C2.1 Lie algebras

Kevin McGerty[1]

Michaelmas 2022

[1]Please contact me at mcgerty@maths.ox.ac.uk if you find any errors or ambiguities in these notes.

Contents

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

Index of Notation

$(D^k(\mathfrak{g}))_{k\geq 0}$	the *derived series* of a Lie algebra $\mathfrak{g}$.
$\mathbb{N}.X$	for any subset X of an abelian group A (*e.g.* a vector space) this denotes the set of all sums of the form $\sum_{y\in Y} n_y.y$ where Y is a finite subset of X and $n_y \in \mathbb{N}$.
$\mathfrak{gl}_V$	the Lie algebra of all endomorphisms of a vector space V equipped with the commutator bracket.
$I \trianglelefteq \mathfrak{g}$	I is an *ideal* in $\mathfrak{g}$, that is, I is a linear subspace and for all $x \in \mathfrak{g}$ and $a \in I$, we have $[a,x] \in I$
$\kappa = \kappa^{\mathfrak{g}}$	the *Killing form*, an invariant symmetric bilinear form on a Lie algebra $\mathfrak{g}$ given by: $$\kappa(x,y) = \mathrm{tr}_{\mathfrak{g}}(\mathrm{ad}(x)\mathrm{ad}(y)).$$
$(C^k(\mathfrak{g}))_{k\geq 0}$	the *lower central series* of a Lie algebra $\mathfrak{g}$.
$[.,.]$	an alternating bilinear map satisfying the Jacobi identity, known as a *Lie bracket*.
$\mathrm{Mat}_{n,m}(\mathsf{k})$	the space of $n \times m$ matrices with entries in a field k.
$\mathrm{Mat}_n(\mathsf{k})$	the space of $n \times n$ matrices with entries in a field k.
$\mathbb{N}$	the natural numbers.
$N_{\mathfrak{g}}(\mathfrak{a})$	the *normalizer* of a subalgebra $\mathfrak{a}$ in a Lie algebra $\mathfrak{g}$.
$\mathrm{rad}(B)$	the *radical* of a symmetric bilinear form B, consisting of all $v \in V$ which satisfy $B(v,w) = 0$ for all $w \in V$.
$\mathrm{rad}(\mathfrak{g})$	the maximal solvable ideal in a Lie algebra $\mathfrak{g}$.
(V,Φ)	an abstract root system.
$\mathfrak{sl}(V)$	the Lie subalgebra $\{\alpha \in \mathfrak{gl}_V : \mathrm{tr}(\alpha) = 0\}$ of traceless endomorphisms of V, known as the *special linear Lie algebra* associated to V.
$\langle X\rangle_F$	if X is a subspace of a k-vector spaces V and $F \leq \mathsf{k}$ is a subfield of k, then the F-linear span of X in V, *i.e.* the intersection of all F-subspaces of V containing X is denoted by either $\langle X\rangle_F$ or $\mathrm{span}_F(X)$.

Modifications

General: The content of Chapters 5 and 6 rebalanced by putting the material on Cartan subalgebras together with the Cartan criteria. Most of the symbols used now should have hyperlinks that link to when the term was first defined in the notes (but there is still an index of notation at the start for paper copies).

Background material: Added a few more details about the examples discussed in lecture 1, but this is just for curiosity (none of the material in the background section is examinable).

(i) Chapter 1: Clarified the definition of a k-algebra to tidy up the unital and non-unital cases: the current version gives a definition of a k-algebra structure on a k-vector space A with or without a unit. In the case where A has a unit, then the definition becomes equivalent to the existence of a homomorphism of rings $\mathsf{k} \to A$ whose image lies in the centre of A.

(ii) Chapter 2: a few more details have been added on things like dual representations. All of these basic properties work exactly the same way for Lie algebras as they did for groups once you figure out how the Lie algebra acts, so in lectures I did not review this material in great detail, but it may be useful to have it written down carefully for reference purposes.

(iii) Chapter 3 explains that the same idea one uses for finite groups – composition series – can be used to show how an arbitrary (finite-dimensional) Lie algebra can be built up from atomic or "almost simple" Lie algebras. The proofs of Jordan-Holder etc. in §3.1 are (still) non-examinable, but they are neater than in the previous version of the notes.[1]

(iv) Chapter 4 has been reordered to cover first the "structural results" on solvable and nilpotent Lie algebras and then the representation theory of both. Some typographical error have been corrected. Two key results of this chapter are Lie's theorem and the theorem that if (V, ρ) is a $\mathfrak{g}$-representation such that $\rho(x)$ is nilpotent for all $x \in \mathfrak{g}$ then the image $\rho(\mathfrak{g})$ of $\mathfrak{g}$ is nilpotent subalgebra of $\mathfrak{gl}_V$ – in the terminology of these notes, if (V, ρ) is a nilpotent representation then $\rho(\mathfrak{g})$ is a nilpotent Lie algebra. The chapter ends with a (fairly coarse) classification of the representations of a nilpotent Lie algebra: they decompose into "generalised weight spaces" in the same way that a vector space decomposes into a direct sum of generalised eigenspaces. The old version of the notes had a mistake in the proof of Proposition 5.3.19. The updated notes replace Proposition 5.3.19 with Lemma 4.3.10, as was done in the lectures, but deduces the main result in slightly neater way to the way it was done in lectures.

(v) Chapter 5: Now starts with the Cartan decomposition. (The proof that $\mathfrak{h} = \mathfrak{g}_0$ some how seems to have been cut from the previous online notes, but it is now part of Lemma 5.1.2). Section 5.2 discusses trace forms, and has a few more remarks which may help to connect it more clearly to Part A Linear Algebra. The proof of Cartan's criterion for solvability is also a little cleaner than the proof given in lectures.

(vi) Chapter 6 discusses the solvable radical, semisimple Lie algebras and the Cartan criterion for semisimplicity. The criterion is then used to show semisimple Lie algebras are a direct sum of simple Lie algebras, as done in lectures. The Jordan decomposition is a pretty easy consequence of the fact that a semisimple Lie algebra is a direct sum of *non-abelian* almost simple Lie algebras, as this shows any derivation is inner. The proof of Weyl theorem is now modelled as closely as possible on the startegy of proof of Maschkhe's theorem for finite groups. (There should also be video lectures for these two topics online now – check the Moodle site).

(vii) Chapter 7 analyses the Cartan decomposition in the semisimple case. It has slightly cleaner proofs than the previous set of notes. The material on abstract root systems is also slightly cleaner and shorter.

[1] And neater than in most textbooks, which, for reasons that I can only assume are historical, pointlessly fixate on phrasing the result as the existence of a permutation of the composition factors on one composition series giving you the composition factors of a second composition series. While it is true such a permutation exists, the content of the theorem is that, for a given simple object S, the number of composition factors in a composition series which are isomorphic to S is independent of the composition series.

*Background

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

From group actions to group representations

In mathematics, group actions give a way of encoding the symmetries of a space or physical system. Formally these are defined as follows: an action of a group G on a space[2] 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. $(g_1g_2).x = g_1.(g_2.x)$ for all $g_1, g_2 \in G$ and $x \in X$.

Natural examples of group actions are that of the general linear group $\mathrm{GL}_n(\mathbb{R})$ on $\mathbb{R}^n$, or the action of the group of rigid motions SO_3 on S^2, the unit sphere $\{x \in \mathbb{R}^3 : \|x\| = 1\}$ in $\mathbb{R}^3$.

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

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

defines a representation of G on $\mathrm{Fun}(X)$. (The inverse is necessary for $\mathrm{Fun}(X)$ is to be a left, rather than right, representation.) 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.

Infinitesimal symmetries

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. 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 $\mathscr{C}^\infty(\mathbb{R}^2)$ 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 $\nu\colon \mathscr{C}^\infty(\mathbb{R}^2) \to \mathscr{C}^\infty(\mathbb{R}^2)$, where if $z = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, then ν is given by

[2] 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 $\mathbb{R}^n$.

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

The operator we obtained in this example, $\nu = y\partial_x - x\partial_y$ is a $\mathscr{C}^\infty(\mathbb{R}^2)$- linear combination of ∂_x and ∂_y. Operators of this form encode "infinitesimal symmetries". The next definition formalises its key properties. We will work with the space $\mathbb{R}^n$ for the rest of this section, but everything we say also applies, *mutatis mutandis* to the context of smooth manifolds.

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

$$\nu(f_1.f_2) = \nu(f_1).f_2 + f_1.\nu(f_2). \tag{0.0.1}$$

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

Lemma. If $\nu\colon \mathscr{C}^\infty(\mathbb{R}^n) \to \mathscr{C}^\infty(\mathbb{R}^n)$ is a derivation, and $a_j = \nu(x_j) \in \mathscr{C}^\infty(\mathbb{R}^n)$, then $\nu = \sum_{j=1}^n a_j\partial_j$, that is, for all $f \in \mathscr{C}^\infty(\mathbb{R}^n)$ we have

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

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

Definition. A *vector field* on $X = \mathbb{R}^n$ is a (smooth) function $\nu\colon \mathbb{R}^n \to \mathbb{R}^n$. To any vector field $\nu = (a_i)_{i=1}^n$ one can associate the derivation $\theta_\nu = \sum_{j=1}^n a_j\partial_j$ which one can think of as giving the infinitesimal direction of a flow (*e.g.* of a fluid, or an electric field say). Thus the space of vector fields $\Theta_{\mathbb{R}^n}$ on $\mathbb{R}^n$ acts on $\mathscr{C}^\infty(\mathbb{R}^n)$, and thus inherits a nonassociative product $[.,.]$ where $\theta_{[\nu_1,\nu_2]} = [\theta_{\nu_1}, \theta_{\nu_2}]$. Explicitly, if $\nu_1 = (a_i)_{i=1}^n$, $\nu_2 = (b_j)_{j=1}^n$ then $[\nu_1, \nu_2] = \Big(\theta_{\nu_1}(b_i) - \theta_{\nu_2}(a_i)\Big)_{i=1}^n$. Such fields can be made to act on functions $f\colon X \to \mathbb{R}$ by differentiation. If $\nu = (a_1, a_2, \dots, a_n)$ in standard coordinates (here $a_i : \mathbb{R}^n \to \mathbb{R}$), then set $\nu(f) = \sum_{i=1}^n a_i\partial_i(f)$. By the previous Lemma, this yields a bijection between vector fields and derivations on $\mathscr{C}^\infty(\mathbb{R}^n)$.

Heuristically, we think of the infinitesimal version of a group action as the collection of derivations on smooth functions obtained 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 this will yield a larger space of derivations).

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: the fact that the commutator product preserves them is a sort of infinitesimal remnant of the group multiplication[4].

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

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

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

But since, for each $g \in \mathrm{SO}_3(\mathbb{R})$, the derivation v_g we obtain in this way, is determined up to scaling by the axis of rotation, and if we conjugate g by an element of $h \in \mathrm{SO}_3(\mathbb{R})$, then hgh^{-1} is a rotation by the same angle around the axis $h(L)$ and $h\gamma(t)h^{-1}$ is a path from the identity to hgh^{-1}. Applying the chain rule as for the case of a circle, noting that a linear map is its own derivative, it follows that the derivation obtained from using the rotation hgh^{-1} in place of g is obtained from that for g simply by applying h. It follows from this that the linear span of all such derivations is in fact a 3-dimensional vector space $\mathfrak{g} = \langle\{x\partial_y - y\partial_x, y\partial_z - z\partial_y, z\partial_x - x\partial_z\}\rangle_{\mathbb{R}}$, 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).

Chapter 1

Lie algebras: Definition and basic notions

1.1 Definitions and Examples

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

Definition 1.1.1. Let R be a commutative ring[1]. An R*-algebra* is a pair $(A, *)$ consisting of an R-module A and an R-bilinear map $*\colon A \times A \to A$, that is, for all $a_1, a_2, b_1, b_2 \in A$ and $r \in R$, the operation $*$ satisfies:

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

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

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

Definition 1.1.2. A Lie algebra over a field k is a k-algebra $(\mathfrak{g}, [.,.]_{\mathfrak{g}})$ 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 1.1.3.
1. 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.
2. If $\mathfrak{g}$ is a Lie algebra, then the alternating property implies that $\mathfrak{g}$ cannot have a unit. It is also almost never the case that an associative product will satisfy the conditions to be a Lie bracket. Thus, viewed a Lie algebra is (usually) a non-commutative, non-associative, and non-unital algebra.[2]
3. We will normally write $[.,.]$ for the Lie bracket on any Lie algebra and decorate it only for emphasis or where there is the potential for confusion.

Definition 1.1.4. Let $(\mathfrak{g}_1, [.,.]_1)$ and $(\mathfrak{g}_2, [.,.]_2)$ be Lie algebras. A k-linear map $\phi\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ is said to be a *homomorphism* of Lie algebras if it respects the Lie brackets. That is:

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

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

[1] All *commutative* rings in this course will have a multiplicative identity.
[2] This makes them sound awful. However, as we will see this is not the way to think about them!

Example 1.1.5. i) If $\dim_{\mathsf{k}}(\mathfrak{g}) = 1$, then the alternating condition forces the Lie bracket to vanish. Thus, up to isomorphism, there is a unique 1-dimensional Lie algebra over k, that is, any 1-dimensional Lie algebra $\mathfrak{g}$ is isomorphic to k equipped with the zero Lie bracket.

ii) If $\mathfrak{a}$ is any vector space then setting the Lie bracket $[.,.]$ to be zero, *i.e.* setting $[a,b] = 0$ for all $a, b \in \mathfrak{a}$, we get a (not very interesting) Lie algebra. Such Lie algebras are called *abelian* Lie algebras.

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

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

where,in the final expression, we have paired terms which can be obtained from each other by cycling x, y and z. Since the terms in these pairs have opposite signs, it is then clear that adding the three expressions obtained by cycling x, y and z gives zero. We will write $\mathfrak{g}_A$ for the Lie algebra $(A, [.,.])$ obtained from an associative algebra in this way.

iv) For a more down-to-earth example, recall that the space of n-by-n matrices $\mathrm{Mat}_n(\mathsf{k})$ with entries in k becomes an associative algebra under matrix multiplication. We therefore obtain a Lie algebra, which we will denote by $\mathfrak{gl}_n(\mathsf{k})$, by equipping $\mathrm{Mat}_n(\mathsf{k})$ with the commutator bracket

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

If the field k is clear from context we will just write $\mathfrak{gl}_n$. Slightly more abstractly, if V is a k-vector space, then we will write $\mathfrak{gl}_V$ for the Lie algebra obtained from the associative algebra $\mathrm{End}_{\mathsf{k}}(V) = \mathrm{Hom}_{\mathsf{k}}(V, V)$ by equipping it with the commutator bracket.[3]

v) If $\dim(V) = 1$ then $\mathfrak{gl}_V = \mathfrak{gl}_1 = \mathsf{k}$: the action of scalars gives an injective map $\mathsf{k} \to \mathrm{End}(V)$ for any nonzero vector space V which is an isomorphism if $\dim(V) = 1$. We will therefore write $\mathfrak{gl}_1$ for k viewed as a Lie algebra with zero Lie bracket.

vi) If $\mathfrak{g}$ is a Lie algebra and $\mathfrak{s} \leq \mathfrak{g}$ is a k-subspace of $\mathfrak{g}$ on which the restriction of the Lie bracket takes values in $\mathfrak{s}$, so that it induces a bilinear operation $[.,.]_{\mathfrak{s}} \colon \mathfrak{s} \times \mathfrak{s} \to \mathfrak{s}$, then $(\mathfrak{s}, [.,.]_{\mathfrak{s}})$ is clearly a Lie algebra, and we say $\mathfrak{s}$ is a (Lie) *subalgebra* of $\mathfrak{g}$. If $\mathfrak{s}$ is a Lie subalgebra of $\mathfrak{g}$ then the inclusion map $i \colon \mathfrak{s} \to \mathfrak{g}$ is a homomorphism of Lie algebras.

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 $\mathrm{Mat}_n(\mathsf{k})$ provided $n > 1$). Similarly we define $\mathfrak{sl}_V$ to be the Lie subalgebra of $\mathfrak{gl}_V$ consisting of endomorphisms of trace 0. These are called *special linear* Lie algebras. More generally we say any Lie subalgebra of $\mathfrak{gl}_V$ for a vector space V is a *linear Lie algebra*.

vii) If $\mathfrak{g}$ is a k-Lie algebra and $x \in \mathfrak{g}$, then the map $\phi_x \colon \mathfrak{gl}_1(\mathsf{k}) \to \mathfrak{g}$ given by $\phi_x(t) = t.x$ is a Lie algebra homomorphism, because the alternating property means that a Lie bracket vanishes on any 1-dimensional subspace of a Lie algebra. This gives a bijection between Lie algebra homomorphisms $\phi \colon \mathfrak{gl}_1(\mathsf{k}) \to \mathfrak{g}$ and the elements of $\mathfrak{g}$ where if $x \in \mathfrak{g}$ we let $x \mapsto \phi_x \colon \mathfrak{gl}_1(\mathsf{k}) \to \mathfrak{g}$ as above, while given $\phi \colon \mathfrak{gl}_1(\mathsf{k}) \to \mathfrak{g}$ we associate to it $\phi(1) \in \mathfrak{g}$.

viii) If $\mathfrak{g}_1, \mathfrak{g}_2$ are Lie algebras, then we may form their *direct sum* $\mathfrak{g}_1 \oplus \mathfrak{g}_2$, which is the direct sum of $\mathfrak{g}_1$ and $\mathfrak{g}_2$ as a vector space, with Lie bracket given by $[(x_1, x_2), (y_1, y_2)] = ([x_1, y_1], [x_2, y_2])$ for all $x_1, y_1 \in \mathfrak{g}_1, x_2, y_2 \in \mathfrak{g}_2$. We may define the direct sum of $k \geq 2$ Lie algebras in the same way.

ix) If $\mathfrak{a}$ is an abelian Lie algebra then if we chose a basis $\{e_1, \dots, e_k\}$ of $\mathfrak{a}$, then we obtain an isomorphism $\theta \colon \mathfrak{gl}_1(\mathsf{k})^{\oplus k} \to \mathfrak{a}$ where $\theta(t_1, \dots, t_k) = \sum_{i=1}^{k} t_i e_i$. Indeed the Lie bracket on both $\mathfrak{a}$ and $\mathfrak{gl}_1(\mathsf{k})^{\oplus k}$ is zero, hence we need only check that θ is an isomorphism of vector spaces, which is clear by construction.

The following definition should be understood as the infinitesimal analogue of an automorphism of a k-algebra.

[3] If it is not clear from context which field k the vector space V is over, we will write $\mathfrak{gl}_{\mathsf{k}}(V)$.

Definition 1.1.6. Generalising the example of vector fields in the previous chapter, 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 $\mathrm{Der}_k(A)$ denotes the k-vector space of k-derivations on A, then $\mathrm{Der}_k(A)$ is stable under taking commutators, that is, if

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

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

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

Example 1.1.7. i) If A is an associative k-algebra, then if $a \in A$ the operation of taking commutator with a is a derivation. That is, if $\delta_a\colon A \to A$ is given by $\delta_a(b) = [a, b]$ for any $b \in A$, then $\delta_a \in \mathrm{Der}_k(A)$. Indeed

$$\delta_a(b).c + b.\delta_a(c) = (ab - ba)c + b(ac - ca) = a.(bc) - (bc).a = \delta_a(b.c)$$

The map $\Delta\colon \mathfrak{g}_A \to \mathrm{Der}_k(A)$ given by $\Delta(a) = \delta_a$ is a homomorphism of Lie algebras, that is, $\Delta([a, b]) = [\delta_a, \delta_b]$. In fact slightly more is true: if $\partial \in \mathrm{Der}_k(A)$ and $b \in A$ then $[\partial, \delta_b] = \delta_{\partial(b)}$. (Applying this to $\partial = \delta_a$ gives the compatibility with commutators). Indeed for all $c \in \mathfrak{g}$ we have

$$[\partial, \delta_b](c) = \partial(bc - cb) - (b\partial(c) - \partial(c).b) = \partial(b).c - c.\partial(b) = \delta_{\partial(b)}(c).$$

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

iii) One way of interpreting the Jacobi identity is that, assuming the alternating property, it is equivalent to the condition that, for any $x \in \mathfrak{g}$, the operation $\mathrm{ad}(x) \in \mathfrak{gl}_{\mathfrak{g}}$ given by $\mathrm{ad}(x)(y) = [x, y]$ lies in $\mathrm{Der}_k(\mathfrak{g})$. Indeed

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

where the equivalence between the third and fourth equalities follows from the alternating property of a Lie bracket.

iv) The Jacobi identity is also equivalent, again assuming the alternating property, to the fact that $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}_{\mathfrak{g}}$ is a homomorphism of Lie algebras: Indeed, for all $x, y, z \in \mathfrak{g}$ we have

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

where the second and fourth equality uses the alternating property, and the third the Jacobi identity.

v) If $(A, *)$ is any k-algebra then A^{op} is the k-algebra with product $*^{\mathrm{op}}$, where $a *^{\mathrm{op}} b = b * a$. If A is commutative, then A^{op} is isomorphic to A. In the case of a Lie algebra $\mathfrak{g}$ then $\mathfrak{g}^{\mathrm{op}}$ is the Lie algebra $(\mathfrak{g}, -[.,.])$. In fact $\mathfrak{g}$ is canonically isomorphic to $\mathfrak{g}^{\mathrm{op}}$: if we let $m\colon \mathfrak{g} \to \mathfrak{g}^{\mathrm{op}}$ be the map $m(x) = -x$, then

$$m([x, y]) = -[x, y] = [y, x] = [x, y]^{\mathrm{op}} = [-x, -y]^{\mathrm{op}} = [m(x), m(y)], \quad x, y \in \mathfrak{g}.$$

hence $m\colon \mathfrak{g} \to \mathfrak{g}^{\mathrm{op}}$ is an isomorphism from $\mathfrak{g}$ to $\mathfrak{g}^{\mathrm{op}}$.

