x_n)$ is the Jordan decomposition of $ad(x)$. In other words, $ad(x)_s = ad(x_s)$ and $ad(x)_n = ad(x_n)$.

Proof. Suppose first that $x \in \mathrm{End}(V)$ be semisimple. Then there is a basis $\{e_i : 1 \leq i \leq n\}$ of V and scalars $\{\lambda_i \in \mathsf{k} : 1 \leq i \leq n\}$ such that $x(e_i) = \lambda_i e_i$, that is, e_i is an eigenvector of x with eigenvalue λ_i. (Note the λ_i are thus not necessarily distinct). Let $\{\delta_1 \ldots, \delta_n\}$ be the basis of V^* dual to $\{e_1, \ldots, e_n\}$. Then for any $i, j \in \{1, \ldots, n\}$, $x(\delta_j)(e_i) = -\delta_j(x(e_i)) = -\lambda_i \delta_j(e_i)$, so that $x(\delta_j) = -\lambda_j \delta_j$.

Now if $\alpha \in \mathrm{End}(V)$ is arbitrary, the matrix $A = (a_{ij}) \in \mathrm{Mat}_n(\mathsf{k})$ associated to it is given by $\alpha(e_j) = \sum_{i=1}^n a_{ij} e_i$, $(\forall j, 1 \leq j \leq n)$. It follows that $\alpha = \sum_{i,j} a_{ij}(\delta_j.e_i)$, that is, under the identification $\alpha \mapsto A = (a_{ij})$, the elementary matrices E_{ij} correspond to $\delta_j.e_i$. But then by the previous paragraph,

$$\mathrm{ad}(x)(\delta_j.e_i) = x \circ (\delta_j.e_i) - (\delta_j.e_i) \circ x = \delta_j x(e_i) + x(\delta_j).e_i = (\lambda_i - \lambda_j)(\delta_j.e_i)$$

so that $\{\delta_j.e_i, 1 \leq i, j \leq n\}$ is a basis of eigenvectors for $\mathrm{ad}(x)$ and hence $\mathrm{ad}(x)$ is semisimple.

If x is nilpotent, then $\mathrm{ad}(x) = \lambda_x - \rho_x$ where λ_x and ρ_x denote left and right multiplication by x. Since λ_x and $-\rho_x$ clearly commute and are both nilpotent if x is, it follows from Lemma 16.1 that $\mathrm{ad}(x)$ is nilpotent.

Since $0 = \mathrm{ad}([x_s, x_n]) = [\mathrm{ad}(x_s), \mathrm{ad}(x_n)]$, it follows that the decomposition $\mathrm{ad}(x) = \mathrm{ad}(x_s) + \mathrm{ad}(x_n)$ satisfies the properties characterizing $\mathrm{ad}(x)_s$, $\mathrm{ad}(x)_n$ and so the final sentence follows from the uniqueness of the decomposition in Proposition 16.2. □

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 16.4. 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 that if $\mathfrak{g} = \mathfrak{gl}(V)$ for some vector space V, then Lemma 16.3 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 16.5. *Let $\mathfrak{a}$ be a semisimple Lie algebra.*

1. *Suppose that $\mathfrak{a}$ is an ideal of a Lie algebra $\mathfrak{g}$. Then there is a unique ideal I in $\mathfrak{g}$ such that $\mathfrak{g}$ is the direct sum of ideals $\mathfrak{a} \oplus I$.*
2. *All derivations of $\mathfrak{a}$ are inner, that is, $Der_{\mathsf{k}}(\mathfrak{a}) = Inn_{\mathsf{k}}(\mathfrak{a})$.*

Proof. For the first part, let κ be the Killing form of $\mathfrak{g}$ and let $\mathfrak{a}^\perp = \{x \in \mathfrak{g} : \kappa(x, a) = 0, \forall a \in \mathfrak{a}\}$, then $\mathfrak{a}^\perp$ is an ideal in $\mathfrak{g}$. Now $\mathfrak{a} \cap \mathfrak{a}^\perp$ is an ideal of $\mathfrak{g}$ on which the Killing form vanishes, so that by Cartan's Criterion[24] it is solvable. But since it is also an ideal of the semisimple Lie algebra $\mathfrak{a}$, it follows that $\mathfrak{a} \cap \mathfrak{a}^\perp = \{0\}$. Now since κ is nondegenerate on $\mathfrak{a}$, we also have $\dim(\mathfrak{a}) + \dim(\mathfrak{a}^\perp) = \dim(\mathfrak{g})$ so that $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$ as required[25]. For uniqueness, note that if $\mathfrak{g} = \mathfrak{a} \oplus I$, is a direct sum of ideals, then $[\mathfrak{a}, I] = 0$ and so clearly $I \subseteq \mathfrak{a}^\perp$.

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

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

[24] Again we use the fact that the Killing form on an ideal is the restriction of the Killing form for the whole Lie algebra.

[25] To see the claim about dimensions, note that κ gives a map $\theta\colon \mathfrak{g} \to \mathfrak{a}^*$, where $\theta(x)(a) = \kappa(x, a)$ for $x \in \mathfrak{g}, a \in \mathfrak{a}$. Clearly $\ker(\theta) = \mathfrak{a}^\perp$. Since κ is non-degenerate on $\mathfrak{a}$, the map θ is injective when restricted to $\mathfrak{a}$, and hence it is surjective. But then it's kernel has the required dimension.

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

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. If $x \in \mathfrak{a}_\lambda$ and $y \in \mathfrak{a}_\mu$, then an easy induction shows that

$$(\delta - (\lambda + \mu))^n([x, y]) = \sum_{r=0}^{n} \binom{n}{r} [(\delta - \lambda)^r(x), (\delta - \mu)^{n-r}(y)]$$

hence $[x, y] \in \mathfrak{a}_{\lambda+\mu}$. It follows immediately that s is a derivation on $\mathfrak{a}$, and since $n = \delta - s$ we see that n is also. □

Theorem 16.7. *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 ad(s) is semisimple, while 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 16.6 so do $\mathrm{ad}(x)_s$ and $\mathrm{ad}(x)_n$. But then by Lemma 16.5, 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 16.8. 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.

17. 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 17.1. *Suppose that $\mathfrak{g}$ is a semisimple Lie algebra and $\mathfrak{h}$ is a Cartan subalgebra.*

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 of κ to $\mathfrak{h}$ is nondegenerate.*

Proof. For the first part, from Lemma 9.1 we know that $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subseteq \mathfrak{g}_{\lambda+\mu}$. Thus 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 $\mathfrak{g}_\alpha^\perp = \mathfrak{g}$, which is impossible since κ is nondegenerate.