***Remark 1.1.8.** Combining *iii*) and *iv*) in the above example we see that the adjoint representation $x \mapsto \mathrm{ad}(x)$ is in fact a Lie algebra homomorphism from $\mathfrak{g}$ to $\mathrm{Der}_k(\mathfrak{g})$. This is, in a sense, where the Jacobi identity comes from: very roughly, the conjugation action of G on itself yields a group homomorphism $G \to \mathfrak{gl}_{\mathfrak{g}}$ (since conjugation preserves the identity $e \in G$) whose image lies in $\mathrm{Aut}(\mathfrak{g})$. The adjoint representation of $\mathfrak{g}$ is then the derivative of this action yields the adjoint representation ad which hence should have image in $\mathrm{Der}_k(\mathfrak{g})$.

1.2 Ideals and isomorphism theorems

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

Definition 1.2.1. An *ideal* in a Lie algebra $(\mathfrak{g}, [.,.]_{\mathfrak{g}})$ is a subspace $\mathfrak{a}$ such that for all $x \in \mathfrak{g}$ and $a \in \mathfrak{a}$ we have $[a, x]_{\mathfrak{g}} \in \mathfrak{a}$. It is easy to check that if $\phi\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ is a homomorphism, then

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

is an ideal of $\mathfrak{g}_1$. We will write $I \trianglelefteq \mathfrak{g}$ to indicate that I is an ideal in $\mathfrak{g}$.

Remark 1.2.2. Notice that because a Lie bracket is alternating, the condition that, for all $x \in \mathfrak{g}$ and $a \in \mathfrak{a}$ one has $[a, x] \in \mathfrak{a}$, is equivalent to the condition that $[x, a] \in \mathfrak{a}$ for all $x \in \mathfrak{g}$, $a \in \mathfrak{a}$. Thus, similarly to commutative rings, the notions of a left, right or two-sided ideal in a Lie algebra are all the same.

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

Theorem 1.2.3. *(The first isomorphism theorem:) Let $\mathfrak{a}$ be an ideal in a Lie algebra $\mathfrak{g}$, and let $q\colon \mathfrak{g} \to \mathfrak{g}/\mathfrak{a}$ be the quotient map of vector spaces. Then there is a unique Lie bracket on $\mathfrak{g}/\mathfrak{a}$ with respect to which q is a homomorphism of Lie algebras, that is, for all $x, y \in \mathfrak{g}$*

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

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

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

Definition 1.2.4. If V, W are subspaces of a Lie algebra $\mathfrak{g}$, then write $[V, W]$ for the linear span of the elements $\{[v, w] : v \in V, w \in W\}$. Notice that if I, J are ideals in $\mathfrak{g}$ then so is $[I, J]$. Indeed to check this, note that by part 8) of Example 1.1.5, if $z \in \mathfrak{g}$, $x \in I$, $y \in J$ then we have

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

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

Remark 1.2.5. If I and J are ideals in a Lie algebra $\mathfrak{g}$ then it is easy to check that their intersection $I \cap J$ is again an ideal in $\mathfrak{g}$, and we have $[I, J] \subseteq I \cap J$. (Thus $[I, J]$ is the Lie algebra analogue of the product of ideals in a commutative ring.) Similarly, it is easy to see that the linear sum $I + J$ of I and J is also an ideal[4].

Definition 1.2.6. Let $\mathfrak{g}$ be a Lie algebra and let $\mathfrak{a} \leq \mathfrak{g}$ be a subalgebra. The *normalizer* of $\mathfrak{a}$ in $\mathfrak{g}$ is

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

This is a subalgebra of $\mathfrak{g}$, as one can check using the formulation of the Jacobi identity given in Definition 1.2.4. It is the largest subalgebra of $\mathfrak{g}$ within which $\mathfrak{a}$ is an ideal.

Definition 1.2.7. If a nontrivial Lie algebra has no nontrivial ideals we say that it is *almost simple*. It it is in addition not abelian, *i.e.* the Lie bracket is not identically zero, then we say that it is *simple*.

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

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

Theorem 1.2.8. *i) If* $\mathfrak{h}$ *is a subalgebra of* $\mathfrak{g}$ *and* I *is an ideal in* $\mathfrak{g}$ *then* $\mathfrak{h}+I$ *is a subalgebra of* $\mathfrak{g}$ *(containing* I *as an ideal)* $\mathfrak{h}\cap I$ *is an ideal in* $\mathfrak{h}$*, and*

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

ii) If $J \subseteq I \subseteq \mathfrak{g}$ *are ideals of* $\mathfrak{g}$ *then we have:*

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

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

Chapter 2

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 (usually more concrete) objects. For Lie algebras, since they are already vector spaces over k, it is natural to study their action on linear spaces, or in other words, "representations".

2.1 Definition and examples

Definition 2.1.1. A *representation* of a Lie algebra $\mathfrak{g}$ is a vector space V equipped with a linear action of $\mathfrak{g}$, that is, 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 also refer to a representation of $\mathfrak{g}$ as a $\mathfrak{g}$-representation or $\mathfrak{g}$-module. A representation is *faithful* if $\ker(\rho) = 0$. When there is no danger of confusion we will normally suppress ρ in our notation, and write $x(v)$ rather than $\rho(x)(v)$, for $x \in \mathfrak{g}, v \in V$.

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

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

Example 2.1.2. i) If V is a k-vector space, then the identity map $\mathfrak{gl}_V \to \mathfrak{gl}_V$ gives 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 action map ρ is just the inclusion map.

ii) Let $\mathfrak{a}$ be an abelian Lie algebra. If (V,ρ) is a representation of $\mathfrak{a}$, then the image $\rho(\mathfrak{a})$ of $\mathfrak{a}$ is a commutative subalgebra of $\mathrm{End}_{\mathsf{k}}(V)$: if $a,b \in \mathfrak{a}$ then $0 = \rho([a,b]) = \rho(a)\rho(b) - \rho(b)\rho(a)$, so that

$$\rho(a)\rho(b) = \rho(b)\rho(a), \forall\, a,b \in \mathfrak{a}.$$

iii) 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 $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}_{\mathfrak{g}}$ from $\mathfrak{g}$ to $\mathfrak{gl}_{\mathfrak{g}}$ is given by

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

Indeed, as noted in iv) of Example 1.1.7, the fact that this map is a homomorphism of Lie algebras is just a rephrasing[2] of the Jacobi identity. Note that while the vector representation is clearly faithful, in general the adjoint representation is not. Indeed the kernel is known as the *centre* of $\mathfrak{g}$:

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

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

[1] If you take the Category Theory course, $\mathrm{Rep}(\mathfrak{g})$ is a category whose objects are representations of $\mathfrak{g}$ and whose morphisms are $\mathfrak{g}$-homomorphisms. The term "collection" is used because of set-theoretic subtleties which we can essentially ignore in this course.

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

iv) If $\mathfrak{g}$ is any Lie algebra, the pair $(\mathsf{k}, 0)$ consisting of the vector space k together with the zero map $0\colon \mathfrak{g} \to \mathfrak{gl}_1$ is a $\mathfrak{g}$-representation. This 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 map $1_V \in \mathrm{GL}(V)$).

v) If (V, ρ) is a $\mathfrak{g}$-representation, then the sum of all subrepresentations of V isomorphic to $(\mathsf{k}, 0)$ is a natural subrepresentation

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

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

vi) If (V, ρ) is a representation of a Lie algebra $\mathfrak{g}$ and $\theta\colon \mathfrak{h} \to \mathfrak{g}$ is a homomorphism of Lie algebras, then we define the *pull-back* of (V, ρ) to the representation of $\mathfrak{h}$ given by $(V, \rho \circ \theta)$. The most common example of a pull-back is *restriction*, when $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$ (and thus θ is the inclusion map from $\mathfrak{h}$ to $\mathfrak{g}$).

The following definitions are useful when studying the the structure of Lie algebra representations:

Definition 2.1.3. 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. A representation V is said to be *indecomposable* if, whenever we have $V = U_1 \oplus U_2$ with U_1, U_2 subrepresentations, either $U_1 = V$ or $U_2 = V$ (and $U_2 = 0, U_1 = 0$ respectively).

It is easy to show (by induction on dimension) that any finite-dimensional representation is a direct sum of indecomposable subrepresentations.

Example 2.1.4. In this example, we classify the representations of the simplest Lie algebra $\mathfrak{gl}_1$: a representation (V, ρ) of $\mathfrak{gl}_1$ is given by a Lie algebra homomorphism $\rho\colon \mathfrak{gl}_1 \to \mathfrak{gl}_V$. But we saw in vii) of Example 1.1.5, that there is a natural bijection between such homomorphisms and elements $x \in \mathfrak{gl}_V$ given by $\rho \mapsto \rho(1)$. Through this correspondence the problem of classifying $\mathfrak{gl}_1$-representations up to isomorphism becomes the problem of classifying vector spaces equipped with an endomorphism up to conjugacy.

If we assume k is algebraically closed this classification of linear endomorphisms is given by the Jordan canonical form. It is a useful exercise to translate statements about linear maps into statements about representations of $\mathfrak{gl}_1$. For example, the irreducible representations of $\mathfrak{gl}_1$ are the one-dimensional ones, and correspond to eigenvectors of $\rho(1)$. What do the indecomposable representations correspond to?

Example 2.1.5. Now suppose that $\mathfrak{g}$ is any finite-dimensional Lie algebra and that (L, λ) is a one-dimensional representation of $\mathfrak{g}$. The canonical identifications $\mathfrak{gl}_L \cong k = \mathfrak{gl}_1(\mathsf{k})$ given in v) of Example 1.1.5 identifies λ with a homomorphism $[\lambda]\colon \mathfrak{g} \to \mathfrak{gl}_1$. But since $\mathsf{k} = \mathfrak{gl}_1$ has the zero Lie bracket, this is just an element of $\mathfrak{g}^*$ which vanishes on $D(\mathfrak{g})$, that is, an element of $D(\mathfrak{g})^0 \cong (\mathfrak{g}/D(\mathfrak{g}))^*$. The homomorphism $[\lambda]$ clearly identifies (L, λ) up to isomorphism, and given any $\mu \in D(\mathfrak{g})^0$, we obtain (k, μ) a canonical representative of this isomorphism class, which we will denote by k_μ. In particular, if $\lambda = 0 \in D(\mathfrak{g})^0$ then k_0 is the trivial representation of $\mathfrak{g}$.

2.2 Subrepresentations, quotients, duals, and composition series

There are a number of standard ways of constructing new representations from old, all of which have their analogues in the context of group representations. We begin with some definitions.

Definition 2.2.1. Let V be a k-vector space and $U \leq V$ a subspace. Write $i\colon U \to V$ for the inclusion map and $p\colon V \to V/U$ for the quotient map. Let

$$\mathfrak{b}_U = \{x \in \mathfrak{gl}_V : x(U) \subseteq U\} = \{x \in \mathfrak{gl}_V : p \circ x \circ i = 0\}.$$

The we have linear maps $i^*\colon \mathfrak{b}_U \to \mathrm{End}(U)$ and $p_*\colon \mathfrak{b}_U \to \mathrm{End}(V/U)$ given by $i^*(x) = x \circ i$, and $p_*\colon \mathfrak{b}_U \to \mathrm{End}(V/U)$ is given by $p_*(x)(v + U) = p(x(v)) = x(v) + U$ for any $x \in \mathfrak{b}_U, v \in V$.

Lemma 2.2.2. *If $\rho\colon \mathfrak{g} \to \mathfrak{gl}_V$ is a $\mathfrak{g}$-representation and $\rho(\mathfrak{g}) \subseteq \mathfrak{b}_U$ for some subspace $U \leq V$, then U and V/U become $\mathfrak{g}$-representations with action maps $i^* \circ \rho$ and $p_* \circ \rho$ respectively.*

Proof. It is clear that $\mathfrak{b}_U$ is an associative subalgebra of $\mathrm{End}_{\mathsf{k}}(V)$ and both i^* and p_* are homomorphisms of associative algebras, hence they are also homomorphisms of the associated Lie algebras. The Lemma follows immediately. □

It will be useful later to have the following definition:

Definition 2.2.3. Let V be a vector space, and let $\mathscr{F} = (F_i)_{i=0}^k$ be a *flag* in V, that is

$$\mathscr{F} = (V = F_0 \supset F_1 \supset F_2 \supset \ldots \supset F_k = 0)$$

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

We let $\mathfrak{b}_{\mathscr{F}} = \bigcap_{1 \leq i \leq k-1} \mathfrak{b}_{F_i} = \{x \in \mathfrak{gl}_V : x(F_i) \subseteq F_i\}$. This is an associative subalgebra of $\mathrm{End}(V)$, and hence a Lie subalgebra. If (V, ρ) is a $\mathfrak{g}$-representation, the elements of the flag are subrepresentations of V if and only if $\rho(\mathfrak{g}) \subseteq \mathfrak{b}_{\mathscr{F}}$.

Definition 2.2.4. If V is a k-vector space and $a \in \mathfrak{gl}_V$, then a induces a linear map $a^\top : V^* \to V^*$ which we call the *adjoint* (or *transpose*) of a, given by $a^\top(f)(v) = (f \circ a)(v)$. However, if $a, b \in \mathfrak{gl}_V, v \in V$,

$$(ab)^\top(f)(v) = (f \circ (ab))(v) = ((f \circ a) \circ b)(v) = b^\top(f \circ a)(v) = b^\top \circ a^\top(f)(v),$$

so that $a \mapsto a^\top$ is an algebra anti-homomorphism from $\mathrm{End}_{\mathsf{k}}(V)$ to $\mathrm{End}_{\mathsf{k}}(V^*)^{\mathrm{op}}$ (see 1.1.5 (v))), so that it is also a Lie algebra homomorphism $\mathfrak{gl}_V \to \mathfrak{gl}_{V^*}{}^{\mathrm{op}}$. But, again by 1.1.5 (v)) the map $x \mapsto -x$ is an isomorphism from $\mathfrak{g} \to \mathfrak{g}^{\mathrm{op}}$ for any Lie algebra, it follows that $x \mapsto -x^\top$ is an isomorphism of Lie algebras from $\mathfrak{gl}_V$ to $\mathfrak{gl}_{V^*}$. For $x \in \mathfrak{gl}_V$, we will write $x^* = -x^\top$. Moreover, if (V, ρ) is a $\mathfrak{g}$-representation, then we may define $\rho^*(x) = (\rho(x))^*$, so that $\rho^* : \mathfrak{g} \to \mathfrak{gl}_{V^*}$ is a Lie algebra homomorphism and hence (V^*, ρ^*) is a $\mathfrak{g}$-representation, the *dual* representation to (V, ρ).

Recall that the annihilator of a subspace $U \leq V$ is the subspace of V^* given by

$$U^0 = \{f \in V^* : f(u) = 0, \forall\, u \in U\}.$$

By considering a basis of V and the corresponding dual basis of V^*, it is easy to see that $\dim(U) + \dim(U^0) = \dim(V)$, and the correspondence $U \mapsto U^0$ is order-reversing for containment, that is, if $U_1 \leq U_2$ then $U_2^0 \leq U_1^0$.

Lemma 2.2.5. *If (V, ρ) is a $\mathfrak{g}$-representation with dual representation (V^*, ρ^*), then the map $U \mapsto U^0$ gives an order-reversing correspondence between the subrepresentations of V and V^* respectively. Since V is finite-dimensional, $(V^*)^*$ is canonically isomorphic to V, and via that canonical identification, this correspondence is an involution, that is, $(U^0)^0 = U$.*

Proof. We need only check that if U is a subrepresentation of V then U^0 is a subrepresentation of V^*. But notice that the adjoint of the inclusion map $i^\top : V^* \to U^*$ has kernel U^0, and the adjoint of the quotient map $p^\top : (V/U)^* \to V^*$ has image U^0. Moreover $p^\top$ is clearly injective, and $i^\top$ is surjective since any functional on U extends to one on V (as you can easily see using *e.g.* dual bases).

Now $\mathfrak{b}_{U^0} = \{y \in \mathfrak{gl}_{V^*} : \iota^\top \circ y \circ p^\top = 0\}$. But if $x \in \mathfrak{gl}_V$ then

$$i^\top \circ (x^*) \circ p^\top = -(i^\top \circ x^\top \circ p^\top) = -(p \circ x \circ i)^\top.$$

Hence $x^* \in \mathfrak{b}_{U^0}$ if and only if $p \circ x \circ i = 0$, that is, if and only if $x \in \mathfrak{b}_U$. It follows that if $\rho(\mathfrak{g}) \subseteq \mathfrak{b}_U(V)$ if and only if $\rho^*(\mathfrak{g}) \subseteq \mathfrak{b}_{U^0}(V^*)$ as required.

□

Definition 2.2.6. If (V, ρ) is a finite-dimensional $\mathfrak{g}$-representation, then we say that a flag $\mathscr{C} = (V = F_0 > F_1 > \ldots > F_d = \{0\})$ in V is a *composition series* for V if each F_i is a subrepresentation of V and F_i/F_{i+1} is an irreducible representation of $\mathfrak{g}$.

The *Jordan-Hölder* theorem for finite-dimensional representations of $\mathfrak{g}$ shows that the isomorphism classes of the irreducible $\mathfrak{g}$-representations F_k/F_{k+1} are independent of the choice of composition series, as is the number of times a given simple occurs. (See Appendix II for a proof.) If S is a simple $\mathfrak{g}$-representation and (V, ρ) any $\mathfrak{g}$-representation, we write $[S : V]$ for the number of composition factors in a composition series for V which are isomorphic to S. The next definition will be crucial later in the course, when hopefully it will appear more natural.

Definition 2.2.7. Let $\mathfrak{g}$ be a Lie algebra and suppose that (V, ρ) is a $\mathfrak{g}$-representation. We define a symmetric bilinear form $t_V : \mathfrak{g} \times \mathfrak{g} \to \mathsf{k}$ by setting $t_V(x, y) = \mathrm{tr}_V(\rho(x) \circ \rho(y))$. Clearly t_V depends only on the isomorphism class of V, that is, if $V_1 \cong V_2$ as $\mathfrak{g}$-representations then $t_{V_1} = t_{V_2}$.

Lemma 2.2.8. *Let (V,ρ) be a finite-dimensional representation of $\mathfrak{g}$ and let $x, y \in \mathfrak{g}$. Then*

$$t_V(x,y) = \sum_{S \text{ irreducible}} [S:V] t_S(x,y)$$

Proof. Picking a composition series $\mathscr{C} = (V = F_0 > F_1 > \ldots > F_d = \{0\})$, the image $\rho(\mathfrak{g})$ of $\mathfrak{g}$ in $\mathfrak{gl}_V$ is contained in $\mathfrak{b}_{\mathscr{C}}$ and since $\mathfrak{b}_{\mathscr{C}}$ is an associative algebra, if $\rho(x), \rho(y) \in \mathfrak{b}_{\mathscr{C}}$ then $\rho(x)\rho(y) \in \mathfrak{b}_{\mathscr{C}}$. Thus the Lemma follows if we can show that for all $\alpha \in \mathfrak{b}_{\mathscr{C}}$

$$\mathrm{tr}_V(\alpha) = \sum_{k=0}^{d-1} \mathrm{tr}_{W_k}(\bar{\alpha}_k)$$

where $\bar{\alpha}_k$ is the linear map induced by α on $W_k = F_k/F_{k+1}$ (where $W_{d-1} = F_{d-1}/F_d = F_{d-1}$). But this is easy to check by picking a basis B for V compatible with the composition series (in the sense that, for each $k, 0 \leq k \leq d$, the intersection $B \cap F_k$ is a basis for F_k).

□

Lemma 2.2.9. *Suppose that V is a $\mathfrak{g}$-representation and that $\{S_i : 1 \leq i \leq k\}$ are its composition factors. Then V^* has composition factors $\{S_i^* : 1 \leq i \leq k\}$ and moreover $[S_i : V] = [S_i^* : V^*]$.*

Proof. Let $\mathscr{C} = (V = F_0 > F_1 > \ldots > F_d = \{0\})$, and suppose that $F_i/F_{i+1} \cong S_{\pi(i)}$ so that $[S_i : V] = |\pi^{-1}(i)|$. Then V^* has a filtration by subrepresentations given by the annihilators $\mathscr{C}^* = (V^* = F_d^0 > F_{d-1}^0 > \ldots > F_0^0 = \{0\})$. Now $F_{i+1}^0 \cong (V/F_{i+1})^*$ via the transpose of the quotient map $q_{i+1} \colon V \to V/F_{i+1}$, and $F_i^* \cong V^*/F_i^0$ via the transpose of the inclusion $p_i \colon F_i \to V$. It follows that $(F_i/F_{i+1})^* \cong F_{i+1}^0/F_i^0$, and hence $F_{i+1}^0/F_i^0 \cong S_{p(i)}^*$ is simple, so that $\mathscr{C}^*$ is a composition series for V^*, with composition factors S_i^* with the simple S_i^* having multiplicity $[S_i^* : V^*] = |\pi^{-1}(i)| = [S_i : V]$ as required. □

Definition 2.2.10. If V is a $\mathfrak{g}$-representation, we let $V^s = \sum_{S \leq V} S$ the *socle* of V, be the sum of all irreducible subrepresentations of V. This is a semisimple subrepresentation of V and hence it can be written as the direct sum of irreducible subrepresentations of V. It is maximal among semisimple subrepresentations of V in the partical order given by containment.

2.2.1 Direct sums and Hom-spaces

Now suppose that $V = V_1 \oplus V_2$ is a k-vector space. For $j = 1, 2$, we have natural inclusion maps $\iota_j \colon V_j \to V$ and projection maps $p_j \colon V \to V_j$ (with kernel V_{3-j}). We claim, for any vector space U, we have natural isomorphisms

$$\begin{aligned} &i) \ \mathrm{Hom}(V,U) \cong \mathrm{Hom}(V_1,U) \oplus \mathrm{Hom}(V_2,U) \\ &ii) \ \mathrm{Hom}(U,V) \cong \mathrm{Hom}(U,V_1) \oplus \mathrm{Hom}(U,V_2). \end{aligned} \tag{2.2.1}$$

In the case of i), the map simply takes the restriction of $\phi \in \mathrm{Hom}(V,U)$ to V_1 and V_2 respectively. In terms of our inclusion and projection maps, for $r = 1, 2$ we have $\phi_{|V_r} = \phi \circ \iota_r$. To see that this map is an isomorphism, note that any $v \in V$ can be written uniquely as $v = v_1 + v_2$ with $v_1 \in V_1$ and $v_2 \in V_2$. Indeed $v_r = \iota_r \circ p_r(v)$, hence $\phi(v) = \phi(\iota_1 \circ p_1(v) + \iota_2 \circ p_2(v)) = \sum_{r=1}^2 (\phi \circ \iota_r) \circ p_r(v)$. In other words, the inverse to $\phi \mapsto (\phi \circ \iota_r)_{r=1,2}$ is the map $(\psi_r)_{r=1,2} \mapsto \sum_{r=1}^2 \psi_r p_r$.

In the case of ii), the morphism simply takes, for $\phi \in \mathrm{Hom}(U,V)$, the components of $\phi(u)$ in V_1 and V_2 respectively, that is $\phi \mapsto (p_s \circ \phi)_{s=1;2}$. Clearly the inverse of this map is given by $(\eta_s)_{s=1,2} \mapsto \sum_{s=1,2} \iota_s \circ \eta_s$.

Now consider $\mathrm{End}(V) = \mathrm{Hom}(V,V)$ where $V = V_1 \oplus V_2$. We may use i) and ii) (twice) to obtain

$$\mathrm{Hom}(V,V) \cong \bigoplus_{r=1}^{2} \mathrm{Hom}(V_r, V) \cong \bigoplus_{r=1}^{2} \left(\bigoplus_{s=1}^{2} \mathrm{Hom}(V_r, V_s) \right) = \bigoplus_{r,s=1}^{2} \mathrm{Hom}(V_r, V_s)$$

by $\phi \mapsto (\phi_r^s)$ where $\phi_r^s = p_s \circ \phi \circ \iota_r \in \mathrm{Hom}(V_r, V_s)$. This decomposition is just that of a matrix into block submatrices, so it can be useful to arrange it in that form:

$$\mathrm{End}(V) = \mathfrak{gl}_V \ni \phi \mapsto \begin{pmatrix} \phi_1^1 & \phi_2^1 \\ \phi_1^2 & \phi_2^2 \end{pmatrix} \in \begin{pmatrix} \mathrm{Hom}_{\mathsf{k}}(V_1, V_1) = \mathfrak{gl}_{V_1} & \mathrm{Hom}_{\mathsf{k}}(V_2, V_1) \\ \mathrm{Hom}_{\mathsf{k}}(V_1, V_2) & \mathrm{Hom}_{\mathsf{k}}(V_2, V_2) = \mathfrak{gl}_{V_2} \end{pmatrix} \tag{2.2.2}$$

This shows that $\mathfrak{gl}_{V_1} \oplus \mathfrak{gl}_{V_2}$ is naturally isomorphic to a subalgebra of $\mathfrak{gl}_V = \mathfrak{gl}_{V_1 \oplus V_2}$, and hence $V_1 \oplus V_2$ is a representation of $\mathfrak{gl}_{V_1} \oplus \mathfrak{gl}_{V_2}$. More interestingly, the summands $\mathrm{Hom}(V_1, V_2)$ and $\mathrm{Hom}(V_2, V_1)$ are clearly all preserved by the action of $\mathfrak{gl}_{V_1} \oplus \mathfrak{gl}_{V_2}$, so that in particular, $\mathrm{Hom}(V_1, V_2)$ is a representation of $\mathfrak{gl}_{V_1} \oplus \mathfrak{gl}_{V_2}$, where if $x = (x_1, x_2) \in \mathfrak{gl}_{V_1} \oplus \mathfrak{gl}_{V_2}$ and $\phi \in \mathrm{Hom}(V_1, V_2)$, then $x(\phi) = x_2 \circ \phi - \phi \circ x_1$.

It follows that if (V_1, ρ_1) and (V_2, ρ_2) are $\mathfrak{g}$-representations, then $\mathrm{Hom}(V_1, V_2)$ is a $\mathfrak{g}$-representation, via

$$\mathfrak{g} \xrightarrow{\Delta} \mathfrak{g} \oplus \mathfrak{g} \xrightarrow{\rho_1 \oplus \rho_2} \mathfrak{gl}_{V_1} \oplus \mathfrak{gl}_{V_2} \longrightarrow \mathfrak{gl}(\mathrm{Hom}(V_1, V_2))$$

where $\Delta\colon \mathfrak{g} \to \mathfrak{g} \oplus \mathfrak{g}$ is the diagonal map $\Delta(x) = (x, x)$. Explicitly, if $x \in \mathfrak{g}$ and $\phi \in \mathrm{Hom}(V, W)$ then $x(\phi) = \rho_2(x) \circ \phi - \phi \circ \rho_1(x)$.

Remark 2.2.11. Note that the previous example actually includes the example of dual spaces: if V is a k-vector space then $V^* = \mathrm{Hom}(V, \mathsf{k})$ becomes a $\mathfrak{gl}_V \oplus \mathfrak{gl}_1$-representation, and then simply using the inclusion $\mathfrak{gl}_V \to \mathfrak{gl}_V \oplus \mathfrak{gl}_1$ we see that V^* becomes a $\mathfrak{gl}_V$-representation, and indeed we obtain the same action: for any $f \in V^*$ and $x \in \mathfrak{gl}_V$ we have $x(f) := (x, 0)(f) = 0 \circ f - f \circ x = -x^\top(f)$.

Remark 2.2.12. For any direct sum decomposition $V = \bigoplus_{r=1}^k V_r$, have natural inclusion maps $\iota_r^V\colon V_r \to V$ and projection maps $p_r^V\colon V \to V_r$, where $\ker(p_r) = \bigoplus_{s\neq r} V_s$. It is easy to see that $1_V = \sum_{r=1}^k \iota_r^V \circ p_r^V$ and $p_s^V \circ \iota_r^V = \delta_{rs} 1_{V_r}$. The above discussion then generalises readily to the case where we have k-vector spaces $V = \bigoplus_{r=1}^k V_r$ and $W = \bigoplus_{s=1}^l W_s$. By considering, for any $\phi \in \mathrm{Hom}(V, W)$ the identity $\phi = 1_W \circ \phi \circ 1_V$, we see that

$$\phi \mapsto (p_s^W \circ \phi \circ i_r^V)_{r,s} \quad \text{and} \quad (\phi_r^s)_{r,s} \mapsto \sum_{r,s} \iota_s^W \circ \phi_r^s \circ p_i^V$$

are mutually inverse and give isomorphisms between $\mathrm{Hom}(V, W)$ and $\bigoplus_{r,s} \mathrm{Hom}(V_r, W_s)$ which again is just the decomposition of $\mathrm{Hom}(V, W)$ into "block matrices" corresponding to the direct sum decompositions of V and W respectively.

2.3 Tensor products

First we note a general Lemma:

Lemma 2.3.1. *Let $\mathfrak{g}_1, \mathfrak{g}_2$ be Lie algebras over k and let $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ be their direct sum (so each of $\mathfrak{g}_1, \mathfrak{g}_2$ is an ideal in $\mathfrak{g}$). If (U, ρ) is a representation of $\mathfrak{g}$, and we set $\rho_i = \rho_{|\mathfrak{g}_i}$, then each ρ_i is a representation of $\mathfrak{g}_i$ for $i = 1, 2$, and $[\rho_1(x), \rho_2(y)] = 0$ for any $x \in \mathfrak{g}_1, y \in \mathfrak{g}_2$. Conversely, if $\rho_i\colon \mathfrak{g}_i \to \mathfrak{gl}_U$ are Lie algebra homomorphisms for $i = 1, 2$ and $[\rho_1(x), \rho_2(y)] = 0$ for all $x \in \mathfrak{g}_1, y \in \mathfrak{g}_2$, then $\rho(x, y) = \rho_1(x) + \rho_2(y)$ is a Lie algebra homomorphism from $\mathfrak{g}$ to $\mathfrak{gl}_U$.*

Proof. Given a representation (U, ρ) of $\mathfrak{g}$, the asserted properties of ρ_1, ρ_2 are immediate. For the converse, note that if $(x_1.x_2), (y_1, y_2) \in \mathfrak{g}$, where $x_1.y_1 \in \mathfrak{g}_1$ and $x_2, y_2 \in \mathfrak{g}_2$, then

$$\begin{aligned}
[\rho(x_1, x_2), \rho(y_1, y_2)] &= [\rho_1(x_1) + \rho_2(x_2), \rho_1(y_1) + \rho_2(y_2)] \\
&= [\rho_1(x_1), \rho_1(y_1)] + [\rho_1(x_1), \rho_2(y_2)] - [\rho_1(y_1), \rho_2(x_2)] + [\rho_2(x_2), \rho_2(y_2)] \\
&= [\rho_1(x_1), \rho_1(y_1)] + [\rho_2(x_2), \rho_2(y_2)] \\
&= \rho((x_1, x_2), (y_1, y_2))
\end{aligned}$$

so that ρ is a homomorphism as required. □

Now $\mathfrak{gl}_V$ and $\mathfrak{gl}_W$ are naturally subalgebras of $\mathfrak{gl}_{V\otimes W}$, via the embeddings i_V and i_W respectively, where $i_V(\alpha) = \alpha \otimes 1_W$ and $i_W(\beta) = 1_V \otimes \beta$ respectively. Since for any $\alpha \in \mathfrak{gl}_V, \beta \in \mathfrak{gl}_W$ we have

$$i_W(\beta) \circ i_V(\alpha) = (1_V \otimes \beta) \circ (\alpha \otimes 1_W) = \alpha \otimes \beta = (\alpha \otimes 1_W) \circ (1_V \otimes \beta) = i_V(\alpha) \circ i_W(\beta)$$

it follows by Lemma 2.3.1 that $d\colon \mathfrak{gl}_V \oplus \mathfrak{gl}_W \to \mathfrak{gl}_{V\otimes W}$ given by

$$d(x, y) = i_V(x) + i_W(y) = x \otimes 1_W + 1_V \otimes y$$

is a Lie algebra homomorphism, and hence $V \otimes W$ is naturally a $\mathfrak{gl}_V \oplus \mathfrak{gl}_W$-representation. It follows immediately that if (V, ρ) is a representation of $\mathfrak{g}_1$ and (W, σ) is a representation of $\mathfrak{g}_2$ then $V \otimes W$ is a representation of $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ via $d \circ (\rho \oplus \sigma)$ and if $\mathfrak{g}_1 = \mathfrak{g}_2 = \mathfrak{g}$ then $V \otimes W$ is a representation of $\mathfrak{g}$ via $d \circ \Delta$, where $\Delta(x) = (x, x) \in \mathfrak{g} \oplus \mathfrak{g}$. More explicitly, if (V, ρ) and (W, σ) are $\mathfrak{g}$-representations then $V \otimes W$ is a $\mathfrak{g}$-representation with

$$x(v \otimes w) = \rho(x) \otimes w + v \otimes \sigma(w), \quad \forall v \in V, w \in W. \tag{2.3.1}$$

2.3.1 Tensoring with one-dimensional representations

If (V, ρ) is any $\mathfrak{g}$-representation, then by Example I.9, we have an isomorphism *of vector spaces* $V \otimes \mathsf{k}_\lambda \to V$ given by the map $v \otimes \lambda \mapsto \lambda.v$. Via this map, one can think of the $\mathfrak{g}$-representation $V \otimes \mathsf{k}_\lambda$ as the same vector space V but now equipped with a new action ρ_λ of $\mathfrak{g}$, where $\rho_\lambda(x) = \rho(x) + \lambda(x).I_V$ (where we write I_V for the identity map.) Note that, in particular, if $\lambda, \mu \in D(\mathfrak{g})^0$ then this shows that $\mathsf{k}_\lambda \otimes \mathsf{k}_\mu \cong \mathsf{k}_{\lambda+\mu}$.

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

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

Examining the formula (2.3.1) for the action on a tensor product of representations given above we see 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. In the case $V = W$, σ becomes an involution on $V \otimes V$ commuting with the $\mathfrak{g}$-action. In other words, S_2, the symmetric group on two letters acts on $V \otimes V$ and the isotypic decomposition of $V \otimes V$ under this action, (equivalently the $(+1)$- and (-1)-eigenspaces of σ) shows that $V \otimes V$ is the direct sum of the subrepresentations of symmetric tensors and skew-symmetric tensors, that is $V \otimes V = \mathrm{Sym}^2(V) \oplus \mathrm{Alt}^2(V)$ where

$$\mathrm{Sym}^2(V) = \mathrm{span}_{\mathsf{k}} \left\{ \frac{1}{2}(v_1 \otimes v_2 + v_2 \otimes v_1) : v_1, v_2 \in V \right\},$$
$$\mathrm{Alt}^2(V) = \mathrm{span}_{\mathsf{k}} \left\{ \frac{1}{2}(v_1 \otimes v_2 - v_2 \otimes v_1) : v_1, v_2 \in V \right\}.$$

Let V and W be k-vector spaces. There is a natural linear map $\theta\colon V^* \otimes W \to \mathrm{Hom}(V, W)$, given by $\theta(f \otimes w) = f.w$ where $(f.w)(v) = f(v).w$ for all $v \in V, f \in V^*$ and $w \in W$. This map is injective, and its image is precisely the space of finite-rank linear maps[3] from V to W. In particular, if $\dim(V) < \infty$ then we have $\mathrm{End}(V) \cong V^* \otimes V$. Similarly, there is a natural map $m\colon V^* \otimes W^* \to (V \otimes W)^*$, where

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

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

Remark 2.3.2. It is worth noticing that this gives a coordinate-free way of defining the trace, and also some explanation for why one needs some finiteness condition in order for the trace to be defined.

Remark 2.3.3. If $\mathfrak{g}$ is a Lie algebra and V and W are $\mathfrak{g}$-representations, then 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 I.13 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 I.2 yield maps of $\mathfrak{g}$-representations.

The following example will be very useful in a number of places later in the course.

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

Example 2.3.4. If $\mathfrak{g}$ is a Lie algebra and (V,ρ) is a $\mathfrak{g}$-representation, then ρ induces a natural bilinear map $a_\rho\colon \mathfrak{g}\times V\to V$, namely $(x,v)\mapsto \rho(x)(v)$. By the universal property of tensor products this yields a linear map $\tilde{a}_\rho\colon \mathfrak{g}\otimes V\to V$. We claim this map is a homomorphism of $\mathfrak{g}$ representations (where $\mathfrak{g}$ is viewed as the adjoint representation). To see this, first notice that the bilinear map $a_\rho\colon \mathfrak{g}\times V\to V$ is equal to $a_V\circ(\rho\times 1_V)$ where $a_V\colon \mathfrak{gl}_V\times V\to V$ is the natural action of $\mathfrak{gl}_V$ on V, $(\phi,v)\mapsto\phi(v)$. Thus it suffices to check the claim for $\mathfrak{gl}_V$ and its vector representation V. Let $\tilde{a}_V\colon \mathfrak{gl}_V\otimes V\to V$ be the linear map induced by a_V. Then if $x,y\in\mathfrak{gl}_V$ and $v\in V$ we have $x(\tilde{a}_V(y\otimes v)) = x(y(v))$, while

$$\tilde{a}_V(x(y\otimes v)) = [x,y]\otimes v + y\otimes x(v) = (xy-yx)(v)+yx(v) = xy(v) = x(\tilde{a}_V(y\otimes v)) \tag{2.3.2}$$

hence as y was arbitrary we have $\tilde{a}_V\circ x = x\circ\tilde{a}_V$ for all $x\in\mathfrak{gl}_V$ so that $\tilde{a}_V\in\mathrm{Hom}_{\mathfrak{gl}_V}(\mathfrak{gl}_V\otimes V, V)$ as required.

Remark 2.3.5. In fact one can also deduce that $\tilde{a}_\rho$ is a homomorphism of $\mathfrak{g}$-representations by observing that under the identification $\mathfrak{gl}_V\cong V^*\otimes V$ the map $\tilde{a}_\rho$ corresponds to the linear map $\iota_{13}\colon V^*\otimes V\otimes V\to V$, that is, the linear map which is the identity on the 2nd tensor factor and is the contraction map ι on the first and third factors:[4] so that $\iota(f_1\otimes v_2\otimes v_3) = f_1(v_3).v_2$ where $v_2,v_3\in V$, $f_1\in V^*$. Since ι is a homomorphism of $\mathfrak{g}$-representations, it follows $\tilde{a}_\rho$ is also.

[4]Here, as usual, we are also identifying $\mathsf{k}\otimes V$ with V equipped with the scalar multiplication map, $s\colon \mathsf{k}\times V\to V$, that is $s(\lambda,v)=\lambda.v$. (Recall that a tensor product $V\otimes W$ is a vector space and a bilinear map $V\times W\to V\otimes W$.)

Chapter 3

Classifying Lie algebras

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

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

3.1 Classification by composition factors

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

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

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

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

For Lie algebras, we can attempt something similar. In fact, it turns out that, at least in characteristic zero, we obtain a far more complete answer about the structure of an arbitrary finite-dimensional Lie algebra than one could hope to obtain in a Part C course on finite group theory. One aspect of this finer information will reveal a sharp distinction between $\mathfrak{gl}_1$ and the non-abelian Lie algebras which have no proper ideals, which is one reason for the following definition:

Definition 3.1.1. A non-zero Lie algebra $\mathfrak{g}$ is said to be *almost simple*[1] if it has no proper ideals. If $\mathfrak{g}$ is almost simple and $\dim(\mathfrak{g}) > 1$ then we say that $\mathfrak{g}$ is *simple*. Equivalently, an almost simple Lie algebra is simple if it is non-abelian. Thus the only almost simple Lie algebra which is not simple is $\mathfrak{gl}_1$.

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

The *Jordan-Hölder theorem* for Lie algebras shows that the almost simple Lie algebras that occur as composition factors of a composition series are in fact independent of the choice of composition series. As we will see later, Cartan's criteria will give stronger results (though only in when working over fields of characteristic zero), so this result is only included for completeness. (Only the statements are examinable.)

Definition 3.1.2. A *composition series* for a finite dimensional Lie algbera $\mathfrak{g}$ is a chain

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

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

It is straight-forward to check by induction on $\dim(\mathfrak{g})$ that any finite-dimensional Lie algebra has a composition series: given a Lie algebra $\mathfrak{g}$, pick a proper ideal $\mathfrak{a}$ whose dimension is maximal among proper ideals. Then by the maximality of $\mathfrak{a}$ the quotient $\mathfrak{g}/\mathfrak{a}$ has no non-trivial proper ideals and hence is almost simple. But $\dim(\mathfrak{a}) < \dim(\mathfrak{g})$, hence by induction $\mathfrak{a}$ has a composition series, say $(\mathfrak{a}_i)_{i=1}^d$, where $\mathfrak{a} = \mathfrak{a}_1 > \mathfrak{a}_1 > \dots > \mathfrak{a}_d = 0$. But then if we set $\mathfrak{g} = \mathfrak{g}_0$ and $\mathfrak{g}_i = \mathfrak{a}_i$ for $i \geq 1$, it follows easily that $(\mathfrak{g}_i)$ is a composition series for $\mathfrak{g}$.

Definition 3.1.3. If $\mathfrak{s}$ is an almost simple Lie algebra and $\mathscr{C} = (\mathfrak{g}_i)_{i=0}^r$ is a composition series for a finite-dimensional Lie algebra $\mathfrak{g}$, define the multiplicity of $\mathfrak{s}$ in $\mathscr{C}$ to be

$$[\mathfrak{s}, \mathscr{C}] = \#\Big\{i \in \{1, \dots r\} : \mathfrak{s} \cong \mathfrak{g}_{i-1}/\mathfrak{g}_i\Big\}$$

The following Lemma shows that a composition series for a Lie algebra $\mathfrak{g}$ induces one on any ideal or quotient of $\mathfrak{g}$.