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 17.2. *Let $\mathfrak{g}$ be a semisimple Lie algebra and $\mathfrak{h}$ a Cartan subalgebra. Then $\mathfrak{h}$ is abelian. Moreover, all the elements of $\mathfrak{h}$ are semisimple.*

Proof. We show that $[\mathfrak{h},\mathfrak{h}] = D\mathfrak{h} = 0$. By part (3) of Proposition 17.1 it is enough to show that $D\mathfrak{h}$ lies in the radical of $\kappa_{|\mathfrak{h}}$. Suppose that $x, y \in \mathfrak{h}$. Then we claim that:

$$\kappa(x,y) = \text{tr}(\text{ad}(x)\text{ad}(y)) = \sum_{\lambda\in\Phi} \dim(\mathfrak{g}_\lambda)\lambda(x)\lambda(y). \tag{17.1}$$

Indeed each $\mathfrak{g}_\lambda$ (for $\lambda \in \Phi \cup \{0\}$) is an $\mathfrak{h}$-subrepresentation, so we may compute the trace of $\text{ad}(x)\text{ad}(y)$ on each in turn and then add the resulting expressions. Now $\mathfrak{h}$ is solvable, so we may find a basis for $\mathfrak{g}_\lambda$ with respect to which each $\text{ad}(x)$ has an upper triangular matrix, and since $\text{ad}(x)$ has the unique eigenvalue $\lambda(x)$ on $\mathfrak{g}_\lambda$, the diagonal entries of the matrix must all equal $\lambda(x)$. It follows that the matrix of $\text{ad}(x)\text{ad}(y)$ is upper triangular with diagonal entries $\lambda(x)\lambda(y)$, and so summing over $\lambda \in \Phi \cup \{0\}$ we obtain Equation (17.1). Now λ is a one-dimensional representation of $\mathfrak{h}$, so it vanishes on $D\mathfrak{h}$. It is then immediate from Equation (17.1) 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 $x \in \text{rad}(\kappa_{|\mathfrak{h}}) = \{0\}$ as claimed.

Next suppose that $x \in \mathfrak{h}$ has Jordan decomposition $x = s + n$ in $\mathfrak{g}$. Now $\text{ad}(s), \text{ad}(n)$ must preserve $\mathfrak{h}$ since $\text{ad}(x)$ does, and thus since $\mathfrak{h}$ is self-normalising, we see that $s, n \in \mathfrak{h}$. Thus it is enough to show that $\mathfrak{h}$ contains no nilpotent elements. But if $n \in \mathfrak{h}$ is nilpotent, then $\text{ad}(n)$ is nilpotent on $\mathfrak{g}$, and hence on each $\mathfrak{g}_\lambda$. But then $0 = \text{tr}(\text{ad}(n)_{|\mathfrak{g}_\lambda}) = \dim(\mathfrak{g}_\lambda)\lambda(n)$, so that $\lambda(n) = 0$. But then we see from Equation (17.1) that n is in the radical of $\kappa_{|\mathfrak{h}}$, so $n = 0$ as required. □

Remark 17.3. Since the restriction of κ to $\mathfrak{h}$ is non-degenerate, it yields an isomorphism from $\mathfrak{h}^*$ 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.)

Remark 17.4. Some textbooks study semisimple Lie algebras $\mathfrak{g}$ via *maximal toral subalgebras*. These are subalgebras of $\mathfrak{g}$ consisting entirely of semisimple elements, maximal with this property. It follows readily from the above Lemma and the Cartan decomposition that Cartan subalgebras of semisimple Lie algebras are maximal toral subalgebras, though we will not use this fact.

Proposition 17.5. *Let $\mathfrak{g}$ be a semisimple Lie algebra and let $\mathfrak{h}$ be a Cartan subalgebra, and $\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha\in\Phi} \mathfrak{g}_\alpha$ the associated Cartan decomposition.*

1. *If $x \in \mathfrak{g}_\alpha$, $y \in \mathfrak{g}$ and $h \in \mathfrak{h}$ then*
$$\kappa(h, [x,y]) = \alpha(h)\kappa(x,y).$$
2. *The roots $\alpha \in \Phi$ span $\mathfrak{h}^*$.*
3. *The subspace $\mathfrak{h}_\alpha = [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subset \mathfrak{h}$ is one-dimensional and $\alpha(\mathfrak{h}_\alpha) \neq 0$.*
4. *If $\alpha \in \Phi$, we may find $e_\alpha \in \mathfrak{g}_\alpha$, $f_\alpha \in \mathfrak{g}_{-\alpha}$ and $h_\alpha \in \mathfrak{h}_\alpha$ 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 (1) we have

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

as required.

For (2), suppose that $W = \text{span}\{\Phi\}$. If W is a proper subspace of $\mathfrak{h}^*$, then we may find an $h \in \mathfrak{h}$ such that $\alpha(h) = 0$ for all $\alpha \in \Phi$. But then it follows from (17.1) that $\kappa(h,x) = 0$ for all $x \in \mathfrak{h}$, which contradicts the nondegeneracy of the form $\kappa_{|\mathfrak{h}}$.

For (3), as in the remark above, since $\kappa_{|\mathfrak{h}}$ is nondegenerate it yields an isomorphism $\mathfrak{h}^* \to \mathfrak{h}$, given by $\lambda \mapsto t_\lambda$ where $\kappa(t_\lambda, h) = \lambda(h)$ 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 $x \in \mathfrak{g}_\alpha$, $y \in \mathfrak{g}_{-\alpha}$. Then by (1) we see that $[x,y] = \kappa(x,y)t_\alpha$, so that $\mathfrak{h}_\alpha \subseteq \text{span}\{t_\alpha\}$. Since κ is nondegenerate on $\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}$ we may find $x \in \mathfrak{g}_\alpha, y \in \mathfrak{g}_{-\alpha}$ such that $\kappa(x,y) \neq 0$, hence $\mathfrak{h}_\alpha = \text{span}\{t_\alpha\}$ as required.

Next we wish to show that $\alpha(\mathfrak{h}_\alpha) \neq 0$. For this note that if $\alpha(\mathfrak{h}_\alpha) = 0$ then pick $x \in \mathfrak{g}_\alpha$, $y \in \mathfrak{g}_{-\alpha}$ so that $z = [x,y] \in \mathfrak{h}_\alpha$ is nonzero. Then $[z,x] = \alpha(z)x = 0 = -\alpha(z)y = [z,y]$, so that $\mathfrak{a} = \mathsf{k}\text{-span}\{x,y,z\}$ is a solvable subalgebra of $\mathfrak{g}$. In particular, by Lie's theorem we may find a basis of $\mathfrak{g}$ with respect to which the matrices of $\text{ad}(\mathfrak{a})$ act by upper triangular matrices, and so $\text{ad}(z) = \text{ad}([x,y])$ acts by a strictly upper triangular matrix, and hence is nilpotent. Since we also know $z \in \mathfrak{h}$ we have $\text{ad}(z)$ is semisimple, hence $\text{ad}(z)$ is both semisimple and nilpotent, which implies it is zero, contradicting $z \neq 0$.

Given $\alpha(\mathfrak{h}_\alpha) \neq 0$, it is clear that there is a unique $h_\alpha \in \mathfrak{h}_\alpha$ such that $\alpha(h_\alpha) = 2$. Next if $e_\alpha \in \mathfrak{g}_\alpha$ is nonzero, then using the nondegeneracy of κ and part (1) we can pick $f_\alpha \in \mathfrak{g}_{-\alpha}$ so that $[e_\alpha, f_\alpha] = h_\alpha$ It is easy to check that $\{e_\alpha, f_\alpha, h_\alpha\}$ span an copy of $\mathfrak{sl}_2$ in $\mathfrak{g}$ which establishes (4).

□

Remark 17.6. 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*. Given $\alpha \in \Phi$, the copy of $\mathfrak{sl}_2$ attached to α in the previous proposition will be denoted $\mathfrak{sl}_\alpha$.