Proposition 3.1.4. *Suppose that $\mathfrak{g}$ has a composition series $\mathscr{C} = (\mathfrak{g} = \mathfrak{g}_0 \rhd \mathfrak{g}_1 \rhd \dots \rhd \mathfrak{g}_n = 0)$ and let $\mathfrak{a}$ be an ideal of $\mathfrak{g}$. Then $\mathscr{C}$ induces a composition series $\mathscr{C}_\mathfrak{a}$ for $\mathfrak{a}$, and a composition series $\mathscr{C}_{\mathfrak{g}/\mathfrak{a}g}$ for the quotient $\mathfrak{g}/\mathfrak{a}$. Moreover, for any almost simple Lie algebra $\mathfrak{s}$ we have $[\mathfrak{s} : \mathscr{C}] = [\mathfrak{s} : \mathscr{C}_\mathfrak{a}] + [\mathfrak{s} : \mathscr{C}_{\mathfrak{g}/\mathfrak{a}}]$.*

Proof. Consider the sequence $(\mathfrak{a}_i)_{i=0}^n$ where $\mathfrak{a}_i = \mathfrak{a} \cap \mathfrak{g}_i$. Note that its terms, while nested, need not be strictly decreasing. Since $\mathfrak{a}$ is an ideal in $\mathfrak{g}$, the intersection $\mathfrak{a}_i = \mathfrak{a} \cap \mathfrak{g}_i$ is an ideal in $\mathfrak{g}_i$ and, by the second isomorphism theorem, its image under the quotient map $p_{i+1} \colon \mathfrak{g}_i \to \mathfrak{g}_i/\mathfrak{g}_{i+1}$ is

$$\mathfrak{a}_i/\mathfrak{a}_{i+1} \cong (\mathfrak{a}_i + \mathfrak{g}_{i+1})/\mathfrak{g}_{i+1} = p_{i+1}(\mathfrak{a}_i) \tag{3.1.1}$$

Similarly, we may consider the sequence $(q(\mathfrak{g}_i))_{i=0}^n$, where $q \colon \mathfrak{g} \to \mathfrak{g}/\mathfrak{a}$ is the quotient map, then $q(\mathfrak{g}_{i+1})$ is an ideal in $q(\mathfrak{g}_i)$, and by the second isomorphism theorem $q(\mathfrak{g}_i) \cong \mathfrak{g}_i/\mathfrak{g}_i \cap \mathfrak{a} = \mathfrak{g}_i/\mathfrak{a}_i$. Under this identification, $q(\mathfrak{g}_{i+1})$ is isomorphic to $(\mathfrak{g}_{i+1} + \mathfrak{a}_i)/\mathfrak{a}_i$, and hence

$$q(\mathfrak{g}_i)/q(\mathfrak{g}_{i+1}) \cong (\mathfrak{g}_i/\mathfrak{a}_i)\big/(\mathfrak{g}_{i+1} + \mathfrak{a}_i)/\mathfrak{a}_i \cong \mathfrak{g}_i/(\mathfrak{g}_{i+1} + \mathfrak{a}_i). \tag{3.1.2}$$

Thus since $\mathfrak{g}_i/\mathfrak{g}_{i+1}$ is almost simple, and $\mathfrak{g}_{i+1} \subseteq \mathfrak{g}_{i+1} + \mathfrak{a}_i \subseteq \mathfrak{g}_i$, we must either have $\mathfrak{g}_{i+1} + \mathfrak{a}_i = \mathfrak{g}_i$, in which case Equations (3.1.1) and (3.1.2) show that $q(\mathfrak{g}_i) = q(\mathfrak{g}_{i+1})$ and $\mathfrak{a}_i/\mathfrak{a}_{i+1} \cong \mathfrak{g}_i/\mathfrak{g}_{i+1}$, or $\mathfrak{g}_{i+1} + \mathfrak{a}_i = \mathfrak{g}_{i+1}$, in which case $\mathfrak{a}_i = \mathfrak{a}_{i+1}$ and $q(\mathfrak{g}_i)/q(\mathfrak{g}_{i+1}) \cong \mathfrak{g}_i/\mathfrak{g}_{i+1}$.

Thus removing repetitions from the sequences $(\mathfrak{a}_i)$ and $(q(\mathfrak{g}_i))$ yields composition series $\mathscr{C}_\mathfrak{a}$ and $\mathscr{C}_{\mathfrak{g}/\mathfrak{a}}$ for $\mathfrak{a}$ and $\mathfrak{g}/\mathfrak{a}$ respectively, and the composition factors of $\mathscr{C}$ correspond to a composition factor of precisely one of $\mathscr{C}_\mathfrak{a}$ or $\mathscr{C}_{\mathfrak{g}/\mathfrak{a}}$. □

The previous proposition gives one natural way to prove the *Jordan-Hölder theorem*:

Corollary 3.1.5. *(Jordan-Hölder theorem for Lie algberas): Let $\mathfrak{g}$ be any (finite-dimensional) Lie algebra $\mathfrak{g}$ and let $\mathscr{C}$ be a composition series for $\mathfrak{g}$. If $\mathfrak{s}$ is an almost simple Lie algebra, then the multiplicity with which $\mathfrak{s}$ occurs as a composition factor of $\mathscr{C}$ is independent of $\mathscr{C}$ and hence equals $[\mathfrak{s} : V$.*

Proof. We use induction on the minimal length $n(\mathfrak{g})$ of a composition series for $\mathfrak{g}$. If $n(\mathfrak{g}) = 1$ then V is irreducible and $(\mathfrak{g} > 0)$ is its unique composition series. If $n = n(\mathfrak{g}) > 1$ then take a composition series $\mathscr{M} = (\mathfrak{m}_i)_{i=0}^n$ of $\mathfrak{g}$ with length n and set $\mathfrak{h} = \mathfrak{m}_1$. Since $(\mathfrak{m}_{i+1})_{i=0}^{n-1}$ is a composition series for $\mathfrak{h}$, we have $n(\mathfrak{h}) \leq n - 1$. Now if $\mathscr{C} = (\mathfrak{g}_i)_{i=0}^d$ is any composition series for $\mathfrak{g}$, by Proposition II.10, it induces composition series $\mathscr{C}_{\mathfrak{h}}$ and $\mathscr{C}_{\mathfrak{g}/\mathfrak{h}}$ of $\mathfrak{h}$ and $\mathfrak{g}/\mathfrak{h}$ respectively. Thus if $\mathfrak{s}$ is almost simple, by the final sentence of Proposition 3.1.4 we have

$$[\mathfrak{s} : \mathscr{C}] = [\mathfrak{s} : \mathscr{C}_{\mathfrak{h}}] + [S : \mathscr{C}_{\mathfrak{g}/\mathfrak{h}}] = [\mathfrak{s} : \mathfrak{h}] + [\mathfrak{s} : \mathfrak{g}/\mathfrak{h}]$$

where the second equality follows by induction since $n(\mathfrak{g}/\mathfrak{h}) = 1$ and $n(\mathfrak{h}) \leq n-1$. Thus $[\mathfrak{s} : \mathscr{C}] = [\mathfrak{s} : \mathfrak{g}]$ is independent of $\mathscr{C}$. $\square$

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

3.2 Exact sequences of Lie algebras

Definition 3.2.1. We say that the sequence of Lie algebras and Lie algebra homomorphisms

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

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

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

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

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

3.2.1 Split extensions

Definition 3.2.2. An extension of Lie algebras

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

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

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

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

This motivates the following definition:

Definition 3.2.3. Suppose that $\mathfrak{g}, \mathfrak{h}$ are Lie algebras, and we have a homomorphism $\phi\colon \mathfrak{g} \to \mathrm{Der}_{\mathsf{k}}(\mathfrak{h})$, the Lie algebra of derivations[2] on $\mathfrak{h}$. Then it is straight-forward to check that we can form a new Lie algebra $\mathfrak{h} \rtimes \mathfrak{g}$, the *semi-direct product*[3] of $\mathfrak{g}$ and $\mathfrak{h}$ by ϕ which as a vector space is just $\mathfrak{g} \oplus \mathfrak{h}$, and where the Lie bracket is given by:

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

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

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

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

Remark 3.2.4. In general, there may be many ways to split an exact sequence of Lie algebras (see Problem Sheet 1).

Example 3.2.5. Let $\mathfrak{s}_2$ be the 2-dimensional Lie algebra with basis $\{x, y\}$ and Lie bracket given by $[x, y] = y$. Then $\mathsf{k}.y$ is an ideal in $\mathfrak{s}_2$, and $\mathfrak{s}_2/\mathsf{k}.y$ is 1-dimensional, hence we have a short exact sequence:

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

where $i(\lambda) = \lambda.y$ and $q(ax + by) = a$, for all $a, b, \lambda \in \mathsf{k}$. Now the map $s(\lambda) = \lambda.x$ is a Lie algebra homomorphism, hence the extension is split. Note that $\mathrm{Der}_{\mathsf{k}}(\mathfrak{a}) = \mathfrak{gl}_{\mathfrak{a}}$ for an Abelian Lie algebra $\mathfrak{a}$, and so $\mathrm{Der}_{\mathsf{k}}(\mathfrak{gl}_1) = \mathfrak{gl}_{\mathfrak{gl}_1} = \mathfrak{gl}_1$, and the map from $\mathfrak{gl}_1$ to $\mathrm{Der}_{\mathsf{k}}(\mathfrak{gl}_1)$ describing $\mathfrak{s}_2$ as a semi-direct product corresponds to the identity map under this identification.

Remark 3.2.6. A short exact sequence of the form

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

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

Remark 3.2.7. There is a close analogy with the notion of a short exact sequence of groups which you have seen in a previous course: here one has a sequence

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

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

3.2.2 Central extensions

Another type of extension which plays an important role in our study of Lie algebras is a *central extension*. In this case, the Lie algebra $\mathfrak{g}_1$ in the sequence of Definition 3.2.1 is assumed to be central in $\mathfrak{g}$, that is $\mathfrak{g}_1 \subseteq \mathfrak{z}(\mathfrak{g})$, and hence in particular $\mathfrak{g}_1$ is Abelian. Picking a *linear* splitting $s\colon \mathfrak{g}_2 \to \mathfrak{g}$, we can write any $x, y \in \mathfrak{g}$ uniquely in the form $x = i(x_1) + s(x_2), y = i(x_2) + s(y_2)$, respectively. Thus, as $i(\mathfrak{g}_1)$ is central, the Lie bracket on $\mathfrak{g}$ is given by

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

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

Definition 3.2.8. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra and let $\mathfrak{z}$ be a vector space. A *2-cocycle* on $\mathfrak{g}$ taking values in the vector space $\mathfrak{z}$ is a map $\alpha\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{z}$ satisfying the conditions:

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

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

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

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

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

Example 3.2.9. It is straight-forward to understand central extensions of a Lie algebra $\mathfrak{g}$ by $\mathfrak{gl}_1$ in low dimensions. If $\mathfrak{g}$ is 1-dimensional, then the fact that α is alternating forces it to vanish, and hence the only central extension of $\mathfrak{gl}_1$ by $\mathfrak{gl}_1$ is the abelian Lie algebra $\mathfrak{gl}_1^{\oplus 2}$.

If $\dim(\mathfrak{g}) = 2$, then if $\mathfrak{g}$ is abelian then condition (*ii*) is automatically satisfied, and there is a unique non-zero alternating bilinear form up to isomorphism: if $\mathfrak{g}$ has basis $\{x, y\}$, then $\alpha(x, y) = 1 = -\alpha(y, x)$, defines a central extension of $\mathfrak{g}$. This is the smallest non-abelian nilpotent Lie algebra, known as the *Heisenberg Lie algebra*. It can be realised as the strictly upper triangular matrices $\mathfrak{n}_3 \subseteq \mathfrak{gl}_3(\mathsf{k})$.

Remark 3.2.10. Split and central extensions are in a loose sense complementary to each other: An extension of $\mathfrak{g}_2$ by $\mathfrak{g}_1$ which is both central and split is just the direct sum $\mathfrak{g}_1 \oplus \mathfrak{g}_2$, where $\mathfrak{g}_1 \cong \mathfrak{gl}_1^{\oplus k}$ and $k = \dim_{\mathsf{k}}(\mathfrak{g}_1)$.

Chapter 4

Gluing $\mathfrak{gl}_1$: Solvable and nilpotent Lie algebras

Conventions: *From this point onwards in these notes, we will assume that all Lie algebras and all representations are finite-dimensional over the field* k, *unless the contrary is explicitly stated, and from §4.3 onwards,* k *will be algebraically closed of characteristic zero.*

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

4.1 Solvable Lie algebras

Definition 4.1.1. A Lie algebra $\mathfrak{g}$ is *solvable* if its only composition factor is $\mathfrak{gl}_1(\mathsf{k})$. This is equivalent to the condition that $\mathfrak{g}$ has a nested sequence of subalgebras

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

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

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

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

where ι_{k+1} is the inclusion map and q_k the quotient map. Thus $\mathfrak{g}_{k-1}$ is an extension of $\mathfrak{gl}_1$ by $\mathfrak{g}_k$. By Remark 3.2.7, this short exact sequence must split, and so $\mathfrak{g}_k$ is a semidirect product of $\mathfrak{g}_{k-1}$ by $\mathfrak{gl}_1(\mathsf{k})$, and so solvable Lie algebras are precisely the Lie algebras one obtains from the zero Lie algebra by taking iterated semidirect products with $\mathfrak{gl}_1(\mathsf{k})$.

Example 4.1.2. Example 3.2.5 shows that $\mathfrak{s}_2$, the 2-dimensional non-abelian Lie algebra, is solvable.

Definition 4.1.3. We can rephrase the condition that a Lie algebra $\mathfrak{g}$ is solvable in terms of a decreasing sequence of ideals in $\mathfrak{g}$: The *derived subalgebra*[2] $D(\mathfrak{g})$ of $\mathfrak{g}$ is defined to be $[\mathfrak{g}, \mathfrak{g}]$ (an ideal in $\mathfrak{g}$ since $\mathfrak{g}$ is). Inductively we define $D^k(\mathfrak{g}) = D(D^{k-1}(\mathfrak{g})) = [D^{k-1}(\mathfrak{g}), D^{k-1}(\mathfrak{g})]$ for each $k \geq 1$. The sequence of ideals $(D^k(\mathfrak{g}))_{k\geq 0}$ is called the *derived series* of $\mathfrak{g}$. Note that, since $\mathfrak{g}$ is an ideal in $\mathfrak{g}$, it follows by induction on k that $D^k(\mathfrak{g}) = [D^{k-1}(\mathfrak{g}), D^{k-1}(\mathfrak{g})]$ is an ideal in $\mathfrak{g}$.

Lemma 4.1.4. *Let $\mathfrak{g}$ be a Lie algebra. Then $D(\mathfrak{g})$ is the smallest ideal in $\mathfrak{g}$ such that $\mathfrak{g}/D(\mathfrak{g})$is abelian. In particular, $\mathfrak{g}$ is solvable precisely when the derived series $(D^k(\mathfrak{g}))_{k\geq 1}$ satisfies $D^k(\mathfrak{g}) = 0$ for sufficiently large k.*

[1] Hence starting with nothing...

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

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

Next note that we have a short exact sequence

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

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

□

Remark 4.1.5. Because the terms of the derived series are ideals in $\mathfrak{g}$, it follows that if $\mathfrak{g}$ is solvable, then there is a filtration of $\mathfrak{g}$ whose terms are ideals in $\mathfrak{g}$ not just subalgebras each of which is an ideal in the previous term of the filtration. In particular, if $\mathfrak{g}$ is solvable, it follows $\mathfrak{g}$ has an non-trivial abelian ideal, since the last non-zero term of the derived series must be such an ideal. This also shows that solvable Lie algebras can be viewed as those Lie algebras which can be obtained from the trivial Lie algebra by successive extensions by abelian Lie algebras.

Remark 4.1.6. If $\mathfrak{g}$ is an arbitrary finite dimensional Lie algebra, then the derived series $(D^k(\mathfrak{g}))_{k\geq 0}$ is a decreasing sequence of ideals in $\mathfrak{g}$, hence it must stabilize, *i.e.* there is a unique integer N such that $D^N(\mathfrak{g}) = D^m(\mathfrak{g})$ for all $m \geq N$, while $D^{N-1}(\mathfrak{g}) \supsetneq D^N(\mathfrak{g})$. We will denote this final term of the derived series by $D^\infty(\mathfrak{g})$. It is a *perfect* Lie algebra, that is, $D^\infty(\mathfrak{g}) = D(D^\infty(\mathfrak{g}))$.

Recall the notion of a *flag* $\mathscr{F} = (V = F_0 > \ldots F_d = \{0\})$ from Definition 2.2.3

Lemma 4.1.7. *Let V be a finite dimensional vector space and let $\mathscr{F} = (V = F_0 \supset F_1 \supset \ldots \supset F_d = \{0\})$ be a flag in V, and set $F_n = \{0\}$ if $n \geq d$. Let, for any $r \in \mathbb{Z}_{\geq 0}$,*

$$\mathfrak{b}^r_{\mathscr{F}} = \{x \in \mathfrak{gl}_V : x(F_i) \subseteq F_{i+r}, \forall\, i, 0 \leq i \leq d\}.$$

(i) If $k, l \geq 0$, then $[\mathfrak{b}^k_{\mathscr{F}}, \mathfrak{b}^l_{\mathscr{F}}] \subseteq \mathfrak{b}^{k+l}_{\mathscr{F}}$.

(ii) If $\mathscr{F}$ is a complete flag, and $\mathfrak{b}_{\mathscr{F}} = \mathfrak{b}^0_{\mathscr{F}}$, then $D(\mathfrak{b}_{\mathscr{F}}) \subseteq \mathfrak{b}^1_{\mathscr{F}}$ and moreover $\mathfrak{b}_{\mathscr{F}}$ is solvable.

Proof. First note that $\mathfrak{b}^r_{\mathscr{F}} \subseteq \mathfrak{b}^s_{\mathscr{F}}$ if $r \geq s$, and that if $x \in \mathfrak{b}^k_{\mathscr{F}}$, $y \in \mathfrak{b}^l_{\mathscr{F}}$, then clearly $x \circ y$ and $y \circ x$ lie in $\mathfrak{b}^{k+l}_{\mathscr{F}}$. It follows that the $\mathfrak{b}^r_{\mathscr{F}}$ form a descending sequence of associative subalgebras of $\mathrm{End}_{\mathsf{k}}(V)$, where the $\mathfrak{b}^r_{\mathscr{F}}$ for $r > 0$ are two-sided ideals in $\mathfrak{b}_{\mathscr{F}} = \mathfrak{b}^0_{\mathscr{F}}$, since $\mathfrak{b}^r_{\mathscr{F}}.\mathfrak{b}^s_{\mathscr{F}} \subseteq \mathfrak{b}^{r+s}_{\mathscr{F}}$. But this immediately implies (i), that is, $[\mathfrak{b}^r_{\mathscr{F}}, \mathfrak{b}^s_{\mathscr{F}}] \subseteq \mathfrak{b}^{r+s}_{\mathscr{F}}$.

If $\mathscr{F}$ is a complete flag, and $x, y \in \mathfrak{b}_{\mathscr{F}}$, then for any $i, (1 \leq i \leq d)$, x and y induce linear maps on F_i/F_{i+1}, and, since $\mathscr{F}$ is complete, $\dim(F_i/F_{i+1}) = 1$, so that $\mathfrak{gl}_{F_i/F_{i+1}}$ is abelian, and thus the map induced by $[x, y]$ on F_i/F_{i+1} is zero. But this exactly says that $[x, y] \in \mathfrak{b}^1_{\mathscr{F}}$, and hence $D(\mathfrak{b}_{\mathscr{F}}) \subseteq \mathfrak{b}^1_{\mathscr{F}}$. But then $D^k(\mathfrak{b}_{\mathscr{F}}) \subseteq D^{k-1}(\mathfrak{b}^1_{\mathscr{F}})$, and using (i) and induction $D^{k-1}(\mathfrak{b}^1_{\mathscr{F}}) \subseteq \mathfrak{b}^{2^{k-1}}_{\mathscr{F}}$, which is $\{0\}$ if $2^{k-1} \geq d = \dim(V)$, and hence $\mathfrak{b}_{\mathscr{F}}$ is solvable.

□

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 $\mathfrak{b}_{\mathscr{F}}$ for some complete flag $\mathscr{F}$. We next note some basic properties of solvable Lie algebras. We establish them using the characterization of solvability in terms of the derived series, but it is also straight-forward to show them using composition series.[3]

Lemma 4.1.8. *Let $\mathfrak{g}$ be a Lie algebra, $\phi\colon \mathfrak{g} \to \mathfrak{h}$ a homomorphism of Lie algebras.*

(i) 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.

(ii) If $\mathfrak{g}$ is solvable then so are all subalgebras of $\mathfrak{g}$.

[3] If one uses composition series to prove this Lemma, note that if $\mathscr{C} = (\mathfrak{g}_k)_{k=0}^d$ is a composition series with all composition factors $\mathfrak{g}_k/\mathfrak{g}_{k+1}$ isomorphic to $\mathfrak{gl}_1$. Now if $\mathfrak{a} \subseteq \mathfrak{g}$ is a subalgebra, then $\mathfrak{a} \cap \mathfrak{g}_k \supseteq \mathfrak{a} \cap \mathfrak{g}_{k+1}$ and for dimension reasons, this containment is either an equality, or the quotient is isomorphic to $\mathfrak{gl}_1$, so that $\mathfrak{a}$ is solvable. Note that if $\mathfrak{g}$ is an arbitrary Lie algebra and $\mathfrak{h}$ is a subalgebra, the composition factors of $\mathfrak{h}$ do not have to be composition factors of $\mathfrak{g}$.