Lemma 17.7. *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*

- *The root spaces $\mathfrak{g}_\alpha$ are one-dimensional.*
- *If $\alpha \in \Phi$ and $c\alpha \in \Phi$ for some $c \in \mathbb{Z}$ then $c = \pm 1$.*

Proof. Choose a nonzero vector $e_\alpha \in \mathfrak{g}_\alpha$. Then as above we may find an element $e_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $[e_\alpha, e_{-\alpha}] = h_\alpha \in \mathfrak{h}$ (since κ restricted to $\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}$ is nondegenerate). We can then choose $e_{-\alpha}$ so 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 $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = \mathsf{k}.h_\alpha$, and $[e_\alpha, h_\alpha] = 2e_\alpha$, 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 semisimply on each of the direct sums defining M, so that

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

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 $\alpha \in \Phi$, this completes the proof. □

17.1. **Some refinements.** We can refine somewhat the structure of the Cartan decomposition we have already obtained, using the same techniques. Suppose that α, β are two roots in $\mathfrak{g}$ such that $\beta \neq k\alpha$ for $k \in \mathbb{Z}$. 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 $\beta + k\alpha \in \Phi$ for each k with $-p \leq k \leq q$, but neither $\beta - (p+1)\alpha$ nor $\beta + (q+1)\alpha$ are in Φ. This set of roots is called the α-*string* through β.

Proposition 17.8. *Let $\beta - p\alpha, \dots, \beta + q\alpha$ be the α-string through β. Then we have*

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

In particular $\beta - \beta(h_\alpha).\alpha \in \Phi$. Moreover, if $\alpha \in \Phi$ and $c \in \mathsf{k}$ has $c\alpha \in \Phi$ then $c \in \{\pm 1\}$.

Proof. We consider the subspace $M = \bigoplus_{-p\leq k\leq q} \mathfrak{g}_{\beta+k\alpha}$. Pick $e_\alpha \in \mathfrak{g}_\alpha$ and $e_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $0 \neq [e_\alpha, e_{-\alpha}] = h_\alpha$ and so that $\{e_\alpha, e_{-\alpha}, h_\alpha\}$ form the standard generators of $\mathfrak{sl}_2$ as above. It is clear that $e_\alpha, h_\alpha, e_{-\alpha}$ preserve M, so we $\mathrm{tr}_{|M}(h_\alpha) = 0$, and so, using the fact root spaces are 1-dimensional, we have the identity:

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

and so

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

and so since $p + q + 1 \neq 0$ and $\alpha(h_\alpha) = 2$, we obtain:

$$\beta(h_\alpha) = p - q.$$

as required. Since $\beta - (p-q)\alpha$ is certainly in the α-string through β it follows that $\beta - \beta(h_\alpha).\alpha \in \Phi$.

For the second part, taking $\beta = c\alpha$ we see that $2c = p - q \in \mathbb{Z}$, and so $c = \frac{1}{2}(p-q)$. But now if $p - q$ is even we can immediately conclude as before that $c = \pm 1$. 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 $\alpha = 2(\frac{1}{2}\alpha)$. □

18. $\mathfrak{h}$ AND INNER PRODUCT SPACES

In this section we work over an algebraically closed field k of characteristic zero. Results of this section were only sketched in the lectures and the proofs are non-examinable.

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 18.1. *Let $\alpha \in \Phi$. Then if $\{e_\alpha, f_\alpha, h_\alpha\}$ are a standard basis for $\mathfrak{sl}_\alpha$ then we have*

(1)

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

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

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

Proof. The first equality follows immediately from the results of the previous lecture. For the second, note that

$$2 = \alpha(h_\alpha) = \kappa(t_\alpha, h_\alpha) = \kappa(t_\alpha, \kappa(e_\alpha, f_\alpha)t_\alpha) = \kappa(t_\alpha, t_\alpha)\kappa(e_\alpha, f_\alpha),$$

and substitute into the first expression. The last expression follows from calculating $\kappa(h_\alpha, h_\alpha)$ using the second expression.

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 17.8:

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

Finally, the first part of the Lemma now immediately gives $\kappa(t_\alpha, t_\beta) \in \mathbb{Q}$. □

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, $(\alpha, \beta) = \kappa(t_\alpha, t_\beta) \in \mathbb{Q}$ for all $\alpha, \beta \in \Phi$.

Lemma 18.2. *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 $\{\alpha_1, \alpha_2, \ldots, \alpha_l\}$ 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 $\{\alpha_i : 1 \leq i \leq l\}$. 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 18.3. *The form $(-,-)$ is positive definite on $\mathfrak{h}^*_{\mathbb{Q}}$.*

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

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

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. □

19. Bases for 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[26], 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 $\mathbb{Q}$. 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 19.1. 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".

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

Definition 19.2. 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.

1. Φ spans V;
2. If $\alpha \in \Phi$ then $c\alpha \in \Phi$ if and only if $c = \pm 1$;
3. If $\alpha \in \Phi$ then $s_\alpha \colon V \to V$ preserves Φ;
4. If $\alpha, \beta \in \Phi$ then $2(\beta,\alpha)/(\alpha,\alpha) \in \mathbb{Z}$.

Remark 19.3. If $(-,-)$ denotes the inner product on $\mathfrak{h}_{\mathbb{Q}}^*$, then results above show that the roots in $\mathfrak{h}^*$ form a root system in the above sense: one simply has to translate the information about the elements $t_\alpha \in \mathfrak{h}$ obtained in our analysis of the Cartan decomposition. The crucial fact is that, for the roots arising from a semisimple Lie algebra we have

$$\frac{2(\beta,\alpha)}{(\alpha,\alpha)} = \kappa\left(t_\beta, \frac{2t_\alpha}{\kappa(t_\alpha,t_\alpha)}\right) = \beta(h_\alpha) \in \mathbb{Z},$$

which is proved in Proposition 17.8. That Proposition also showed that $s_\alpha(\beta) = \beta - \beta(h_\alpha).\alpha \in \Phi$, and that the only possible rational multiple of a root which is also a root is its negative; hence properties $2) - 4)$ hold. $1)$ follows from Lemma 18.2. So we see that the roots Φ in $\mathfrak{h}_{\mathbb{Q}}^*$ do indeed form a root system.

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 19.4. 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 19.5. 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

$$\mathfrak{h}_{\mathbb{Q}}^* = \left\{\sum_{i=1}^n c_i\varepsilon_i : c_i \in \mathbb{Q}\right\} \Big/ \{\mathbb{Q}.(\varepsilon_1 + \ldots + \varepsilon_n)\},$$

[26]That is, one with a notion of distance and angle. Apart from working over $\mathbb{Q}$ rather than $\mathbb{R}$, this is pretty much the vector geometry of Geometry I.

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.

A key step in the classification process is to find a good basis for V: we assume that Φ spans V, so certainly we may find a subset of Φ which is a basis, but it turns out we can find a class of particularly well adapted bases.

Definition 19.6. Let (V, Φ) be a root system, and let Δ be a subset of Φ. We say that Δ is a *base* for Φ if

1. Δ is a basis for V.
2. Every $\beta \in \Phi$ can be written as $\sum_{\alpha \in \Delta} c_\alpha \alpha$ where $c_\alpha \in \mathbb{Z}$ and the non-zero c_αs all have the same sign.

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 first crucial point is that the angles between roots are very constrained. For convenience we will write

$$\langle \alpha, \beta \rangle = \frac{2(\alpha, \beta)}{(\alpha, \alpha)}, \tag{19.1}$$

which we call a *Cartan integer*.

Lemma 19.7. *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:

$$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4 \cos(\theta)^2 < 4. \tag{19.2}$$

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	$\|\|\alpha\|\|^2/\|\|\beta\|\|^2$
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

□

19.1. Existence of bases.

In fact, root systems have bases, and there is in fact a remarkable relationship between bases and elements of the Weyl group.

Theorem 19.8. *Given a root system (V, Φ), it has at least one base Δ. Suppose that Δ and Δ_1 are bases of (V, Φ). Then there is a $w \in W$ such that $w(\Delta) = \Delta_1$.*

If $v \in V$, we say that v is regular if $(v, \alpha) \neq 0$ for all $\alpha \in \Phi$. Since V is not a finite union of proper subspaces, regular vectors certainly exist. Given such a v, write $\Phi^+(v) = \{\alpha \in \Phi : (v, \alpha) > 0\}$, and $\Phi^-(v) = -\Phi^+(v)$. Clearly $\Phi = \Phi^+(v) \bigsqcup \Phi^-(v)$. We say a root $\alpha \in \Phi^+(v)$ is *decomposable* if we can find roots $\alpha_i \in \Phi^+(v)$ and nonnegative integers n_i, $(1 \leq i \leq s, s \geq 2)$, such that $\alpha = \sum_{i=1}^s n_i \alpha_i$. Let $\Delta(v)$ be the set of indecomposable roots, *i.e.* the roots which are not decomposable, in $\Phi^+(v)$.

Proposition 19.9. *Let (V, Φ) be a root system, and let $v \in V$ be a regular vector. Then $\Delta(v)$ is a base for (V, Φ) and every base is of this form.*

Proof. Let $\Delta(v)$ be the indecomposable roots in $\Phi^+(v)$ as above. We first show that every root in $\Phi^+(v)$ is a nonnegative integer combination of indecomposable roots. Indeed let $m = \min\{(v,\alpha) : \alpha \in \Phi^+\}$. If $\gamma \in \Phi^+(v)$, and we have $\gamma = \sum_{i=1}^s n_i\alpha_i$ where $n_i \in \mathbb{Z}_{>0}$ and $\alpha_i \in \Phi^+(v)$, then clearly $(v,\gamma) \geq \sum_{i=1}^s n_i.m$, hence for any such expression the sum $\sum_{i=1}^s n_i$ is bounded above by $(v,\alpha)/m$. If we write $\gamma = \sum_{i=1}^s n_i\alpha_i$ where $\sum_{i=1}^s n_i$ is maximal possible, then we claim each γ_i is indecomposable. Indeed otherwise there is some i such that γ_i is decomposable, so that we may write $\gamma_i = \sum_{j=1}^t m_j\beta_j$ for some $t \geq 2$, positive integers m_j, and $\beta_j \in \Phi^+(v)$, $(1 \leq j \leq t)$. But then $\gamma = \sum_{k\neq i} n_i\alpha_i + \sum_{j=1}^t m_j n_i \beta_j$ and $\sum_{k\neq i} n_k + (\sum_{j=1}^t m_j)n_i \geq \sum_{k\neq i} n_k + 2n_i > \sum_{i=1}^s n_i$ contradicting our maximality assumption. Since Φ also spans V, this shows that $\Delta(f)$ spans V, hence since $\Phi = \Phi^+(v) \sqcup -\Phi^+(v)$ in order to show that $\Delta(v)$ is a base it only remains to show that $\Delta(f)$ is linearly independent.

We check this in a number of steps.

Step 1: The angle between any two vectors $\alpha, \beta \in \Delta(v)$ is obtuse, that is $(\alpha,\beta) \leq 0$: To see this suppose that $(\alpha,\beta) > 0$. Then one can check[27] using the table from the proof of the Lemma from last time on the angles between roots that if we take $(\alpha,\alpha) \geq (\beta,\beta)$ then $\langle\beta,\alpha\rangle = -1$. But then $s_\alpha(\beta) = \beta - \alpha \in \Phi$, so that either $\beta - \alpha$ or $\alpha - \beta$ lie in Φ^+, and hence either α or β will be decomposable.

Step 2: Linear independence: Suppose that we have $\sum_{\alpha\in\Delta(f)} c_\alpha\alpha = 0$ for $c_\alpha \in \mathbb{Q}$, $(\alpha \in \Delta)$. Gathering all the positive and nonpositive coefficients let

$$z = \sum_{c_\alpha > 0} c_\alpha\alpha = \sum_{c_\beta \leq 0} (-c_\beta)\beta$$

Then we have:

$$(z,z) = \sum_{\alpha,\beta: c_\alpha > 0, c_\beta \leq 0} c_\alpha.(-c_\beta)(\alpha,\beta) \leq 0,$$

so by positive definiteness we must have $z = 0$. But since $(v,z) = \sum_{c_\alpha > 0} c_\alpha(v,\alpha) \geq 0$ with equality if and only if the sum is empty we conclude that for all $\beta \in \Delta$ we have $c_\beta \leq 0$, but then we have $0 = \sum_{c_\beta \leq 0}(-c_\beta)\beta$, and pairing with v again we deduce that for all β we have $c_\beta = 0$, and so Δ is a linearly independent set.

Step 3: Finally we need to show that any base Δ is of the form $\Delta(v)$ for some regular $v \in V$. For $\alpha \in \Delta$, let $P_\alpha = \{v \in V : (v,\alpha) > 0\}$. This is a half-space in V. It can be checked that $\bigcap_{\alpha\in\Delta} P_\alpha$ is non-empty (indeed this holds for any linearly independent set in V, not just a base[28]) so we may pick some $v \in V$ in this intersection. Then since any root is a nonpositive or nonnegative combination of the roots in Δ, it follows the vector v is regular, and moreover it is clear that $\Phi^+ \subseteq \Phi^+(v)$ the set of positive roots associated with Δ, and since $\Phi^- = -\Phi^+$, we see similarly that $\Phi^- \subseteq \Phi^-(v)$, and since Φ is the disjoint union of both $\Phi^+ \sqcup \Phi^-$ and $\Phi^+(v) \sqcup \Phi^-(v)$ it follows that $\Phi^\pm(v) = \Phi^\pm$. Finally, since $\Delta(v)$ are the indecomposable roots in $\Phi^+ = \Phi^+(v)$, we must certainly have $\Delta(v) \subseteq \Delta$, and since both are bases of V it follows that $\Delta = \Delta(v)$ as required. ☐

Our proof shows that for a given root system one can always find a base, but it will not be unique: different choices of regular v can lead to different positive systems $\Phi^+(v)$ and hence different bases. This ambiguity is however controlled by the action of the Weyl group, as we will now see. From now on fix a base Δ for our root system, and thus have the corresponding positive and negative roots Φ^+, Φ^-. Let W_0 be the subgroup of W generated by the reflections s_α for $\alpha \in \Delta$. The elements of Δ are called *simple roots*, and the corresponding reflections s_α are called *simple reflections*.

Lemma 19.10. *If $\alpha \in \Delta$, then the reflection s_α preserves the set $\Phi^+\backslash\{\alpha\}$. Moreover, if $\rho = \frac{1}{2}\sum_{\alpha\in\Phi^+}\alpha$, then $(\rho,\gamma) = (1/2)||\gamma||^2 > 0$ for all $\gamma \in \Delta$ and hence $\Phi^+ = \Phi^+(\rho)$*

Proof. Let $\gamma \in \Phi^+\backslash\{\alpha\}$. Since Δ is a base, we may write $\gamma = \sum_{\beta\in\Delta} c_\beta\beta$ where $c_\beta \in \mathbb{Z}_{\geq 0}$ for all $\beta \in \Delta$, and so

$$s_\alpha(\gamma) = (c_\alpha - \langle\alpha,\gamma\rangle)\alpha + \sum_{\beta\in\Delta,\beta\neq\alpha} c_\beta\beta \in \Phi.$$

Now since $\gamma \neq \pm\alpha$ (since it is positive) there must be some $\beta_0 \in \Delta\backslash\{\alpha\}$ with $c_{\beta_0} > 0$. But then by the definition of a base we must have $s_\alpha(\gamma) \in \Phi^+$, and clearly $s_\alpha(\gamma) \neq \alpha$ since this would imply

[27]This either follows just inspecting the table, or by noting that by inequality proved there for α, β with $\beta \neq \pm\alpha$ we have $\langle\alpha,\beta\rangle.\langle\beta,\alpha\rangle < 4$ and so since $\langle\alpha,\beta\rangle$ and $\langle\beta,\alpha\rangle$ are integers at least one must be ± 1.