(iii) 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. It is immediate from the definitions that $\phi(D(\mathfrak{g})) = D(\phi(\mathfrak{g}))$, and hence by induction we have $\phi(D^k(\mathfrak{g})) = D^k(\phi(\mathfrak{g}))$, from which (i) follows immediately. If $\mathfrak{h} \subseteq \mathfrak{g}$ is a subalgebra, then clearly $D(\mathfrak{h}) \subseteq D(\mathfrak{g})$, and again by induction we see that $D^k(\mathfrak{h}) \subseteq D^k(\mathfrak{g})$, which certainly implies that if $\mathfrak{g}$ is solvable, then $\mathfrak{h}$ is solvable.

Finally, for (iii), the second sentence follows from the first applied to the quotient map $q\colon \mathfrak{g} \to \mathfrak{g}/I$. To establish the first sentence, note that if im(ϕ) is solvable, then for some N we have $D^N\mathrm{im}(\phi) = \{0\}$, so that by part (i) we have $D^N(\mathfrak{g}) \subset \ker(\phi)$, hence applying (ii) we see that $D^N(\mathfrak{g})$ is solvable since $\ker(\phi)$ is. But since the derived series of $D^N(\mathfrak{g})$ is a tail of that of $\mathfrak{g}$, it follows $\mathfrak{g}$ is solvable. □

4.2 Nilpotent Lie algebras

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

Definition 4.2.1. A Lie algebra $\mathfrak{g}$ is said to be *nilpotent* if it can be obtained from 0, the trivial Lie algebra, by iterated central extensions. If $\mathfrak{g}$ can be obtained by precisely k iterated extentions, we say $\mathfrak{g}$ is *k-step nilpotent*. Thus, for example, a Lie algebra is 1-step nilpotent if and only if it is abelian.

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

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

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

Definition 4.2.2. For $\mathfrak{g}$ a Lie algebra, let $C^0(\mathfrak{g}) = \mathfrak{g}$, and $C^i(\mathfrak{g}) = [\mathfrak{g}, C^{i-1}(\mathfrak{g})]$ for $i \geq 1$. This sequence of ideals of $\mathfrak{g}$ is called the *lower central series* of $\mathfrak{g}$.

Remark 4.2.3. Notice that $C^1(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}]$ is the *derived subalgebra*[4] of $\mathfrak{g}$ and, as we have seen, this is also denoted[5] $D(\mathfrak{g})$ and sometimes $\mathfrak{g}'$.

Proposition 4.2.4. *Suppose that $\mathfrak{g}$ is nilpotent and $(\mathfrak{q}_i)_{i=0}^k$ the sequence of ideals associated to a realization of $\mathfrak{g}$ as an iterated sequence of central extensions. Then*

(i) For each $i \geq 0$ we have $C^i(\mathfrak{g}) \subseteq \mathfrak{q}_i$ and hence $C^k(\mathfrak{g}) = 0$.

(ii) Conversely, if $\mathfrak{g}$ is such that, for some $N \geq 0$ we have $C^N(\mathfrak{g}) = 0$, then $\mathfrak{g}$ is at most N-step nilpotent.

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

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

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

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

Lemma 4.2.5. *Let $\mathfrak{g}$ be a Lie algebra. Then*

(i) If $\mathfrak{g}$ is nilpotent, any subalgebra or quotient of $\mathfrak{g}$ is nilpotent.

(ii) If $\mathfrak{g}$ is nilpotent, then the centre $\mathfrak{z}(\mathfrak{g})$ is non-zero if $\mathfrak{g}$ is. Moreover, $\mathfrak{g}/\mathfrak{z}(\mathfrak{g})$ is nilpotent if and only if $\mathfrak{g}$ is.

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

[5] 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.

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

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

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

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

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

Part (ii) is trivial since a non-trivial central extension always has a non-trivial centre. (Alternatively, as $C^k(\mathfrak{g})/C^{k+1}(\mathfrak{g})$ is central in $\mathfrak{g}/C^{k+1}(\mathfrak{g})$, clearly the last non-zero term of the lower central series is central.) □

Remark 4.2.6. Notice that if $\mathfrak{a}$ is an arbitrary ideal in $\mathfrak{g}$, and $\mathfrak{a}$ and $\mathfrak{g}/\mathfrak{a}$ are nilpotent it does *not* follow that $\mathfrak{g}$ is nilpotent. Indeed recall from Example 3.2.5 the non-abelian 2-dimensional Lie algebra $\mathfrak{s}_2$, with basis $\{x, y\}$ where $[x, y] = y$. Then $\mathsf{k}.y$ is a 1-dimensional ideal in $\mathfrak{s}_2$ but it is not central. Indeed $\mathfrak{z}(\mathfrak{s}_2) = 0$ so $\mathfrak{s}_2$ is not nilpotent, even though the ideal $\mathsf{k}.y$ and the quotient $\mathfrak{s}_2/\mathsf{k}.y$ are (since they are both abelian). Note that this shows that $\mathfrak{s}_2$ cannot be written as a central extension of $\mathfrak{gl}_1$ by itself.

Remark 4.2.7. The characterisation of the property of nilpotence in terms of the lower central series is similar to the characterisation of solvable Lie algebras in terms of the derived series. This is one reason it is commonly used. There is, however, another nature nested sequence of ideals which can be used to characterize nilpotence: If $\mathfrak{g}$ is any Lie algebra, set $Z^0(\mathfrak{g}) = \mathfrak{g}$, and, assuming $Z^k(\mathfrak{g})$ is defined, let $q_k \colon \mathfrak{g} \to \mathfrak{g}/Z^k(\mathfrak{g})$ be the quotient map, and set $Z^{k+1}(\mathfrak{g}) = q_k^{-1}(\mathfrak{z}(\mathfrak{g}_k))$. This process yields an *increasing* sequence of ideals of $\mathfrak{g}$ known as the *upper central series*. If it exhausts $\mathfrak{g}$, that is, if for some $n \geq 0$ we have $Z^k(\mathfrak{g}) = \mathfrak{g}$ for all k large enough, the $\mathfrak{g}$ is nilpotent. If $\mathfrak{g}$ is not nilpotent, the upper central series will stabilize at a maximal nilpotent ideal of $\mathfrak{g}$.

In terms of the adjoint representation, the centre of a Lie algebra $\mathfrak{g}$ can be viewed as $\ker(\mathrm{ad})$, the kernel of the adjoint action, but it can also be viewed as the invariants in $\mathfrak{g}$, that is

$$\mathfrak{g}^{\mathfrak{g}} = \{z \in \mathfrak{g} : ad(x)(z) = 0, \forall\, x \in \mathfrak{g}\}.$$

Using either the upper or lower central series, it is easy to see that the only composition factor of $(\mathfrak{g}, \mathrm{ad})$ is the trivial representation.

We now wish to show that the notion of a flag in a vector space gives us a large supply of nilpotent Lie algebras. In the next Lemma we use the notation of Lemma 4.1.7.

Lemma 4.2.8. *Suppose that $\mathscr{F}$ is a (not necessarily complete) flag in a finite-dimensional vector space V. Then the Lie algebra $\mathfrak{n}_{\mathscr{F}} = \mathfrak{b}^1_{\mathscr{F}} \subseteq \mathfrak{gl}_V$ is nilpotent.*

Proof. By (i) of Lemma 4.1.7, $[\mathfrak{n}_{\mathscr{F}}, \mathfrak{n}_{\mathscr{F}}] \subseteq \mathfrak{b}^2_{\mathscr{F}}$, and by induction $[\mathfrak{n}_{\mathscr{F}}, C^k(\mathfrak{n}_{\mathscr{F}})] \subseteq \mathfrak{b}^{k+1}_{\mathscr{F}}$, so that $C^k(\mathfrak{g}) \subseteq \mathfrak{b}^k_{\mathscr{F}}$, and hence $C^k(\mathfrak{n}) = 0$ if $\mathfrak{b}^k_{\mathscr{F}} = 0$, which is true whenever $k \geq d = \dim(V)$. □

Example 4.2.9. When $\mathscr{F}$ is a complete flag, so that $\dim(V) = d$, if we pick a basis $\{e_1, e_2, \ldots, e_d\}$ of V such that $F_k = \langle\{e_{k+1}, e_2, \ldots, e_d\}\rangle_{\mathsf{k}}$, then the matrix A representing an element $x \in \mathfrak{n}_{\mathscr{F}}$ with respect to this basis is strictly upper triangular, that is, $a_{ij} = 0$ for all $i \geq j$. But then if $\mathfrak{n}_d \subseteq \mathfrak{gl}_d$ denotes the space of strictly upper-triangular matrices, it is easy to see that $\dim(\mathfrak{n}_d) = \binom{d}{2}$. When $d = 2$ we just get the 1-dimensional Lie algebra $\mathfrak{gl}_1$, thus the first nontrivial case is when $n = 3$ and in this case $\mathfrak{n}_3$ is the 3-dimensional 2-step nilpotent Lie algebra we constructed previously as a central extension.

On the other hand, if $\mathfrak{b}_d \subseteq \mathfrak{gl}_d$ denotes the upper-triangular matrices, *i.e.* $\mathfrak{b}_d$ corresponds to the subalgebra $\mathfrak{b}_{\mathscr{F}}$ of $\mathfrak{gl}_V$, and we set $\mathfrak{t}_d$ to be the set of diagonal matrices in $\mathfrak{b}_n$, then it is straight-forward to show by considering the subalgebra $\mathfrak{t}_n$ of diagonal matrices in $\mathfrak{b}_n$ that $[\mathfrak{t}_d, \mathfrak{n}_d] = \mathfrak{n}_d$, so that, as $\mathfrak{b}_d = \mathfrak{t}_d \oplus \mathfrak{n}_d$, it follows that $\mathfrak{b}_d$ is not nilpotent.

Remark 4.2.10. Note that in Example 4.1.7, unlike in Lemma 4.2.8, it is essential that $\mathscr{F}$ is a complete flag. If $\mathscr{F}$ is not a complete flag the corresponding subalgebra $\mathfrak{b}_{\mathscr{F}}$ will not be solvable (since, for example, if $\dim(F_i/F_{i+1}) > 1$, then there is a surjective homomorphism $\mathfrak{b}_{\mathscr{F}} \to \mathfrak{gl}_{F_i/F_{i+1}}$, which is not solvable.)

Remark 4.2.11. Note that the subalgebra $\mathfrak{t}_d \subset \mathfrak{gl}_d$ of diagonal matrices is abelian, and hence nilpotent, but the only nilpotent endomorphism of k^d that lies in $\mathfrak{t}_d$ is 0. Thus a nilpotent linear Lie algebra need not consist of nilpotent endomorphisms. It turns out that, in some sense, the example of $\mathfrak{t}_d$ is the *only* way in which a nilpotent Lie algebra $\mathfrak{n} \subseteq \mathfrak{gl}_d$ can fail to consist of nilpotent endomorphisms. We will make this precise in 4.3.2.

4.2.1 Nilpotent representations

Definition 4.2.12. Let $\mathfrak{g}$ be a Lie algebra and (V,ρ) a representation of $\mathfrak{g}$. We say that (V,ρ) is *nilpotent* if, for all $x \in \mathfrak{g}$, the endomorphism $\rho(x) \in \mathfrak{gl}_V$ is a nilpotent linear map (that is, for some $n \geq 1$, $\rho(x)^n = 0$).

Lemma 4.2.13. *Let A be an associative algebra, and suppose $a, b \in A$ are nilpotent i.e. for some $n > 0$, we have $a^n = b^n = 0$. Then if a and b commute, $a+b$ is also nilpotent.*

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

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

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

Lemma 4.2.14. *Suppose $\mathfrak{g}$ is a Lie algebra and (V,ρ) and (W,σ) are representation of $\mathfrak{g}$.*

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

ii) If V is nilpotent, then any subrepresentation and any quotient representation of V is also nilpotent.

Proof. By definition, the action of x on $V \otimes W$ is given by $\rho(x) \otimes 1_W + 1_V \otimes \sigma(x)$. Since the two terms in this sum commute, the claim follows from Lemma 4.2.13 (taking $A = \mathrm{End}(V \otimes W)$.)

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

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

For part *ii)* if U is a subrepresentation of V then, as in Lemma 2.2.2, $\rho(\mathfrak{g}) \subseteq \mathfrak{b}_U = \{x \in \mathfrak{gl}_V : x(U) \subseteq U\}$. But $\mathfrak{b}_U$ is an associative subalgebra of $\mathrm{End}(V)$ and the maps i^* and p_* from $\mathfrak{b}_U$ to $\mathrm{End}(U)$ and $\mathrm{End}(V/U)$ are also maps of associative algebras, hence if $\rho(x) \in \mathfrak{b}_U$ is nilpotent, its images in $\mathrm{End}(U)$ and $\mathrm{End}(V/U)$ are also. ☐

The next proposition is the key result in this section. For the proof we will need the notion of the *normalizer* $N_{\mathfrak{g}}(\mathfrak{a})$ of a subalgebra $\mathfrak{a}$ of a Lie algebra $\mathfrak{g}$ given in Definition 1.2.6. We have

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

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

Proposition 4.2.15. *Let $\mathfrak{g}$ be a Lie algebra, and let (V,ρ) be a nilpotent representation of $\mathfrak{g}$.*

i) The invariant subspace

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

is non-zero.

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

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

Now let $\mathscr{S} = \{\mathfrak{b} \subsetneq \mathfrak{g} : \mathfrak{b} \text{ is a proper subalgebra of } \mathfrak{g}\}$ denote the set of proper subalgebras of $\mathfrak{g}$, and pick $\mathfrak{a} \in \mathscr{S}$. Now by Lemma 4.2.14, $\mathfrak{a} \subsetneq \mathfrak{g} \subseteq \mathfrak{gl}_V = V^* \otimes V$ are all nilpotent representations of $\mathfrak{a}$, since the restriction of (V,ρ) to $\mathfrak{a}$ is. But then, by the same Lemma, $\mathfrak{g}/\mathfrak{a}$ is also a nilpotent representation, and since $\dim(\mathfrak{a}) < \dim(\mathfrak{g})$, it follows by induction that the $\mathfrak{a}$-invariants $(\mathfrak{g}/\mathfrak{a})^{\mathfrak{a}}$ form a non-zero subrepresentation. Let $x \in \mathfrak{g}$ be such that $0 \neq x + \mathfrak{a} \in (\mathfrak{g}/\mathfrak{a})^{\mathfrak{a}}$. Then $\mathrm{ad}(a)(x) \in \mathfrak{a}$ for all $a \in \mathfrak{a}$, or equivalently, since $\mathrm{ad}(a)(x) = -\mathrm{ad}(x)(a)$, for all $a \in \mathfrak{a}$, we have $\mathrm{ad}(x)(a) \in \mathfrak{a}$, that is, $x \in N_{\mathfrak{g}}(\mathfrak{a})$. Thus the normalizer of $\mathfrak{a}$ is a subalgebra of $\mathfrak{g}$ which is strictly larger than $\mathfrak{a}$.

Thus if we take $\mathfrak{a} \in \mathscr{S}$ of maximal dimension, we must have $N_{\mathfrak{g}}(\mathfrak{a}) = \mathfrak{g}$, that is $\mathfrak{a}$ is an ideal in $\mathfrak{g}$. But then if $z \in \mathfrak{g}\backslash\mathfrak{a}$, it is easy to see that $\mathsf{k}.z \oplus \mathfrak{a}$ is a subalgebra[6] of $\mathfrak{g}$, hence again by maximality, we must have $\mathfrak{g} = \mathsf{k}.z \oplus \mathfrak{a}$. By induction, we know that $V^{\mathfrak{a}} = \{v \in V : a(v) = 0, \forall\, a \in \mathfrak{a}\}$ is a nonzero subspace of V. We claim that z preserves $V^{\mathfrak{a}}$. Indeed

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

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

For *ii*), let $\mathscr{C} = (V = F_m > F_{m-1} > \ldots > F_1 > F_0 = \{0\})$ be a composition series for V. It suffices to show that each of the composition factors are trivial. But if $1 \leq k \leq m$, then F_k is a subrepresentation of V and hence it is nilpotent. Similarly $Q_k = F_k/F_{k+1}$, as a quotient of F_k must be nilpotent. But then by part (1), its invariants $Q_k^{\mathfrak{g}}$ are a non-zero subrepresentation of Q_k, and since Q_k is simple it follows that Q_k is the trivial representation as required.

□

Corollary 4.2.16. *(Engel's theorem.) A Lie algebra $\mathfrak{g}$ is nilpotent if and only if $ad(x)$ is nilpotent for every $x \in \mathfrak{g}$, i.e the adjoint representation is nilpotent.*

Proof. If $\mathfrak{g}$ is nilpotent, then since by definition $\mathrm{ad}(x)(C^i(\mathfrak{g})) \subseteq C^{i+1}(\mathfrak{g})$, we see that $\mathrm{ad}(x)^k = 0$ for all $x \in \mathfrak{g}$ if $\mathfrak{g}$ is k-step nilpotent. Now suppose that $\mathrm{ad}(x)$ is nilpotent for all $x \in \mathfrak{g}$. Then $(\mathfrak{g}, \mathrm{ad})$ is a nilpotent representation, and hence by part *ii*) of Proposition 4.2.15, we see that $\mathrm{ad}(\mathfrak{g})$ is nilpotent. But since $\mathrm{ad}(\mathfrak{g}) \cong \mathfrak{g}/\mathfrak{z}(\mathfrak{g})$ it follows that $\mathfrak{g}$ is nilpotent as required. □

4.3 Representations of solvable Lie algebras

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

4.3.1 Lie's theorem

Our first goal is the following theorem:

Theorem 4.3.1. *(Lie's theorem) Let $\mathfrak{g}$ be a solvable Lie algebra and V is a $\mathfrak{g}$-representation. Then there is a homomorphism $\lambda\colon \mathfrak{g} \to \mathfrak{gl}_1(\mathsf{k})$ and a nonzero vector $v \in V$ such that $x(v) = \lambda(x).v$ for all $x \in \mathfrak{g}$. Equivalently, any finite-dimensional irreducible representation of a solvable Lie algebra is one-dimensional.*

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

The crucial observation that is needed to prove Lie's theorem is given in the following Lemma:

Lemma 4.3.2. *(Lie's Lemma) Let $\mathfrak{g}$ be a Lie algebra, let $I \subset \mathfrak{g}$ be an ideal, and let V be a finite dimensional $\mathfrak{g}$-representation. Suppose that $\lambda\colon I \to \mathfrak{gl}_1(\mathsf{k})$ is a homomorphism of Lie algebras for which the subspace $V_{\lambda,I} = \{v \in V : h(v) = \lambda(h).v, \forall\, h \in I\}$ is nonzero. Then λ vanishes on $[\mathfrak{g}, I] \subset I$, and $V_{\lambda,I}$ is a $\mathfrak{g}$-subrepresentation of V.*

Proof. Fix $x \in \mathfrak{g}$ and $v \in V_{\lambda,I}\backslash\{0\}$. For each $m \in \mathbb{N}$, let $W_m = \langle\{v, x(v), \ldots, x^m(v)\}\rangle_{\mathsf{k}}$. The W_m form an increasing sequence of subspaces of V with $\dim(W_{m+1}/W_m) \leq 1$, from which it is easy to see that there is some d with $W_{-1} := \{0\} < W_0 < W_1 < \ldots < W_{d-1} = W_d$, where $x(W_d) \subseteq W_d$.

Claim:

$$h(x^m(v)) \in \lambda(h).x^m(v) + W_{m-1}, \quad \forall\, h \in I, m \geq 0$$

Proof of claim: This obvious for $m = 0$ and if $m \geq 1$, since $[h,x] \in I$, by induction we have

$$\begin{aligned} hx^m(v) &= xhx^{m-1}(v) + (hx - xh)x^{m-1}(v) \\ &\in x\left(\lambda(h).x^{m-1}(v) + W_{m-2}\right) + \lambda([h,x])x^{m-1}(v) + W_{m-2} \\ &\subseteq \lambda(h)x^m(v) + W_{m-1}, \end{aligned} \tag{4.3.1}$$

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

since $x(W_{m-2}) \subseteq W_{m-1}$ and $[h,x] \in I$.

It follows that W_d is stable under the action of I and x, and that if $h \in I$, then the matrix of the action of h on W_d with respect to the basis $\{v, x(v), \dots, x^{d-1}(v)\}$ is upper-triangular with each diagonal entry equal to $\lambda(h)$. In particular, $\mathrm{tr}(h_{|W_d}) = \lambda(h).\dim(W_d) = \lambda(h).d$. But then

$$0 = \mathrm{tr}([h,x]) = \lambda([h,x]).d,$$

and since $d > 0$ and $\mathrm{char}(\mathsf{k}) = 0$, it follows that $\lambda([h,x]) = 0$. But now considering (4.3.1) with $m = 1$ we see that $\lambda([h,x]) = 0$ implies that $hx(v) = \lambda(h).x(v)$, so that $x(v) \in V_{\lambda,I}$ if $v \in V_{\lambda,I}$ as required.