[28]Just extend your linearly independent set to a basis, and take the dual basis with respect to the inner product– the intersection of half-spaces then become the set of vectors whose coordinates with respect to the appropriate dual basis vectors are positive.

$\gamma = s_\alpha(\alpha) = -\alpha \in \Phi^-$. For the final part, note that if $\gamma \in \Delta$ we must have

$$\begin{aligned} s_\gamma(\rho) &= s_\gamma(\frac{1}{2}\gamma) + s_\gamma(\frac{1}{2}\sum_{\alpha \in \Phi^+ \setminus \{\gamma\}} \alpha) \\ &= -\frac{1}{2}\gamma + \frac{1}{2}\sum_{\alpha \in \Phi^+ \setminus \{\gamma\}} \alpha \\ &= \rho - \gamma \end{aligned}$$

now since by definition $s_\gamma(\rho) = \rho - \frac{2(\rho,\gamma)}{(\gamma,\gamma)}\gamma$ the result follows immediately. The fact that $\Phi^+ = \Phi^+(\rho)$ (in the notation of Proposition 19.9 then follows immediately from the proof of that proposition. □

Definition 19.11. Given a root system equipped with a base $\Delta = \{\alpha_1, \dots, \alpha_r\}$ we may define a *height function* $\text{ht}\colon V \to \mathbb{Q}$ by setting $\text{ht}(v) = \sum_{i=1}^n c_i$ where $v = \sum_{i=1}^r c_i\alpha_i$. Note that it follows from the definition of a base that the height function is integer-valued on Φ, and $\text{ht}(\beta) = 1$ if and only if $\beta \in \Delta$.

Note that this height function is a special case of the height function used in the proof of Weyl's theorem – if $\{\varpi_i : 1 \le i \le r\}$ denotes the dual basis to the base Δ with respect to the inner product, then $\text{ht}(\beta) = (\delta, \beta)$ where $\delta = \sum_{i=1}^r \varpi_i$.

Proposition 19.12. *Suppose that $\beta \in \Phi$. Then there is a $w \in W_0$ and an $\alpha \in \Delta$ such that $w(\beta) = \alpha$.*

Proof. First suppose that $\beta \in \Phi^+$. We prove the statement by induction on the height of β. The statement being clear if β is of height 1, we may assume $h(\beta) > 0$. We claim there is some $\gamma \in \Delta$ such that $(\beta, \gamma) > 0$. If not then writing $\beta = \sum_{\gamma \in \Delta} c_\gamma \gamma$, we see that

$$(\beta, \beta) = \sum_{\gamma \in \Delta} c_\gamma(\beta, \gamma) \le 0,$$

so by positive definiteness we would have $\beta = 0$.

Now taking γ with $(\beta, \gamma) > 0$, we see that $s_\gamma(\beta) \in \Phi^+$ using the previous Lemma (since β is certainly not equal to α as $h(\beta) > 1$, and moreover $h(s_\gamma(\beta)) = h(\beta) - \langle \gamma, \beta \rangle < h(\beta)$. Thus by induction we are done. Finally we need to consider the case $\beta \in \Phi^-$. But then we have $-\beta \in \Phi^+$, and so there is a $w \in W_0$ such that $w(-\beta) = \gamma$ for some $\gamma \in \Delta$. But then clearly $(s_\gamma w)(\beta) = \gamma$ and we are done. □

Corollary 19.13. *The Weyl group W is generated by the reflections $\{s_\gamma : \gamma \in \Delta\}$, that is $W = W_0$.*

Proof. If $\beta \in \Phi$ then we have just shown in the previous proposition that there is a $w \in W_0$ such that $w(\beta) = \gamma$ for some $\gamma \in \Delta$. 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. □

Proof. (Proof of Theorem 19.8) By Proposition 19.9 we know that bases exist, and any base is determined by a regular element $v \in V$: given v, the base it yields is the set of indecomposable roots in $\Phi^+(v)$. Thus it is enough to show that there is a w in W such that $w(\Phi_1^+) = \Phi^+$, where Φ_1^+ and Φ^+ are the positive roots corresponding to Δ_1 and Δ respectively.

Suppose v is a regular vector such that $\Phi^+(v) = \Phi_1^+$. We claim that there is a $w \in W$ such that $w(v)$ satisfies $(w(v), \alpha) > 0$ for all $\alpha \in \Delta$. This immediately implies that $(w(v), \alpha) > 0$ for all $\alpha \in \Phi^+$, so that $\Phi^+ = \Phi^+(w(v))$, and thus $w(\Phi_1^+) = \Phi$, since $(w(v), \alpha) > 0$ if and only if $(v, w^{-1}(\alpha)) > 0$.

To prove the claim, first consider $\rho = \frac{1}{2}\sum_{\alpha \in \Phi} \alpha$. By the Lemma 19.10 we know that $s_\alpha(\rho) = \rho - \alpha$ for all $\alpha \in \Delta$. Now choose $w \in W$ such that $(\rho, w(v))$ is as large as possible. Then if $\alpha \in \Delta$, we must have

$$(w(v), \rho) \ge (s_\alpha.w(v), \rho) = (w(v), s_\alpha(\rho)) = (w(v), \rho - \alpha) = (w(v), \rho) - (w(v), \alpha).$$

Thus it follows that $(w(v), \alpha) \ge 0$ for all $\alpha \in \Delta$, and since $(v, \alpha) \ne 0$ for all $\alpha \in \Phi$ (as f was assumed to be generic) the claim follows.

□

Remark 19.14. In fact W acts simply transitively on the bases of (V, Φ), that is if $w(\Delta) = \Delta$ then $w = 1$. The proof (which we will not give) consists of examining the minimal length expression for w in terms of these generators $\{s_\alpha : \alpha \in \Delta\}$.

20. Cartan matrices and Dynkin diagrams

In this section we describe the data which is used in the classification of semisimple Lie algebras.

Definition 20.1. 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 $\{\alpha_1, \alpha_2, \dots, \alpha_\ell\} = \Delta$ 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 20.2. 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 $\{\alpha_1, \dots, \alpha_l\}$, 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:

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

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 20.3. *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 $\Delta = \{\alpha_1, \dots, \alpha_\ell\}$ of (V, Φ) and a base $\Delta' = \{\beta_1, \dots, \beta_\ell\}$ 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'_is 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 $\phi(\Phi) = \Phi'$.

Next, the linearity of ϕ and of $\langle -, - \rangle$ in the second variable immediately implies that $\langle \alpha_i, \gamma \rangle = \langle \phi(\alpha_i), \phi(\gamma) \rangle$ for any $\alpha_i \in \Delta$, $\gamma \in \Phi$. Finally, as in the previous paragraph, if $\alpha \in \Phi$ is arbitrary, then we may find $w \in W$ such that $\alpha = w(\alpha_j)$ for some $\alpha_j \in \Delta$, 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. □

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

Definition 20.4. 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 $V_1 = \mathrm{span}(\Phi_1)$ and $V_2 = \mathrm{span}(\Phi_2)$, 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 straight-forward to check that a root system is irreducible if and only if its associated Dynkin diagram is connected.