□

Completion of the proof of Lie's Theorem:

Use induction on $\dim(\mathfrak{g})$. Since $\mathfrak{g}$ is solvable, $D(\mathfrak{g})$ is a proper ideal in $\mathfrak{g}$, and by induction the theorem holds for $D(\mathfrak{g})$. But then let $\lambda\colon D(\mathfrak{g}) \to \mathsf{k}$ be a homomorphism of Lie algebras such that

$$W = \{v \in V : h(v) = \lambda(h).v, \ \forall h \in D(\mathfrak{g})\} \neq \{0\}.$$

But then by Lemma 4.3.2, W is a $\mathfrak{g}$-subrepresentation, and if we let $\rho\colon \mathfrak{g} \to \mathfrak{gl}_W$, then $\rho(\mathfrak{g}) \subseteq \mathfrak{gl}_W$ has $\lambda(D(\mathfrak{g})).I_W = \rho(D(\mathfrak{g})) = D(\rho(\mathfrak{g}))$. But $D(\mathfrak{gl}_W) = \mathfrak{sl}(W)$, and since $\mathrm{char}(\mathsf{k}) = 0$, $\mathsf{k}.I_W \cap \mathfrak{sl}(W) = \{0\}$, so that as $D(\mathfrak{g}_1) \subseteq \mathfrak{sl}(W) \cap \mathsf{k}.I_W = \{0\}$. It follows that the action of $\mathfrak{g}$ on W factors through $\mathfrak{g}/D(\mathfrak{g})$, which is abelian, and the result is then clear, since commuting linear maps on a non-zero vector space always have a common eigenvector.

Remark 4.3.3. The proof of Lemma 4.3.2 relies on a trick which permeates the course, namely that one can often compute a trace in two different ways to obtain important information. One way will be by observing that one is computing the trace of a commutator, which is therefore zero. The other will, in one fashion or another, follow from consideration of the generalised eigenspaces of the linear map in question.

Corollary 4.3.4. *Let $\mathfrak{g}$ be a solvable Lie algebra and let (V,ρ) be a $\mathfrak{g}$-representation. Then there is a complete flag $\mathscr{F} = (V = F_0 \supset F_1 \supset \dots \supset F_d = \{0\})$ where each F_i is a $\mathfrak{g}$-subrepresentation, so that $\rho(\mathfrak{g}) \subseteq \mathfrak{b}_{\mathscr{F}}$. In particular, if $\mathfrak{g}$ is solvable, then it has a composition series each of whose terms is an ideal in all of $\mathfrak{g}$.*

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

Definition 4.3.5. Recall from Example 2.1.5 that the isomorphism classes of one-dimensional representations of a Lie algebra $\mathfrak{g}$ are given by the elements of $(\mathfrak{g}/D(\mathfrak{g}))^* = D(\mathfrak{g})^0$: a homomorphism $\lambda\colon \mathfrak{g} \to \mathfrak{gl}_1$ is just a linear map $\lambda\colon \mathfrak{g} \to \mathsf{k}$ which vanishes on $D(\mathfrak{g})$. Recall that we write k_λ for the representation (k, λ). We will refer to an element of $D(\mathfrak{g})^0$ (equivalently, an isomorphism class of 1-dimensional $\mathfrak{g}$-representations) as a *weight* of $\mathfrak{g}$. In the case where $\mathfrak{g}$ is solvable, Lie's theorem shows that the weights are exactly the isomorphism classes of irreducible $\mathfrak{g}$-representations.

Remark 4.3.6. Note that, since it is a k-vector space, $D(\mathfrak{g})^0$ is an abelian group. This abelian group structure can also be seen from the point of view of one-dimensional representations: since the tensor product of 1-dimensional vector spaces is 1-dimensional, the tensor product restricts to an operation on 1-dimensional vector spaces. This gives the set of isomorphism classes of one-dimensional representations the structure of an abelian group: the operation is commutative because the map $\sigma\colon L_1 \otimes L_2 \to L_2 \otimes L_1$ given by $\sigma(v_1 \otimes v_2) = v_2 \otimes v_1$ is an isomorphism of $\mathfrak{g}$-representations (for any two $\mathfrak{g}$-representations L_1,L_2) and if L is any one-dimensional representation then $L \otimes L^* \cong \mathsf{k}_0$ via the evaluation (or contraction) map induced by the natural bilinear pairing $L \times L^* \to \mathsf{k}_0$.

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

4.3.2 Representations of nilpotent Lie algebras

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

Definition 4.3.7. Let $\mathfrak{g}$ be a Lie algebra and let $\mathscr{S}$ be a set of irreducible representation of $\mathfrak{g}$. Let

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

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

Proposition 4.3.8. *Let $\mathfrak{g}$ be a Lie algebra and (V, ρ) a representation of $\mathfrak{g}$. If $\mathscr{S}$ is a set of irreducible $\mathfrak{g}$-representation then $\operatorname{Rep}_{\mathscr{S}}(\mathfrak{g}, V)$ has a unique element $V_{\mathscr{S}}$ which is maximal with respect to containment, that is $V_{\mathscr{S}} \in \operatorname{Rep}_{\mathscr{S}}(\mathfrak{g}, V)$ and if $U \in \operatorname{Rep}_{\mathscr{S}}(\mathfrak{g}, V)$ then $U \leq V_{\mathscr{S}}$.*

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

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

Definition 4.3.9. Recall that the isomorphism classes of 1-dimensional representations of $\mathfrak{g}$ can be identified with $D(\mathfrak{g})^0 \subseteq \mathfrak{g}$, and given $\lambda \in D(\mathfrak{g})^0$, we write k_λ for the 1-dimensional representation (k, λ). Given a $\mathfrak{g}$-representation (V, ρ), we will write V_λ and $\operatorname{Rep}_\lambda(\mathfrak{g}, V)$ instead of V_{k_λ} and $\operatorname{Rep}_{\mathsf{k}_\lambda}(\mathfrak{g}, V)$. When $\lambda \in D(\mathfrak{g})^0$ we will refer to V_λ as the *λ-weight space* of V.[7] If V is a finite-dimensional representation of a Lie algebra $\mathfrak{g}$, let

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

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

If $\varphi\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ and (V, ρ) is a representation of $\mathfrak{g}_2$, then $(V, \varphi^*(\rho))$ is a representation of $\mathfrak{g}_1$, where $\varphi^*(\rho) = \rho \circ \varphi$. Since $\varphi(D(\mathfrak{g}_1)) \subseteq D(\mathfrak{g}_2)$, the transpose $\varphi^\top : \mathfrak{g}_2^* \to \mathfrak{g}_1^*$ restricts to give a map $\varphi^\top : D(\mathfrak{g}_2)^0 \to D(\mathfrak{g}_1)^0$, and $\Psi_{\varphi^*(V)} = \varphi^\top(\Psi_V)$. Now if $x \in \mathfrak{g}$ and $i_x\colon \mathfrak{gl}_1 \to \mathfrak{g}$ is the homomorphism $i_x(t) = t.x$ ($\forall\, t \in \mathsf{k} = \mathfrak{gl}_1$), and $\lambda \in D(\mathfrak{g})^0$ then $i_x^\top(\lambda) = \lambda(x)$. The weights of the $\mathfrak{gl}_1$-representation $\rho \circ i_x$ are just the eigenvalues of $\rho(x)$, as in Example 2.1.4, it follows that the eigenvalues of $\rho(x)$ are $\{\lambda(x) : \lambda \in \Psi_V\}$, and the μ-generalised eigenspace of $\rho(x)$ is $\bigoplus_{\lambda \in \Psi_V : \lambda(x) = \mu} V_\lambda$.

Lemma 4.3.10. *Suppose that $\mathfrak{g}$ is a Lie algebra and $\lambda, \mu \in D(\mathfrak{g})^0$ are weights of $\mathfrak{g}$. If V and W are $\mathfrak{g}$-representations then*

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

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

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

$$x(e_k \otimes f_l) = x(e_k) \otimes f_l + e_k \otimes x(f_l) \in (\lambda + \mu)(e_k \otimes f_l) + H_{k+l+1} \tag{4.3.2}$$

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

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

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

[7] This is somewhat nonstandard – the λ-isotypical subrepresentation of V_λ is usually called the *λ-generalised weight space* of V, with its socle, V_λ^s being the λ-weight space.

Proposition 4.3.11. *Let $\mathfrak{g}$ be a nilpotent Lie algebra, $\mathfrak{h} \subseteq \mathfrak{g}$ be a subalgebra of $\mathfrak{g}$, and (V, ρ) a representation of $\mathfrak{g}$. Then if $\mu \in (\mathfrak{h}/D(\mathfrak{h}))^* \cong D(\mathfrak{h})^0/\mathfrak{h}^0 \subseteq \mathfrak{g}^*/\mathfrak{h}^0 \cong \mathfrak{h}^*$ is a weight of $\mathfrak{h}$, and V_μ is the μ-isotypic subrepresentation of $\mathrm{Res}^{\mathfrak{g}}_{\mathfrak{h}}(V)$, the restriction of V to $\mathfrak{h}$, then V_μ is a $\mathfrak{g}$-subrepresentation of V. In particular, taking $\mathfrak{h} = \mathsf{k}.x$ for $x \in \mathfrak{g}\backslash\{0\}$, any generalised eigenspace $V_{\mu,x}$ of $\rho(x)$ is a $\mathfrak{g}$-subrepresentation.*

Proof. Since $\mathfrak{g}$ is nilpotent, we have $\mathfrak{g} = \mathfrak{g}_0$ as an $\mathfrak{h}$-representation. But then by Lemma 4.3.10, we have $\mathfrak{g} \otimes V_\mu = \mathfrak{g}_0 \otimes V_\mu \subseteq (\mathfrak{g} \otimes V)_\mu$, and since the map $\tilde{a}_\rho \colon \mathfrak{g} \otimes V \to V$ given by $\tilde{a}_\rho(x \otimes v) = \rho(x)(v)$ is a homomorphism of $\mathfrak{h}$-representations by Example 2.3.4, it follows that $\tilde{a}_\rho(\mathfrak{g} \otimes V_\mu) = \rho(\mathfrak{g})(V_\mu) \subseteq V_\mu$, that is, V_μ is a $\mathfrak{g}$-subrepresentation as required. □

Definition 4.3.12. Let $\mathfrak{g}$ be a nilpotent Lie algebra and let (V, ρ) be a representation of $\mathfrak{g}$. Say $x \in \mathfrak{g}$ is V*-generic* if, for all $\lambda, \mu \in \Psi_V$ we have $\lambda(x) = \mu(x)$ if and only if $\lambda = \mu$.

If $D_V = \{\lambda - \mu : \lambda, \mu \in \Psi_V\}\backslash\{0\}$, then x is V-generic if and only if $x \notin \bigcup_{\nu \in D_V} \ker(\nu)$. If k is infinite,[8] it is an elementary exercise to show that a nonzero k-vector space cannot be written as the union of finitely many hyperplanes, hence V-generic elements of $\mathfrak{g}$ exist for any finite-dimensional $\mathfrak{g}$-representation V.

Theorem 4.3.13. *Let $\mathfrak{g}$ be a nilpotent Lie algebra and (V, ρ) a finite-dimensional representation of $\mathfrak{g}$. For each $\lambda \in (\mathfrak{g}/D\mathfrak{g})^*$, let*

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

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

Proof. Since $\mathfrak{g}$ is nilpotent, it is solvable, hence for any $\mathfrak{g}$-representation (U, σ) its composition factors all lie in Ψ_U and, as in Definition 4.3.9, if $x \in \mathfrak{g}$ then $\sigma(x)$ has spectrum $\{\lambda(x) : \lambda \in \Psi_U\}$. In particular, taking $U = V_\lambda$ we see that $\rho(x)_{|V_\lambda}$ has $\lambda(x)$ as its sole eigenvalue, that is, $V_\lambda \subseteq V_{\lambda(x),x}$. It follows that $V_\lambda \subseteq W_\lambda$.

Now if $x \in \mathfrak{g}$, we have $V = \bigoplus_{\lambda(x):\lambda \in \Psi_V} V_{\lambda(x),x}$ and by Proposition 4.3.11 each $V_{\lambda(x),x}$ is a $\mathfrak{g}$-subrepresentation of V, hence taking $U = V_{\lambda(x),x}$ we see that if k_ν is a composition factor, then $\nu(x) = \lambda(x)$. It follows that if we take x_0 to be V-generic, the generalised eigenspace $V_{\lambda(x_0),x_0}$ has λ as its unique composition factor, so that $V_{\lambda(x_0),x_0} \subseteq V_\lambda$. Hence $V_{\lambda(x_0),x_0} = V_\lambda = W_\lambda$ and $V = \bigoplus_{\lambda \in \Psi_V} V_\lambda$.

□

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

Chapter 5

Cartan subalgebras, trace forms and Cartan's criteria

5.1 Nilpotent Lie algebras as measurements: 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 from Definition 1.2.6 that if $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$ then the normalizer $N_{\mathfrak{g}}(\mathfrak{h})$ of $\mathfrak{h}$ in $\mathfrak{g}$ is $N_{\mathfrak{g}}(\mathfrak{h}) = \{x \in \mathfrak{g} : [x, h] \in \mathfrak{h}, \forall\, h \in \mathfrak{h}\}$, the largest subalgebra of $\mathfrak{g}$ in which $\mathfrak{h}$ is an ideal.

Definition 5.1.1. A subalgebra $\mathfrak{h}$ is said to be a *Cartan subalgebra* if it is *i*) nilpotent and *ii*) self-normalizing, that is, $N_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$. We will call a pair $(\mathfrak{g}, \mathfrak{h})$ a *Cartan pair* if $\mathfrak{h}$ is a Cartan subalgebra of a Lie algebra $\mathfrak{g}$.

Lemma 5.1.2. *If* $(\mathfrak{g}, \mathfrak{h})$ *is a Cartan pair and* $\mathfrak{g} = \bigoplus_{\alpha\in\Phi_0} \mathfrak{g}_\alpha$, *is the decomposition of* $\mathfrak{g}$ *into* $\mathfrak{h}$*-isotypical subrepresentations where* $\Phi_0 \subseteq \mathfrak{h}^*$ *is the finite subset of* $\mathfrak{h}$*-weights for which* $\mathfrak{g}_\alpha \neq \{0\}$, *then* $0 \in \Phi_0$ *and* $\mathfrak{g}_0 = \mathfrak{h}$.

Proof. Consider $\mathfrak{g}/\mathfrak{h}$ as an $\mathfrak{h}$-representation: if $x \in N_{\mathfrak{g}}(\mathfrak{h})$, then $x + \mathfrak{h} \in (\mathfrak{g}/\mathfrak{h})^{\mathfrak{h}}$ of $(\mathfrak{g}/\mathfrak{h})$, the invariants of $\mathfrak{g}/\mathfrak{h}$. Thus if $\mathfrak{h} = N_{\mathfrak{g}}(\mathfrak{h})$ is self-normalising, then $(\mathfrak{g}/\mathfrak{h})^{\mathfrak{h}} = (\mathfrak{g}/\mathfrak{h})^s_0 = 0$, hence $(\mathfrak{g}/\mathfrak{h})_0 = \{0\}$. On the other hand, since $\mathfrak{h}$ is nilpotent, $\mathfrak{h} \subseteq \mathfrak{g}_0$, □

It is not clear from this definition whether a Lie algebra necessarily contains a Cartan subalgebra. We will for the moment assume this result, in order to show how they provide a powerful tool to study the structure of an arbitrary finite-dimensional Lie algebra.

5.1.1 The Cartan Decomposition

Definition 5.1.3. Let $(\mathfrak{g}, \mathfrak{h})$ be a Cartan pair and let (V, ρ) be a $\mathfrak{g}$-representation. Then, by Theorem 4.3.13, restricting V and $\mathfrak{g}$, the adjoint representation, to $\mathfrak{h}$, we may write them as a direct sum of their isotypic subrepresentations:

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha\in\Phi} \mathfrak{g}_\alpha, \quad \text{and } V = \bigoplus_{\lambda\in\Psi_V} V_\lambda \tag{5.1.1}$$

is the *weight-space* decomposition of V. The elements of Φ are called the *roots* of $\mathfrak{g}$ and the elements of Ψ_V are called the *weights* of V.

For $\mathfrak{g}$, as noted in the Lemma above, $\mathfrak{g}_0 = \mathfrak{h}$, so we let $\Phi = \{\alpha \in D(\mathfrak{g})^0\backslash\{0\} : \mathfrak{g}_\alpha \neq \{0\}\}$. If we also set $\Psi_V = \{\lambda \in D(\mathfrak{g})^0 : V_\lambda \neq \{0\}$, then the Cartan decomposition of $\mathfrak{g}$ is

Remark 5.1.4. When k is algebraically closed with $\mathrm{char}(\mathsf{k}) = 0$ it is known that the set of all Cartan subalgebras of a k-Lie algebra $\mathfrak{g}$ form a single orbit under the group of *inner automorphisms* of $\mathfrak{g}$. This shows that the Cartan Decomposition of $\mathfrak{g}$ is unique up to automorphisms of $\mathfrak{g}$.

The following simple Lemma will, along with Cartan's criterion for semisimplicity, be the key to the classification of semisimple Lie algebras. It shows that the $\mathfrak{h}$-weights of $\mathfrak{g}$ give a kind of grading of $\mathfrak{g}$ and its representations.

Lemma 5.1.5. *For any $\alpha \in \{0\} \cup \Phi$ and $\lambda \in \Psi$, we have $\rho(\mathfrak{g}_\alpha)(V_\lambda) \subseteq V_{\lambda+\alpha}$. In particular, if $(V, \rho) = (\mathfrak{g}, ad)$ then for any $\alpha, \beta \in \{0\} \cup \Phi$ we have $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$.*

Proof. This follows using the techniques of the proof of Proposition 4.3.11: the action map ρ induces a homomorphism of $\mathfrak{h}$-representations $\tilde{\rho}\colon \mathfrak{g} \otimes V \to V$. Now by part (i) of Lemma 4.3.10 we see $\mathfrak{g}_\alpha \otimes V_\lambda \subseteq (\mathfrak{g} \otimes V)_{\lambda+\alpha}$, and then part (ii) shows that its image under $\tilde{\rho}$ lies in $V_{\lambda+\alpha}$. But by definition $\tilde{\rho}(\mathfrak{g}_\alpha \otimes V_\lambda) = \rho(\mathfrak{g}_\alpha)(V_\lambda)$, and so the proof is complete. □

5.1.2 Existence of Cartan subalgebras

Note that, if $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$, then the Theorem 4.3.13 shows that there is an $x_0 \in \mathfrak{h}$ such that $\mathfrak{h} = \mathfrak{g}_{0,x_0}$. This motivates the following definition:

Definition 5.1.6. If $x \in \mathfrak{g}$, let $\mathfrak{g}_{0,x}$ be the generalized 0-eigenspace of $\mathrm{ad}(x)$. Note that we always have $x \in \mathfrak{g}_{0,x}$ so that $\dim(\mathfrak{g}_{0,x}) \geq 1$. We say that $x \in \mathfrak{g}$ is *regular* if $\mathfrak{g}_{0,x}$ is of minimal dimension.

Proposition 5.1.7. *Let $\mathfrak{g}$ be a Lie algebra over a field* k.

i) If $x \in \mathfrak{g}$ is any element, then $\mathfrak{g}_{0,x}$ is a self-normalizing subalgebra of $\mathfrak{g}$.

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

Proof. Part *i)* is straight-forward: It follows immediately from Lemma 4.3.11 applied to the adjoint representation that $\mathfrak{h} = \mathfrak{g}_{0,x}$ is a subalgebra of $\mathfrak{g}$. To see that $\mathfrak{h}$ is a self-normalizing in $\mathfrak{g}$. Indeed if $z \in N_\mathfrak{g}(\mathfrak{h})$ then $[x, z] \in \mathfrak{h}$ (since certainly $x \in \mathfrak{h}$), so that for some n we have $\mathrm{ad}(x)^n([x, z]) = 0$, and hence $\mathrm{ad}(x)^{n+1}(z) = 0$ and $z \in \mathfrak{h}$ as required.

To establish part *ii)*, assume that x is regular, and let $\mathfrak{h} = \mathfrak{g}_{0,x}$. To see that $\mathfrak{h}$ is nilpotent, by Engel's theorem it suffices to show that, for each $y \in \mathfrak{h}$, the map $\mathrm{ad}(y)$ is nilpotent as an endomorphism of $\mathfrak{h}$. To see this, we consider the characteristic polynomials of $\mathrm{ad}(y)$ on $\mathfrak{g}$, $\mathfrak{h}$ and $\mathfrak{g}/\mathfrak{h}$: Since $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, the characteristic polynomial $\chi^y(t) \in \mathsf{k}[t]$ of $\mathrm{ad}(y)$ 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 $\{h_1, h_2, \dots, h_r\}$ a basis of $\mathfrak{h}$ (so that $\dim(\mathfrak{h}) = r$). Then if we write $y = \sum_{i=1}^r y_i h_i$, the coefficients $\{c_k(y)\}_{k=0}^n$ of $\chi^y(t)$ are polynomial functions of the coordinates $\{y_i : 1 \leq i \leq r\}$. Similarly we have

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

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

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

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

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

□

In the course of the proof of the above Proposition we used the fact that the coefficients of the characteristic polynomial were polynomial functions of the coordinates of $y \in \mathfrak{h}$ with respect to a basis of $\mathfrak{h}$. This was crucial because, whereas the product of two arbitrary nonzero functions may well be zero, the product of two nonzero *polynomials* (over a field) is never zero. Lemma I.3 in Appendix I establishes a slightly more general statemetn which applied to $V = \mathfrak{g}$, $A = \mathfrak{h}$ and $\varphi = \mathrm{ad}$ gives a proof of this polynomial property.[1]

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

5.2 Trace forms and Cartan's criterion for Solvability

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[2] is given in §I.3 in Appendix 1 of these notes.

5.2.1 Invariant bilinear forms

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

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

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

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

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

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

Definition 5.2.1. We say that a bilinear form $B \in \mathrm{Bil}(\mathfrak{g})$ is *invariant* if it is an invariant vector for the action of $\mathfrak{g}$ on $\mathrm{Bil}(\mathfrak{g}) \cong (\mathfrak{g} \otimes \mathfrak{g})^*$, that is, if $B(\mathrm{ad}(x)(y),z) = B(y, -ad(x)(z)) = 0$ for all $x,y,z \in \mathfrak{g}$. This is often written as

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

Remark 5.2.2. If (V,ρ) is a $\mathfrak{g}$-representation and $B \in \mathrm{Bil}(V)$ is a bilinear form, then it defines a linear map $\theta\colon V \to V^*$ where $\theta(v)(w) = B(v,w)$ $(\forall\, v,w \in V)$. If B is invariant, that is $B \in \mathrm{Bil}(V)^{\mathfrak{g}}$, then we have $\theta(\rho(x)(v))(w) = B(\rho(x)(v),w) = B(v,-\rho(x)(w)) = \rho^*(x)(\theta(v))(w)$ for all $v,w \in V$, hence $\theta(\rho(x)(v)) = \rho(x)^*(\theta(v))$, that is, $\theta \in \mathrm{Hom}(V,V^*)^{\mathfrak{g}} = \mathrm{Hom}_{\mathfrak{g}}(V,V^*)$ is a homomorphism of $\mathfrak{g}$-representations.

If θ is an isomorphism, we say B is *nondegenerate* and in that case, for any linear map $\alpha \in \mathrm{End}(V)$ we may define $\alpha^* = \theta^{-1} \circ \alpha^\top \circ \theta \in \mathrm{End}(V)$, the *adjoint* of α with respect to B. If V is a $\mathfrak{g}$-representation and B is nondegenerate, then the condition that B is invariant can be expressed as $\rho(x)^* = -\rho(x)$ for all $x \in \mathfrak{g}$, where $\rho\colon \mathfrak{g} \to \mathfrak{gl}_V$ is the action map, that is, $\rho(\mathfrak{g})$ consists of skew-adjoint endomorphisms of with respect to the bilinear form B.

Definition 5.2.3. 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 5.2.4. *Let V be a k-vector space. The trace form $b_V\colon \mathfrak{gl}_V \otimes \mathfrak{gl}_V \to \mathsf{k}$ given by*

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

is a nondegenerate invariant symmetric bilinear form on $\mathfrak{gl}_V$.

[2] Part A Algebra focused more on positive definite and Hermitian forms, but there is a perfectly good theory of symmetric bilinear forms over an arbitrary field k. When k is algebraically closed, the theory is also straight-forward!

Proof. Let $\vartheta\colon V\otimes V^*\to \mathfrak{gl}_V$ be give by $\vartheta(w\otimes f)=f.w\in\mathfrak{gl}_V$, where $(f.w)(v)=f(v).w$ for all $v\in V$. For V finite-dimensional, this map is an isomorphism and if $\iota\colon V\otimes V^*\to \mathsf{k}$ is the natural "contraction" map induced by the evaluation map $(v,f)\mapsto f(v)$, then $\mathrm{tr}(\vartheta(v\otimes f))=\mathrm{tr}(f.v)=f(v)=\iota(v\otimes f)$ (see Lemma I.13 for details).

Moreover, the isomorphism $\vartheta\otimes\vartheta\colon V\otimes V^*\otimes V\otimes V^*\to\mathfrak{gl}_V\otimes\mathfrak{gl}_V$ identifies the composition map $(a,b)\to a\circ b$ with the contraction map on the 2nd and 3rd factors. Indeed for any $v_1,v_2\in V, f_1,f_2\in V^*$ we have

$$\vartheta(v_1\otimes f_1)\circ\vartheta(v_2\otimes f_2)=(f_1.v_1)\circ(f_2.v_2)=f_1(v_2).(f_2.v_1)=\vartheta(\iota_{23}(v_1\otimes f_1\otimes v_2\otimes f_2))$$

Thus we see that $(a,b)\mapsto\mathrm{tr}(ab)$ corresponds under $\vartheta\otimes\vartheta$ to the map $\iota_{14}\otimes\iota_{32}\colon V\otimes V^*\otimes V\otimes V^*\to\mathsf{k}$, where we write ι_{kl} for the contraction map acting on the k-th and l-th tensor factors if the kth is V and the lth is V^*. The composition $(a,b)\mapsto ab$ gives the contraction ι_{23} and then taking trace corresponds to the contraction ι_{14}. Taking $\mathrm{tr}(ba)$ gives the same value since $(a,b)\mapsto\mathrm{tr}(ba)$ simply contracts the factors in the opposite order, so that $\mathrm{tr}(ab)=\mathrm{tr}(ba)$ and thus b_V is a symmetric bilinear form. To show it is invariant, since $b_V=\iota_{14}\otimes\iota_{23}$ it suffices to check that ι is invariant. But this is clear, since $x(v\otimes f)=x(v)\otimes f-v\otimes(f\circ x)$, thus $\iota(x(v\otimes f))=f(x(v))-(f\circ x)(v)=0$, while $x(\iota(v\otimes f))=x(f(v))=0$, since $\mathfrak{gl}_V$ acts by 0 on k, the trivial representation. We leave it as an exercise to check the nondegeneracy of b_V. □

Remark 5.2.5. One can also of course check the invariance property by a direct calculation: 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 $\mathrm{tr}((ac).b)$.

Definition 5.2.6. If $\mathfrak{g}$ is a Lie algebra, and let (V,ρ) be a representation of $\mathfrak{g}$. we may define a bilinear form $t_V\colon\mathfrak{g}\times\mathfrak{g}\to\mathsf{k}$ on $\mathfrak{g}$, known as a trace form of the representation (V,ρ), to be $\rho^*(b_V)$. Explicitly, we have

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

Definition 5.2.7. 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 5.2.8. *Let $\mathfrak{a}$ be an ideal of $\mathfrak{g}$. The Killing form $\kappa^{\mathfrak{a}}$ of $\mathfrak{a}$ is given by the restriction of the Killing form $\kappa^{\mathfrak{g}}$ on $\mathfrak{g}$, that is:*

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

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

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

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

where $A\in\mathrm{End}(\mathfrak{a})$ and $B\in\mathrm{Hom}_{\mathsf{k}}(\mathfrak{a},W)$. Then clearly $\mathrm{tr}(\mathrm{ad}(a_1)\mathrm{ad}(a_2))=\mathrm{tr}(A)$. Since A is clearly given by $\mathrm{ad}(a_1)_{|\mathfrak{a}}\mathrm{ad}(a_2)_{|\mathfrak{a}}$, we are done. To see that $\mathfrak{a}^\perp$ is an ideal, we must check that for any $x\in\mathfrak{g}$ and $y\in\mathfrak{a}^\perp$ we have $[x,y]\in\mathfrak{a}^\perp$. But if $a\in\mathfrak{a}$ then $\kappa(a,[x,y])=\kappa([a,x],y)=0$ since $[a,x]\in\mathfrak{a}$.

□

5.2.2 Cartan criterion for solvable Lie algebras

For the rest of 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. Recall that, if (V,ρ) is a representation of a nilpotent Lie algebra $\mathfrak{h}$, then it decomposes as the direct sum $V = \bigoplus_{\lambda\in\Psi_V} V_\lambda$, where $\Psi_V \subseteq D(\mathfrak{h})^0$ denotes the set of one-dimensional representations of $\mathfrak{h}$ which occur as composition factors of V, and V_λ is the maximal subrepresentation of V whose only composition factor is k_λ. When $(\mathfrak{g},\mathfrak{h})$ is a Cartan pair and we view $\mathfrak{g}$ as a $\mathfrak{h}$ representation via the inclusion $\mathfrak{h}\subseteq\mathfrak{g}$, the set $\Psi_\mathfrak{g} = \{0\}\cup\Phi$, where $\mathfrak{g}_0 = \mathfrak{h}$.

Definition 5.2.9. Let $(\mathfrak{g},\mathfrak{h})$ be a Cartan pair and let $\mathfrak{g} = \mathfrak{h}\oplus\bigoplus_{\alpha\in\Phi}\mathfrak{g}_\alpha$ be the associated decomposition of $\mathfrak{g}$. For each $\alpha\in\Phi$ we set $\mathfrak{h}_\alpha = [\mathfrak{g}_\alpha,\mathfrak{g}_{-\alpha}]\subseteq\mathfrak{h}$. Note that if $-\alpha\notin\Phi$, then $\mathfrak{h}_\alpha = \{0\}$.

Lemma 5.2.10. *Let* $(\mathfrak{g},\mathfrak{h})$ *be a Cartan pair and let* $\Phi\subseteq D(\mathfrak{h})^0$ *be the roots of* $\mathfrak{g}$ *associated to its Cartan decomposition. Let* (V,ρ) *be a* $\mathfrak{g}$*-representation and let* $V = \bigoplus_{\lambda\in\Psi_V} V_\lambda$ *be its decomposition into its isotypical summands as an* $\mathfrak{h}$*-representation. If* $\alpha\in\Phi$ *and* $\mathfrak{h}_\alpha\subseteq\mathfrak{h}$ *is as in Definition 5.2.9, then for any* $\lambda\in\Psi_V$*, there is an* $r_\lambda\in\mathbb{Q}$ *such that* $\lambda_{|\mathfrak{h}_\alpha} = r_\lambda.\alpha_{|\mathfrak{h}_\alpha}$.

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 by Lemma 5.1.5,

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

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 $\operatorname{tr}(\rho(z),M) = 0$. On the other hand, we may also compute the trace of $\rho(z)$ on M directly: for any $h\in\mathfrak{h}$, $\rho(h)$ acts on an isotypical summand V_μ with unique eigenvalue $\mu(h)$, hence $\operatorname{tr}_{V_\mu}(\rho(h)) = \dim(V_\mu).\mu(h)$. Applying this to $\rho(z)$ we find

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

Since $\dim(M)\geq\dim(V_\lambda)>0$, this can be rearranged to give $\lambda(z) = r_\lambda.\alpha(z)$ as required. □

Definition 5.2.11. Let $\mathfrak{g}$ be a Lie algebra over a field k. We say that $\mathfrak{g}$ is *perfect* if it satisfies $\mathfrak{g} = D(\mathfrak{g}) = [\mathfrak{g},\mathfrak{g}]$. A perfect Lie algebra therefore has no nontrivial abelian quotients.

Proposition 5.2.12. *Let* V *be a finite-dimensional* k*-vector space and let* $b_V\colon\mathfrak{gl}_V\times\mathfrak{gl}_V\to\mathsf{k}$*, be the trace form,* $b_V(x,y) = tr_V(xy)$*, for all* $x,y\in\mathfrak{gl}_V$*. If* $\mathfrak{a}$ *is a non-zero perfect subalgebra of* $\mathfrak{gl}_V$ *then there is an* $x\in\mathfrak{a}$ *such that* $b_V(x,x)\neq 0$*, so that* b_V *does not vanish identically on* $\mathfrak{a}$.

Proof. Suppose that $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{a}$ so that that $\mathfrak{a} = \bigoplus_{\lambda\in\Phi\cup\{0\}}\mathfrak{a}_\lambda$ is the associated Cartan decomposition, where $\mathfrak{h} = \mathfrak{a}_0$. If we let $V = \bigoplus_{\mu\in\Psi} V_\mu$ be the decomposition of V into generalised $\mathfrak{h}$-weight spaces as in Theorem 4.3.13, then since $\mathfrak{a}\subseteq\mathfrak{gl}_V\cong V^*\otimes V$, it follows that $\Phi\bigcup\{0\}\subseteq\{\mu_1-\mu_2 : \mu_1,\mu_2\in\Psi\}$. Now if $\mathfrak{a} = \mathfrak{g}_0$, then $\mathfrak{h} = \mathfrak{a}$ is nilpotent and hence solvable, but by assumption $\mathfrak{a} = D(\mathfrak{a})$, so this is a contradiction. It follows that Φ must be non-empty, and so in particular there must be some non-zero $\lambda\in\Psi_V$. Next observe that

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

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

$$\mathfrak{h} = [\mathfrak{h},\mathfrak{h}] + \sum_\alpha[\mathfrak{a}_\alpha,\mathfrak{a}_{-\alpha}] = D(\mathfrak{h}) + \sum_\alpha\mathfrak{h}_\alpha$$

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

$$b_V(x,x) = \operatorname{tr}(x^2) = \sum_{\mu \in \Psi} \dim(V_\mu)\mu(x)^2.$$

But now by Lemma 5.2.10 for each $\mu \in \Psi$ there is an $r_\mu \in \mathbb{Q}$ such that $\mu(x) = r_\mu.\alpha(x)$ for all $x \in [\mathfrak{a}_\alpha, \mathfrak{a}_{-\alpha}]$. In particular, $0 \neq \lambda(x) = r_\lambda \alpha(x)$ so that $r_\lambda \neq 0$ and $\alpha(x) \neq 0$. Hence we see that

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

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

□

Recall that for any finite dimensional Lie algebra, the derived series stablizes to an ideal which we denote as $D^\infty(\mathfrak{g})$. It has the property that it is equal to its own derived subalgebra, *i.e.* $D^\infty(\mathfrak{g})$ is perfect.

Theorem 5.2.13. *Let $\mathfrak{g}$ be a Lie algebra and let (V,ρ) be a $\mathfrak{g}$-representation. Then if t_V vanishes on $D(\mathfrak{g})$ then $\rho(\mathfrak{g})$ is solvable, or equivalently, $D^\infty(\mathfrak{g}) \subseteq \ker(\rho)$.*

Proof. Since $\rho(D(\mathfrak{g})) = D(\rho(\mathfrak{g}))$, replacing $\mathfrak{g}$ by its image $\rho(\mathfrak{g})$, we may assume that $\mathfrak{g} \subseteq \mathfrak{gl}_V$ and b_V vanishes on $D(\mathfrak{g})$. We must show that $\mathfrak{g}$ is solvable, that is $D^\infty(\mathfrak{g}) = \{0\}$. But if this is not the case, then setting $\mathfrak{a} = D^\infty(\mathfrak{g})$ it follows that $\mathfrak{a}$ is a non-zero perfect subalgebra of $\mathfrak{gl}_V$. But then the previous Proposition shows there is some $x \in \mathfrak{a}$ for which $b_V(x,x) \neq 0$. But by assumption b_V vanishes identically on $D(\mathfrak{g}) \supseteq \mathfrak{a}$, which gives a contradiction. □

Corollary 5.2.14. *(Cartan's Criterion for Solvability) Let $\mathfrak{g}$ be a (finite-dimensional) k-Lie algebra and let κ denote its Killing form. Then the following are equivalent:*

i) $\kappa(D(\mathfrak{g}), D(\mathfrak{g})) = 0$,

ii) $\mathfrak{g}$ *is solvable,*

iii) $\kappa(\mathfrak{g}, D(\mathfrak{g})) = 0$, *that is,* $D(\mathfrak{g}) \subseteq rad(\kappa)$.

is solvable if and only if the Killing form vanishes on $D(\mathfrak{g})$.

Proof. Clearly *iii*) $\implies$ *i*) so it suffices to show *i*) $\implies$ *ii*) and *ii*) $\implies$ *iii*).

For *i*) $\implies$ *ii*) note that *i*) is the hypothesis of Theorem 5.2.13 for the representation $(\mathfrak{g}, \mathrm{ad})$, and hence the theorem shows that $D^\infty(\mathfrak{g}) \subseteq \ker(\mathrm{ad}) = \mathfrak{z}(\mathfrak{g})$. But since $D^\infty(\mathfrak{g})$ is perfect, this shows that $D^\infty(\mathfrak{g}) = D(D^\infty(\mathfrak{g})) \subseteq D(\mathfrak{z}(\mathfrak{g})) = \{0\}$ and hence *ii*) holds.

For *ii*) $\implies$ *iii*) note that if $\mathfrak{g}$ is solvable, by Lie's theorem if $\mathscr{C} = (V = F_0 > F_1 > \dots > F_d = \{0\})$ is a composition series for $(\mathfrak{g}, \mathrm{ad})$ then each subquotient F_i/F_{i+1} must be one-dimensional, *i.e.* $\mathscr{C}$ is a complete flag. Pick a basis $B = \{e_1, \dots, e_d\}$ of $\mathfrak{g}$ such that $F_k = \langle e_i : i \leq d-k \rangle_{\mathsf{k}}$, and let $\alpha \mapsto [\alpha]_B$ denote the isomorphism $\mathfrak{gl}_\mathfrak{g} \to \mathfrak{gl}_d(\mathsf{k})$ where $[\alpha]_B = (\alpha_{ij}^B)_{1 \leq i,j \leq d}$. This isomorphism identifies $\mathfrak{b}_\mathscr{C} \supset D(\mathfrak{b}_\mathscr{C}) = \mathfrak{n}_\mathscr{C}$ with $\mathfrak{b}_d \supset \mathfrak{n}_d$ the space of upper-triangular matrices and strictly upper triangular matrices in $\mathfrak{gl}_d$ respectively. In particular, for any $x, y \in \mathfrak{b}_\mathscr{C}$ we have $\operatorname{tr}_\mathfrak{g}(xy) = \sum_{i=1}^d x_{ii}^B.y_{ii}^B$, and in particular $\operatorname{tr}_\mathfrak{g}(xy) = 0$ if $x \in \mathfrak{b}_\mathscr{C}, y \in \mathfrak{n}_\mathscr{C}$. But as $\mathscr{C}$ is a composition series for $\mathfrak{g}$, we have $\mathrm{ad}(\mathfrak{g}) \subseteq \mathfrak{b}_\mathscr{C}$, and hence $\mathrm{ad}(D(\mathfrak{g})) = D(\mathrm{ad}(\mathfrak{g})) \subseteq D(\mathfrak{b}_\mathscr{C}) = \mathfrak{n}_\mathscr{C}$. It follows immediately that $\kappa(\mathfrak{g}, D(\mathfrak{g})) \subseteq \operatorname{tr}_\mathfrak{g}(\mathfrak{b}_\mathscr{C}, \mathfrak{n}_\mathscr{C}) = 0$ as required. □

Chapter 6

Semisimple Lie algebras

In this section we assume that our field k *is algebraically closed of characteristic zero, and all representations are assumed to be finite dimensional over* k.

6.1 The solvable radical, semisimplicity, and Cartan's criterion

Suppose that $\mathfrak{g}$ is a Lie algebra, and $\mathfrak{a}$ and $\mathfrak{b}$ are solvable Lie ideals of $\mathfrak{g}$. It is easy to see that $\mathfrak{a}+\mathfrak{b}$ is again solvable (for example, because $0 \subseteq \mathfrak{a} \subseteq \mathfrak{a}+\mathfrak{b}$, and $\mathfrak{a}$ and $(\mathfrak{a}+\mathfrak{b})/\mathfrak{a} \cong \mathfrak{b}/(\mathfrak{a}\cap\mathfrak{b})$ are both solvable). It follows that if $\mathfrak{g}$ is finite dimensional, then it has a largest solvable ideal $\mathfrak{r}$. Note that this is in the strong sense: every solvable ideal of $\mathfrak{g}$ is a subalgebra of $\mathfrak{r}$ (*c.f.* Definition 4.3.7 where the same strategy was used to define the subrepresentation V_S of a $\mathfrak{g}$-representation given an irreducible representation S of $\mathfrak{g}$).

Definition 6.1.1. Let $\mathfrak{g}$ be a finite dimensional Lie algebra. The largest solvable ideal $\mathfrak{r}$ of $\mathfrak{g}$ is known as the (*solvable*) *radical* of $\mathfrak{g}$, and will be denoted $\mathrm{rad}(\mathfrak{g})$. We say that $\mathfrak{g}$ is *semisimple* if $\mathrm{rad}(\mathfrak{g}) = 0$, that is, if $\mathfrak{g}$ contains no non-zero solvable ideals.

Lemma 6.1.2. *The Lie algebra* $\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$ *is semisimple, that is, it has zero radical.*

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

Example 6.1.3. The Lemma shows that any Lie algebra $\mathfrak{g}$ contains a canonical solvable ideal $\mathrm{rad}(\mathfrak{g})$ such that $\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$ is a semisimple Lie algebra. Thus we have a short exact sequence:

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

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

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

Theorem 6.1.4. *(Levi's theorem) Let* $\mathfrak{g}$ *be a finite dimensional Lie algebra over a field* k *of characteristic zero, and let* $\mathfrak{r}$ *be its radical. Then there exists a subalgebra* $\mathfrak{s}$ *of* $\mathfrak{g}$ *such that* $\mathfrak{g} \cong \mathfrak{r} \ltimes \mathfrak{s}$. *In particular* $\mathfrak{s} \cong \mathfrak{g}/\mathfrak{r}$ *is semisimple.*

6.1.1 Cartan's Criterion for semisimplicity

The Killing form gives us a way of detecting when a Lie algebra is semisimple. Recall that, given a symmetric bilinear form $B\colon V \times V \to \mathsf{k}$, the *radical* of B is

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

The form B said to be nondegenerate if $\mathrm{rad}(B) = \{0\}$. We first note the following simple result.