Definition 20.5. (*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 $\Delta = \{\alpha_1, \alpha_2, \dots, \alpha_l\}$ 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).

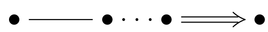
Such a set of vectors is called an *admissible set*.

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 $(\ell \geq 1)$:

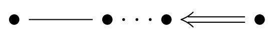
$$V = \{v = \sum_{i=1}^{\ell} c_i e_i \in \mathbb{Q}^\ell : \sum c_i = 0\}, \Phi = \{\varepsilon_i - e_j : 1 \leq i \neq j \leq \ell\}$$
$$\Delta = \{\varepsilon_{i+1} - \varepsilon_i : 1 \leq i \leq \ell - 1\}$$

- Type B_ℓ $(\ell \geq 2)$:



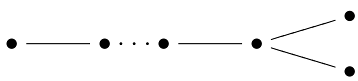
$$V = \mathbb{Q}^\ell, \Phi = \{\pm\varepsilon_i \pm \varepsilon_j : 1 \leq i, j \leq \ell, i \neq j\} \cup \{\varepsilon_i : 1 \leq i \leq \ell\},$$
$$\Delta = \{\varepsilon_1, \varepsilon_{i+1} - \varepsilon_i : 1 \leq i \leq \ell - 1\}$$

- Type C_ℓ $(\ell \geq 3)$:



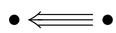
$$V = \mathbb{Q}^\ell, \Phi = \{\pm\varepsilon_i \pm \varepsilon_j : 1 \leq i, j \leq \ell, i \neq j\} \cup \{2\varepsilon_i : 1 \leq i \leq \ell\},$$
$$\Delta = \{2\varepsilon_1, \varepsilon_{i+1} - \varepsilon_i : 1 \leq i \leq \ell - 1\}$$

- Type D_ℓ $(\ell \geq 4)$:



$$V = \mathbb{Q}^\ell, \Phi = \{\pm\varepsilon_i \pm \varepsilon_j : 1 \leq i, j \leq \ell, i \neq j\},$$
$$\Delta = \{\varepsilon_1 + \varepsilon_2, \varepsilon_{i+1} - \varepsilon_i : 1 \leq i \leq \ell - 1\}$$

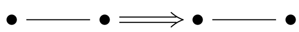
- Type G_2.



Let $e = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 \in \mathbb{Q}^3$, then:

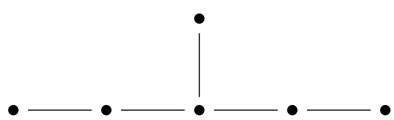
$$V = \{v \in \mathbb{Q}^3 : (v, e) = 0\}, \Phi = \{\varepsilon_i - \varepsilon_j : i \neq j\} \cup \{\pm(3\varepsilon_i - e) : 1 \leq i \leq 3\}$$
$$\Delta = \{\varepsilon_1 - \varepsilon_2, e - 3\varepsilon_1\}$$

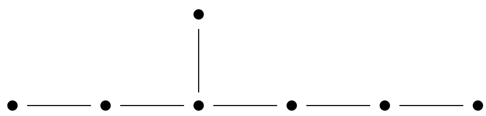
- Type F_4:

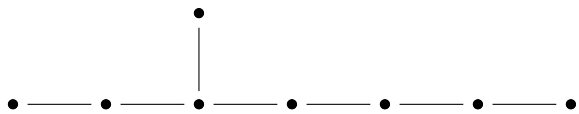


$$V = \mathbb{Q}^4,$$
$$\Phi = \{\pm\varepsilon_i : 1 \leq i \leq 4\} \cup \{\pm\varepsilon_i \pm e_j : i \neq j\} \cup \{\frac{1}{2}(\pm\varepsilon_1 \pm \varepsilon_2 \pm \varepsilon_3 \pm \varepsilon_4)\}$$
$$\Delta = \{\varepsilon_2 - \varepsilon_3, \varepsilon_3 - \varepsilon_4, \varepsilon_4, \frac{1}{2}(\varepsilon_1 - \varepsilon_2 - \varepsilon_3 - \varepsilon_4)\}.$$

- 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 20.6. 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!)

21. 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 21.1. *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 21.2. *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.[29]

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 $\{e_\alpha, f_\alpha, h_\alpha : \alpha \in \Phi\}$ 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.

[29]This is completely rigorous, but feels like cheating (to me).

22. APPENDIX 1: (MULTI)-LINEAR ALGEBRA

22.1. **Tensor Products.**

22.1.1. *Defintion.* Tensor products were studied in Part B Introduction to Representation Theory. We review their basic properties here.

Definition 22.1. Let V and W be vector spaces over a field k. The tensor product $V \otimes W$ is a k-vector space $V \otimes W$ equipped with a bilinear map $t\colon V \times W \to V \otimes W$ (where we write $v \otimes w$ for $t(v, w)$) which has the following universal property: If $B\colon V \times W \to U$ is a bilinear map taking values in a k-vector space U, then there exists a unique linear map $b\colon V \otimes W \to U$ such that $B = b \circ t$.

Remark 22.2. It is possible to construct $V \otimes W$ in various ways. If V and W are finite dimensional, we may pick a basis $\{e_1, \ldots, e_n\}$ of V and a basis $\{f_1, \ldots, f_m\}$ of W. Set $V \otimes W$ to be the vector space with basis $\{e_i \otimes f_j : 1 \leq i \leq n, 1 \leq j \leq m\}$, and if $v = \sum_{i=1}^{n} \lambda_i e_i$ and $w = \sum_{j=1}^{m} \mu_j f_j$, we set

$$t(v.w) = \sum_{i,j} (\lambda_i \mu_j) e_i \otimes f_j,$$

Given $B\colon V \times W \to U$ a bilinear map, we can define $b\colon V \otimes W \to U$ to be the unique linear map with $b(e_i \otimes f_j) = B(e_i, f_j)$. Provided you are happy with the axiom of choice, so that every vector space has a basis, the same construction gives the existence of the tensor product of arbitrary vector spaces.

Remark 22.3. 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, $V^{\otimes n} = V \otimes \ldots \otimes V$, 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 $\tau(v_1 \otimes \ldots \otimes v_n) := v_{\tau(1)} \otimes \ldots \otimes \tau_{\tau(n)}$.

Lemma 22.4. *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. (*c.f.* the proof that $\mathrm{ad}(x)$ is semisimple when x is). The map $(\alpha, w) \mapsto [v \mapsto \alpha(v).w]$ is bilinear[30], and so induces 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

$$\sum_{i \in I : (i,k) \in S} \lambda_{i,k} \delta_i \otimes f_k = \Big(\sum_{i \in I : (i,k) \in S} \lambda_{i,k} \delta_i \Big) \otimes f_k,$$

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 $v \in V$

$$0 = \theta(\gamma)(v) = \sum_{k \in S_K} \phi_k(v).f_k,$$

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 $\{\delta_1, \ldots, \delta_n\}$ of V^* to be dual to the basis $\{e_1, \ldots, e_n\}$ 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 22.5. 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 $\{e_1, \ldots, e_n\}$ is a basis of V and $\{f_1, \ldots, f_m\}$ is a basis of W, then taking the dual basis $\{\delta_1, \ldots, \delta_n\}$ of V^* it is easy to see that the images of $\delta_i \otimes f_j$

[30] 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.

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 22.4). 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.)

22.1.2. *Linear maps between tensor products.* If $\alpha\colon V_1 \to V_2$ and $\beta\colon W_1 \to W_2$, then 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

$$\mathrm{Hom}(V_1, W_1) \otimes \mathrm{Hom}(V_2, W_2) \to \mathrm{Hom}(V_1 \otimes V_2, W_1 \otimes W_2).$$

When all the vector spaces V_1, V_2, W_1, W_2 are finite dimensional, this map is an isomorphism, indeed using Lemma 22.4 you can check that

$$\begin{aligned} \mathrm{Hom}(V_1, W_1) \otimes \mathrm{Hom}(V_2, W_2) &\cong (V_1^* \otimes W_1) \otimes (V_2^* \otimes W_2) \\ &\cong (V_1^* \otimes V_2^*) \otimes (W_1 \otimes W_2) \\ &\cong \mathrm{Hom}(V_1 \otimes V_2, W_1 \otimes W_2), \end{aligned}$$

where the second isomorphism simply permutes the second and third tensor factors.

The contraction map ι defined in Lemma 22.4 allows us to decribe the composition map

$$\mathrm{Hom}(U, V) \times \mathrm{Hom}(V, W) \to \mathrm{Hom}(U, W).$$

Since composition is bilinear, it induces a linear map on the corresponding tensor product, thus it is given by a linear map

$$\mathrm{Hom}(U, V) \otimes \mathrm{Hom}(V, W) = (U^* \otimes V) \otimes (V^* \otimes W) \to \mathrm{Hom}(U, W) = U^* \otimes W.$$

But this map is simply $1_U \otimes \iota \otimes 1_W$.

22.1.3. *Tensor products and duality.* Suppose that V and W are 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 $(V \otimes W)^*$. Now by definition, the vectors in $(V \otimes W)^*$ are just the k-valued bilinear map on $V \times W$. But if $\delta \in V^*$ and $\eta \in W^*$, clearly their product, $\delta.\eta$ gives a bilinear map $(v, w) \mapsto \delta(v).\eta(w)$ $(v \in V, w \in W)$. Thus mulitplication defines a map $m\colon V^* \times W^* \to (V \otimes W)^*$. Since multiplication is bilinear, it thus induces a linear map which, by abuse of notation, we will again write as $m\colon V^* \otimes W^* \to (V \otimes W)^*$. Arguing in a similar fashion to the proof of Lemma 22.4, you can check that this map is always injective, and hence when V and W are finite-dimensional, since both spaces then have dimension $\dim(V).\dim(W)$, the map is an isomorphism.

One can also relate the map m to the contraction map ι: Indeed let c be the composition

$$(V \otimes V) \otimes (V^* \otimes V^*) \cong (V \otimes V^*) \otimes (V \otimes V^*) \to \mathsf{k} \otimes \mathsf{k} \cong \mathsf{k}$$

where the first isomorphism permutes the middle two factors, the second is the map $\iota \otimes \iota$ 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 c can be views as a k-valued bilinear pairing between $V \otimes V$ and $V^* \otimes V^*$, which in turn can be viewed as a linear map from $d\colon V^* \otimes V^* \to (V \otimes V)^*$. To see that this is just the multiplication map from above, note that

$$\begin{aligned} d(\delta_1 \otimes \delta_2)(v_1 \otimes v_2) &:= c((v_1 \otimes v_2) \otimes (\delta_1 \otimes \delta_2)) \\ &= \iota(v_1 \otimes \delta_1).\iota(\delta_2 \otimes v_2) = \delta_1(v_1).\delta_2(v_2) \\ &= m(\delta_1 \otimes \delta_2)(v_1 \otimes v_2). \end{aligned}$$

Remark 22.6. Notice that this means we obtain, for V finite dimensional, we have an isomorphism d given by $(V \otimes V^*)^* \cong V^* \otimes V^{**} \cong V^* \otimes V$ where in the second isomorphism we use the inverse of the natural map from V to V^{**}. The map d is non-trivial, and you can check that $d(\iota) = \theta^{-1}(I_V)$, where I_V is the identity map $I_V \in \mathrm{Hom}(V, V) = V^* \otimes V$, and $\theta\colon V^* \otimes V \to \mathrm{Hom}(V, V)$ is the map in Lemma 22.4.

22.2. **Background on 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 22.7. 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

$$B(\lambda_1 v_2 + \lambda_2 v_2, w) = \lambda_1 B(v_1, w) + \lambda_2 B(v_2, w),$$

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 $B(v, w) = B(w, v)$. The *radical* of a symmetric bilinear form is the subspace:

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

We say that B is *nondegenerate* if $\text{rad}(B) = \{0\}$. Let $\text{Bil}(V)$ denote the vector space of bilinear forms on V, and[31] $\text{SBil}(V)$ for the space of symmetric bilinear forms on V.

Remark 22.8. Note that $\text{Bil}(V)$ is isomorphic to $(V \otimes V)^*$, since a bilinear map $V \otimes V \to \mathsf{k}$ induces a linear map $V \otimes V \to k$. When V is finite dimensional, $(V \otimes V)^*$ is isomorphic to $V^* \otimes V^*$. One can thus use the isomorphism $V^* \otimes W \cong \text{Hom}(V, W)$ to give another proof of the following Lemma.

Lemma 22.9. *There is a natural isomorphism* $\Theta\colon \mathit{Bil}(V) \to \mathit{Hom}(V, V^*)$.

Proof. Suppose that $B \in \text{Bil}(V)$. Then define $\Theta(B) = \theta \in \text{Hom}(V, V^*)$ as follows: given $v \in V$, let $\theta(v)\colon V \to \mathsf{k}$ be the function given by $\theta(v)(w) = B(v, w)$. The linearity of B in the second variable then ensures that $\theta(v) \in V^*$, while the linearity of B in the first variable ensures that the map $v \mapsto \theta(v)$ is linear, so that θ is induce a linear map from V to V^*.

Conversely, given $\theta\colon V \to V^*$, define B_θ by $B_\theta(v, w) = \theta(v)(w)$. It is easy to see that $B_\theta \in \text{Bil}(V)$, and that $\theta \mapsto B_\theta$ gives a linear map $\text{Hom}(V, V^*) \to \text{Bil}(V)$ which is clearly inverse to the map Θ, so they are both isomorphisms. □

Remark 22.10. The lemma shows that giving a bilinear form on V is equivalent to giving a linear map from V to V^*. Note that we made a choice in the above construction, since given $B \in \text{Bil}(V)$ we could have defined θ by $\theta(v)(w) = B(w, v)$. For symmetric bilinear forms the two possible choices agree, but for arbitrary bilinear forms they establish different isomorphisms.

From now on we will only work with symmetric bilinear forms. The kernel of $\theta = \Theta(B)$ is clearly the subspace of $v \in V$ such that $\theta(v)(w) = 0$ for all $w \in V$, which is exactly the definition of $\text{rad}(B)$ as required.

Now fix $B \in \text{SBil}(V)$. Then if U is a subspace of V, we define

$$U^\perp = \{v \in V : B(v, w) = 0, \forall w \in U\}.$$

If $\theta\colon V \to V^*$ is the associated map from V to V^*, then clearly $U^\perp = \theta^{-1}(U^0)$. When B is nondegenerate, so that θ is an isomorphism, this shows that $\dim(U^\perp) = \dim(V) - \dim(U)$. The next Lemma shows that this can be refined slightly.

Lemma 22.11. *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:*

(1) $\dim(U) + \dim(U^\perp) \geq \dim(V)$.
(2) *The restriction of B to U is nondegenerate if and only if $V = U \oplus U^\perp$.*

Proof. For the first part, define a map $\phi\colon U \to V^*$ by $\phi(u)(v) = B(u, v)$, $(u \in U, v \in V)$. Then $U^\perp$ is by definition exactly the subspace of $(\text{im}(\phi))^0$ (*i.e.* the annihilator of $\text{im}(\phi)$ under the natural identification of V with its double dual $(V^*)^*$). It follows that $\dim(U^\perp) + \dim(\text{im}(\phi)) = \dim(V)$, and hence certainly

$$\dim(U^\perp) + \dim(U) \geq \dim(V).$$

For the second part, it is immediate from the definitions that B is nondegenerate when restricted to U if and only if $U \cap U^\perp = \{0\}$, *i.e.* if and only if the sum of U and $U^\perp$ is direct. But then the equivalence follows immediately from the dimension inequality in the first part.

□

[31] This is not standard notation – it would be more normal to write something like $\text{Sym}^2(V^*)$ but then I'd have to explain why...

22.2.1. *Classification of symmetric bilinear forms.* This subsection is not needed for the course[32] 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 22.12. *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[33]

$$B(v, w) = \frac{1}{2}\big(B(v + w, v + w) - B(v, v) - B(w, w)\big),$$

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 22.13. 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").

23. APPENDIX 2: REPRESENTATION THEORY BACKGROUND

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.

Definition 23.1. 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 23.2. *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. Moreover, 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

[32] So in particular you don't need to know it for any exam...

[33] Note that this identity holds unless 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.

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. □

Definition 23.3. 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.

Lemma 23.4. *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 23.5. *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 15.3 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[34] $(X \oplus W) \cap U = \{0\}$, which contradicts the maximality of W, so we are done. □

Lemma 23.6. *Let $\mathfrak{g}$ be a Lie algebra. A representation is semisimple if and only if every surjective map of representations $q\colon V \to W$ has a right inverse, that is there is a map of representations $s\colon W \to V$ such that $q \circ s = id_W$.*

Proof. Suppose that V is semisimple. Then given a surjection $q\colon V \to W$, we may find a complement U to the subrepresentation $\ker(q)$. Clearly then $q_{|U}$ is an isomorphism from U to W, and its inverse gives the required homomorphism s. For the converse, if U is a subrepresentation, then the quotient map $q\colon V \to V/U$ has a right inverse s. We claim that $V = U \oplus s(V/U)$. Suppose that $u \in V$ lies in the intersection. Then $u = s(w)$ for some $w \in V/U$, and hence (since U is the kernel of q) we have $0 = q(u) = q(s(w) = w$, so that $u = 0$. Thus $U \cap s(V/U) = \{0\}$, and thus by dimension counting we see that $V = U \oplus s(V/U)$ and $s(V/U)$ is a complement as required. □

Remark 23.7. Our proof of Weyl's theorem showed any surjective map of representations splits, hence any representation is semisimple, and hence completely reducible. Thus it only uses the chain of implications:

$$\text{surjections split} \implies \text{semisimple} \implies \text{completely reducible},$$

while above we have shown that in fact all three statements are equivalent.

As we saw in our study of representations of nilpotent Lie algebras, it is not always the case that a representation is a direct sum of irreducible representations, *i.e.* a representation is not always completely reducible, and for these representations the notion of a composition series is useful: It shows that, even if a representation is not completely reducible, we may still think of it as being built out of irreducibles. Indeed using induction on dimension, it is easy to see that any representation may be *filtered* by subrepresentations $0 = V_0 \subsetneq V_1 \subsetneq \ldots \subsetneq V_k = V$ where each successive quotient V_i/V_{i-1} is irreducible. We call such a filtration a *composition series* for V, and the irreducible representations V_i/V_{i-1} $(1 \leq i \leq k)$ are called *composition factors*.

[34] Since both are just expressing the fact that the sum $X + W + U$ is direct.

Although we don't use it in the course, 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 23.8. *(Jordan-Holder). The numbers $[L, V]$ are independent of the composition series.*

Proof. Use induction on the dimension of V. Clearly if this is 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 23.9. If we take $\mathfrak{g} = \mathfrak{gl}_1(\mathsf{k})$, then a representation (V, ρ) of $\mathfrak{g}$ is completely determined by the linear map $\alpha = \rho(1)$. Subrepresentations of V correspond to subspaces which are invariant for the linear map α, and the existence of eigenvalues shows that the irreducible (finite dimensional) representations of $\mathfrak{gl}_1(\mathsf{k})$ are one-dimensional and hence completely reducible representations are the ones for which α is diagonalisable (or in the terminology favoured in this course, "semisimple"!) Notice that Lemma 23.4 then shows that if a linear map is diagonalisable, it is diagonalisable on any invariant subspace.

24. Appendix 3: Exact sequences

The notion of exact sequences and extensions does not play a prominent role in the course, but some of the results are best expressed in this language.

24.1. **Exact sequences of Lie algebras.**

Definition 24.1. We say that the sequence of Lie algebras and Lie homomorphisms

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

is *exact at* $\mathfrak{g}$ if $\text{im}(i) = \ker(q)$. 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 $\text{im}(i) = \ker(q)$. In this case, we say that $\mathfrak{g}$ is an *extension* of $\mathfrak{g}_2$ by $\mathfrak{g}_1$.

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 map $s \colon \mathfrak{g}_2 \to \mathfrak{g}$ such that $q \circ s = \text{id}_{\mathfrak{g}_2}$.

Notice that if an exact sequence

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

is split, then 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. In fact we then have $\mathfrak{g} \cong i(\mathfrak{g}_1) \rtimes s(\mathfrak{g}_2)$ where $s(\mathfrak{g}_2)$ acts by derivations on $i(\mathfrak{g}_1)$ via the adjoint action (since $\mathfrak{g}_1$ is an ideal in $\mathfrak{g}$).

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

Remark 24.2. Part A Group theory discusses split extensions of groups without using this terminology: Indeed there is a similar notion of a short exact sequence of groups: namely it is a sequence

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

where 1 denotes the trivial group, and exactness at G is the condition that $\ker(q) = \text{im}(i)$, while exactness at G_1 implies i is injective and exactness at G_2 implies that q is surjective. Such a sequence is *split* if there is a homomorphism $p \colon G_2 \to G$ such that $q \circ p = \text{id}_{G_2}$. In this case you can check that $G = G_1 \rtimes G_2$, where G_2 acts on G_1 by the composition of p and the conjugation action of G (as G_1 is normal in G). You can check that $G = G_1 \times G_2$ precisely if $p(G_2)$ is normal in G.

Example 24.3. Lemma 12.2 of the main text says 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. This translates into the existence of an exact sequence

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

which in turn says that any Lie algebra is an extension of the semisimple Lie algebra $\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$ by the solvable Lie algebra $\mathrm{rad}(\mathfrak{g})$. By Levi's theorem, Theorem 12.4, this extension is split in characteristic zero.

24.2. **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 24.4. 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. Once again, an exact sequence

$$0 \longrightarrow U \xrightarrow{\alpha} V \xrightarrow{\beta} W \longrightarrow 0$$

is called split if β admits a right inverse.

Example 24.5. Let $\mathfrak{g}$ be a semisimple Lie algebra. Then any short exact sequence of $\mathfrak{g}$-representations splits. This is just a restatement of Corollary 15.15.

Example 24.6. 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 $t \mapsto (t, 0)$. In 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 7.3.

Lemma 24.7. *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.