Lemma 6.1.5. *A finite dimensional Lie algebra* $\mathfrak{g}$ *is semisimple if and only if it does not contain any non-zero abelian ideals.*

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

We have the following simple characterisation of semisimple Lie algebras.

Theorem 6.1.6. *A Lie algebra $\mathfrak{g}$ is semisimple if and only if the Killing form is nondegenerate.*

Proof. Let $\mathfrak{g}^{\perp} = \{x \in \mathfrak{g} : \kappa(x,y) = 0, \forall\, y \in \mathfrak{g}\}$. Then by Lemma 5.2.8 $\mathfrak{g}^{\perp}$ 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 5.2.8 the ideal $\mathfrak{g}^{\perp}$ is solvable. It follows that if $\mathfrak{g}$ is semisimple we must have $\mathfrak{g}^{\perp} = \{0\}$ and hence κ is non-degenerate.

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

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

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

Remark 6.1.7. It is worth noting that the proof of the previous theorem establishes two facts: first, that $\mathfrak{g}^{\perp}$ is a solvable ideal in $\mathfrak{g}$ for any Lie algebra $\mathfrak{g}$, and secondly, that any abelian ideal of $\mathfrak{g}$ is contained in $\mathfrak{g}^{\perp}$. Combined with the previous Lemma this shows that $\mathfrak{g}^{\perp} = \{0\} \iff \mathrm{rad}(\mathfrak{g}) = \{0\}$, but in general the containment $\mathfrak{g}^{\perp} \subseteq \mathrm{rad}(\mathfrak{g})$ need not be an equality.

6.2 Simple and semisimple Lie algebras

Definition 6.2.1. Recall from Definition 3.1.1 that a Lie algebra $\mathfrak{g}$ is said to be *almost simple* if it has no non-trivial proper ideals. We say that $\mathfrak{g}$ is *simple* if it is nonabelian and has no nontrivial proper ideal, *i.e.* $\mathfrak{g}$ is almost simple and nonabelian. We now show that this notion is closed related to our notion of a semisimple Lie algebra.

Lemma 6.2.2. *Let V be a k-vector space equipped with a symmetric bilinear form B. Then for any subspace U of V we have*

i) $\dim(U) + \dim(U^{\perp}) \geq \dim(V)$,

ii) the restriction $B_{|U}$ is non-degenerate if and only if $U \oplus U^{\perp} = V$.

Proof. See Lemma I.21 in Appendix I. □

Proposition 6.2.3. *Let $\mathfrak{g}$ be a Lie algebra, and let I be an ideal of $\mathfrak{g}$.*

i) If $\mathfrak{g}$ is semisimple then $\mathfrak{g} = I \oplus I^{\perp}$, and both I and $I^{\perp}$ are semisimple, hence any ideal and any quotient of $\mathfrak{g}$ is semisimple.

ii) I is semisimple if and only if $\mathfrak{g} = I \oplus I^{\perp}$.

Proof. For part *i*) consider $I \cap I^{\perp}$. The Killing form κ 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 part i) of Lemma 6.2.2 we must have $\mathfrak{g} = I \oplus I^{\perp}$. Since this is evidently an orthogonal direct sum, the Killing form must be nondegenerate on both I and $I^{\perp}$, and since they are ideals, Cartan's criterion then implies they are both semisimple. Since the quotient map induces an isomorphism $I^{\perp} \cong \mathfrak{g}/I$ it follows that any quotient of $\mathfrak{g}$ is also semisimple.

For *ii*), note that by part *ii*) of Lemma 6.2.2, κ is non-degenerate on I if and only if $\mathfrak{g} = I \oplus I^{\perp}$. But the restriction of $\kappa^{\mathfrak{g}}$ to I is the Killing form of I, and so Cartan's criterion completes the proof. □

Corollary 6.2.4. *Let $\mathfrak{g}$ be a semisimple Lie algebra and let $\mathrm{Der}_{\mathsf{k}}(\mathfrak{g})$ be the Lie algebra of derivations of $\mathfrak{g}$. Then ad: $\mathfrak{g} \to \mathrm{Der}_{\mathsf{k}}(\mathfrak{g})$ is an isomorphism, so that in particular any derivation of $\mathfrak{g}$ is inner.*

Proof. Suppose that $\delta \in \mathrm{Der}_{\mathsf{k}}(\mathfrak{g})$. Then we may form $\mathfrak{g}_1 = \mathfrak{g} \rtimes_\delta \mathfrak{gl}_1$, the semi-direct product[1] of $\mathfrak{g}$ and $\mathfrak{gl}_1$. Now $\mathfrak{g}$ is a semisimple ideal in $\mathfrak{g}_1$, so by Proposition 6.2.3, $\mathfrak{g}_1 = \mathfrak{g} \oplus \mathfrak{g}^\perp$. But then $[\mathfrak{g}^\perp, \mathfrak{g}] \subseteq \mathfrak{g}^\perp \cap \mathfrak{g} = \{0\}$. and so if $a \in \mathfrak{a}$ is such that $(a, -1) \in \mathfrak{g}^\perp$ then for all $x \in \mathfrak{g}$ we have

$$0 = [(a, -t), (x, 0)] = \mathrm{ad}(a)(x) - \delta(x),$$

and hence $\delta = \mathrm{ad}(a)$ is an inner derivation as required.

□

Theorem 6.2.5. *Let $\mathfrak{g}$ be a semisimple Lie algebra.*

i) There exist ideals $\mathfrak{g}_1, \mathfrak{g}_2, \dots \mathfrak{g}_k \subseteq \mathfrak{g}$ which are simple Lie algebras and for which the natural map:

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

is an isomorphism. In particular, if $\mathfrak{g}$ is semisimple it is perfect, i.e. $D(\mathfrak{g}) = \mathfrak{g}$..

ii) Any simple ideal $\mathfrak{a} \in \mathfrak{g}$ is equal to some $\mathfrak{g}_i$ $(1 \leq i \leq k)$. In particular the decomposition in part i) is unique up to reordering.

Proof. For part *i)* 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. It follows that $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$, hence any ideal in $\mathfrak{a}$ is also an ideal in $\mathfrak{g}$, thus since $\mathfrak{a}$ is a minimal ideal, it must be simple, and so $\mathfrak{g}$ is a direct sum of simple ideals as required. Since a simple Lie algebra is trivially seen to be perfect, each $\mathfrak{g}_i$ is perfect and hence so is $\mathfrak{g}$.

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

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

(the ideals $[\mathfrak{g}_i, \mathfrak{a}]$ are contained in $\mathfrak{g}_i$ so the last sum remains direct). But $\mathfrak{a}$ is simple, so direct sum decomposition must have exactly one nonzero summand and we have $\mathfrak{a} = [\mathfrak{g}_i, \mathfrak{a}]$ for some i $(1 \leq i \leq k)$. Finally, using the simplicity of $\mathfrak{g}_i$ we see that $\mathfrak{a} = [\mathfrak{g}_i, \mathfrak{a}] = \mathfrak{g}_i$ as required.

□

Remark 6.2.6. For any finite-dimensional Lie algebra $\mathfrak{g}$, its solvable radical $\mathrm{rad}(\mathfrak{g})$ is a solvable ideal, and hence has $\mathfrak{gl}_1$ as its only composition factor, while $\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$ is semisimple. The previous Theorem thus shows that $\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$ is a direct sum of simple Lie algebras. Combined with Proposition 6.2.3, this gives a substantial refinement of the Jordan-Hölder theorem of Chapter 3.

Let $\mathscr{C} = (\mathfrak{g} = \mathfrak{g}_0 \rhd \mathfrak{g}_1 \rhd \dots \rhd \mathfrak{g}_d = \{0\})$ be a composition series for $\mathfrak{g}$ with composition factors $\mathfrak{s}_i = \mathfrak{g}_i/\mathfrak{g}_{i+1}$. If $\mathfrak{s}_i$ is simple, then applying Proposition 6.2.3 to $\mathfrak{s}_i \lhd \mathfrak{g}_{i-1}/\mathfrak{g}_{i-1}$ we see that $\mathfrak{g}_{i-1}/\mathfrak{g}_{i-1}$ is a direct sum which (by abuse of notation) we write as $\mathfrak{s}_i \oplus \mathfrak{s}_{i-1}$. It follows that we may modify $\mathfrak{g}_i$ to obtain a new composition series with the composition factor $\mathfrak{s}_i$ now occuring as $\mathfrak{s}_{i-1}$. Applying this repeatedly, we may modify any composition series to obtain a composition series $\mathscr{C}'$ with the property that there is some $k \leq d$ such that $\mathfrak{g}_k = \mathrm{rad}(\mathfrak{g})$, and hence the composition factor $\mathfrak{g}_s/\mathfrak{g}_{s+1} \cong \mathfrak{gl}_1$ for all $s \geq k$, while the composition factors $\mathfrak{g}_s/\mathfrak{g}_{s+1}$ for $s < k$ are all in fact direct summands of $\mathfrak{g}_0/\mathfrak{g}_k$.

6.2.1 The Jordan Decomposition

The following proposition is a consequence of the primary decomposition theorem in linear algebra. For completeness, we provide a proof in the appendices – see Proposition I.2.

[1] Recall that a semidirect product of Lie algebras $\mathfrak{a} \rtimes \mathfrak{b}$ requires, in addition to the two Lie algebras, a homomorphism $\varphi\colon \mathfrak{b} \to \mathrm{Der}_{\mathsf{k}}(\mathfrak{a})$. If $\mathfrak{b} = \mathfrak{gl}_1$ however, φ is determined by $\varphi(1) \in \mathrm{Der}_{\mathsf{k}}(\mathfrak{a})$, *i.e.* we only need to specify the derivation by which $\mathrm{ad}(1)$ acts on $\mathfrak{a}$ in the semidirect product.

Proposition 6.2.7. *Let V be a finite-dimensional* k*-vector space. For any $x \in \mathrm{End}(V)$, there is a direct sum decomposition*

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

The subspace $V_{\lambda,x}$ is known as the λ-generalized eigenspace of x.

If the only $\lambda \in \mathsf{k}$ for which $V_{\lambda,x} \neq \{0\}$ is $\lambda = 0$, so that $x^N = 0$ for some $N \in \mathbb{N}$, we say that x is *nilpotent*. If, for each $v \in V_{\lambda,x}$ we may take $N = 1$, then we say that x is *semisimple*. Equivalently, $x \in \mathrm{End}_{\mathsf{k}}(V)$ is semisimple if there is a direct sum decomposition $V = \bigoplus_{1 \leq i \leq n} L_i$ such that $x(L_i) \subseteq L_i$, *i.e.* in terms of the $\mathfrak{gl}_1$-representation $\rho_x \colon \mathfrak{gl}_1 \to \mathfrak{gl}_V$ given by $\rho_x(c) = c.x$, the representation (V, ρ_x) is semisimple. The generalised eigenspace decomposition above can be used to give a decomposition of the endomorphism x in a semisimple (or diagonalisable) and nilpotent part:

Lemma 6.2.8. *Let $x \colon V \to V$ be a linear map. Then there exists a diagonalisable linear map x_s and a nilpotent linear map x_n such that $x = x_s + x_n$ and $[x_s, x_n] = 0$. Moreover, if $U \leq V$ is x-stable, so that $x(U) \subseteq U$, then x_s and x_n also preserve U.*

Proof. Let $V = \bigoplus_{\lambda \in \mathsf{k}} V_\lambda$ be the generalised eigenspace decomposition of V given by the action of x. Suppose that $\{\lambda_1, \dots, \lambda_k\}$ are the distinct eigenvalues of x, and let $(e_i)_{i=1}^k$ be the projection maps to V_{λ_i}. Then if $x_s = \sum_{i=1}^k \lambda_i.e_i$, clearly x_s is semisimple, and $[x, x_s] = 0$ (since this is evident on each V_{λ_k}). Setting $x_n = x - x_s$, and noting that on each V_{λ_i} the map $x - x_s$ is equal to $x - \lambda_i$, which is nilpotent, we conclude that x_n is nilpotent as required.

Now suppose that $U \leq V$ and $x(U) \subseteq U$. Then we must have $U = \bigoplus_{\lambda \in S(x_{|U})} U_{\lambda,x}$, and it follows directly from the definition that $U_{\lambda,x} = U \cap V_{\lambda,x}$, so that $U = \bigoplus_{\lambda \in S(x)} (U \cap V_\lambda)$. But x_s clearly preserves $U \cap V_\lambda$ for each λ, and thus x_s preserves U. Since $x_n = x - x_s$, it also preserves U as required. □

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

Lemma 6.2.9. *Let V be a* k*-vector space and $x \in \mathrm{End}(V)$.*

i) If $n \in \mathrm{End}(V)$ is such that $[x, n] = 0$ and n is nilpotent. Then we have $V_{\lambda,x} = V_{\lambda,x+n}$.

ii) If $x = s + n$ where $[s, n] = 0$ and s is semisimple, n is nilpotent, then $s = x_s$ and $n = x_n$, that is, the Jordan decomposition is unique.

Proof. For part i) it suffices to show that $V_{\lambda,x} \subseteq V_{\lambda,x+n}$ for all such pairs (x, n) in $\mathrm{End}(V)$. Indeed the lemma clearly follows once one also knows the reverse inclusion, but this follows by considering the pair $(x + n, -n)$. To prove the inclusion, note that since $[x, n] = 0$, we have $n(V_{\lambda,x}) \subseteq V_{\lambda,x}$. But by definition $(x - \lambda)$ is nilpotent on $V_{\lambda,x}$ and hence $(x + n) - \lambda = (x - \lambda) + n$, when restricted to $V_{\lambda,x}$, is the sum of two commuting nilpotent endomorphisms of $V_{\lambda,x}$. It follows from Lemma 4.2.13 that $(x + n) - \lambda$ acts nilpotently on $V_{\lambda,x}$, and hence $V_{\lambda,x} \subseteq V_{\lambda,x+n}$ as required.

For part *ii*), note that by part *i*), x_s and $x_s + x_n = x$ have the same generalised eigenspaces. But as x_s is semisimple, its generalized eigenspaces are precisely its eigenspaces and hence it is completely determined by these. It follows x_s is unique, and hence $x_n = x - x_s$ is also. □

Lemma 6.2.10. *Let V be a vector space and $x \in End(V)$. If x is semisimple then*

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

is also semisimple, and similarly if x is nilpotent.

Proof. First note that the action of $\mathrm{ad}(x)$ on $\mathfrak{gl}_V$ is just the action of x on the tensor product $V^* \otimes V$. When x is nilpotent, the result is proved in Lemma 4.2.14. Alternatively, for $r, s \geq 0$ let $G_{r,s} = x^r \circ \mathfrak{gl}_V \circ x^s \subseteq \mathfrak{gl}_V$, and let $F_k = \sum_{r+s=k} G_{r,s}$. Clearly $G_{r,s} \subseteq G_{r',s'}$ if $r \geq r', s \geq s'$ so that $F_k \supseteq F_{k+1}$ and, if $x^d = 0$, then $G_{r,s} = 0$ if $\max\{r, s\} \geq d$, and hence $F_k = 0$ if $k \geq 2d - 1$. Moreover, $\mathrm{ad}(x)(G_{r,s}) \subseteq G_{r+1,s} + G_{r,s+1} \subseteq F_{r+s+1}$, hence $\mathrm{ad}(x)(F_k) \subseteq F_{k+1}$. It follows $\mathrm{ad}(x)$ is nilpotent as required.

If x is semisimple, then we may write $V = \bigoplus_{i=1}^n L_i$ where $x(L_i) \subseteq L_i$ and $\dim(L_i) = 1$. But then

$$\mathrm{Hom}(V, V) = \mathrm{Hom}\left(\bigoplus_{i=1}^n L_i, \bigoplus_{j=1}^n L_j\right) = \bigoplus_{i,j=1}^n \mathrm{Hom}(L_i, L_j),$$

and clearly the one-dimensional spaces $\mathrm{Hom}(L_i, L_j)$ are preserved by $\mathrm{ad}(x)$, so that $\mathrm{Hom}(V, V)$ decomposes into a direct sum of one-dimensional $\mathrm{ad}(x)$-stable subspaces, and hence $\mathrm{ad}(x)$ is semisimple. $\square$

Corollary 6.2.11. *Let $x \in \mathrm{End}(V)$, and suppose $x = x_s + x_n$ is the Jordan decomposition of x. Then $ad(x) = ad(x_s) + ad(x_n)$ is the Jordan decomposition of $ad(x)$.*

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

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

Definition 6.2.12. Suppose that $\mathfrak{g}$ is a Lie algebra and $x \in \mathfrak{g}$. The endomorphism $\mathrm{ad}(x) \in \mathfrak{gl}_{\mathfrak{g}}$ has a unique Jordan decomposition $\mathrm{ad}(x) = \mathrm{ad}(x)_s + \mathrm{ad}(x)_n$ in $\mathfrak{gl}_{\mathfrak{g}}$. Then if $s, n \in \mathfrak{g}$ are such that $\mathrm{ad}(s) = \mathrm{ad}(x)_s$ and $\mathrm{ad}(n) = \mathrm{ad}(x)_n$, we say the Lie algebra elements s, n are an *abstract Jordan decomposition* of x. Note that if $z \in \mathfrak{z}(\mathfrak{g}) \neq \{0\}$ then if (s, n) is a Jordan decomposition of x so is $(s + z, n - z)$, thus the Jordan decomposition is unique if and only if $\mathfrak{z}(\mathfrak{g}) = \{0\}$.

Note that that if $\mathfrak{g} = \mathfrak{gl}_V$ for some vector space V, then Lemma 6.2.10 shows that the naive Jordan decomposition gives an abstract Jordan decomposition for an element $x \in \mathfrak{gl}_V$, and moreover if $x \in \mathfrak{sl}_V$ then the naive Jordan decomposition $x = x_s + x_n$ has $\mathrm{tr}(x_s) = \mathrm{tr}(x) = 0$ so $x_s, x_n \in \mathfrak{sl}_V$ and the naive Jordan decomposition is the abstract Jordan decomposition by the Corollary above.

Lemma 6.2.13. *Let $\mathfrak{a}$ be a Lie algebra and $\mathrm{Der}_{\mathsf{k}}(\mathfrak{a}) \subset \mathfrak{gl}_{\mathfrak{a}}$ the Lie algebra of k-derivations on $\mathfrak{a}$. Let $\delta \in \mathrm{Der}_{\mathsf{k}}(\mathfrak{a})$. If $\delta = s + n$ is the Jordan decomposition of δ as an element of $\mathfrak{gl}_{\mathfrak{a}}$, then $s, n \in \mathrm{Der}_{\mathsf{k}}(\mathfrak{a})$.*

Proof. We may decompose $\mathfrak{a} = \bigoplus_\lambda \mathfrak{a}_\lambda$ where $\mathfrak{a}_\lambda$ is the generalized eigenspace of δ with eigenvalue $\lambda \in \mathsf{k}$ say. Now since δ is a derivation the map $\mathfrak{a}_\lambda \otimes \mathfrak{a}_\mu \to \mathfrak{a}$ given by $x \otimes y \mapsto [x, y]$ is compatible with the action of δ. But then by Lemma 4.3.11, if $x \in \mathfrak{a}_\lambda$ and $y \in \mathfrak{a}_\mu$, we have $[x, y] \in \mathfrak{a}_{\lambda+\mu}$. It follows immediately that s is a derivation on $\mathfrak{a}$, and since $n = \delta - s$ we see that n is also. $\square$

Theorem 6.2.14. *Let $\mathfrak{g}$ be a semisimple Lie algebra. Then 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 $I = \mathrm{ad}(\mathfrak{g})$, so by Lemma 6.2.13 so do $\mathrm{ad}(x)_s$ and $\mathrm{ad}(x)_n$. But then by part (ii) of Proposition 6.2.3, 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. $\square$

6.3 Representations of semisimple Lie algebras: Weyl's theorem

The goal of this section is to establish the following theorem:

Theorem 6.3.1. *(Weyl's theorem.) Let $\mathfrak{g}$ be a semisimple Lie algebra. If V is a finite-dimensional representation of $\mathfrak{g}$ and U is a subrepresentation of V, then there is a complementary subrepresentation W, that is, W is a subrepresentation and $V = U \oplus W$.*

Remark 6.3.2. The property that every subrepresentation has a complement is called the *semisimplicity* of a representation, so the theorem can be phrased as saying that the finite dimensional representations of a semisimple Lie algebra are semisimple! Note that we showed in Theorem 6.2.5 that any ideal in a semisimple Lie algebra has a complementary ideal, which establishes the semisimplicity of the adjoint representation.

It is easy to see that a semisimple representation is completely reducible, that is, is a direct sum of irreducible subrepresentations (indeed Maschke's theorem for finite groups establishes the same semisimplicity result for suitable representations of finite groups and the argument used to deduce complete reducibility in that setting works in this context also).

Definition 6.3.3. Let (V, ρ) be a $\mathfrak{g}$-representation. We define

$$V^{\mathfrak{g}} = \left\{v \in V : \rho(x)(v) = 0, \forall x \in \mathfrak{g}\right\}, \quad \mathfrak{g}.V = \mathrm{span}_{\mathsf{k}}\left\{\rho(x)(v) : x \in \mathfrak{g}, v \in V\right\}$$

Note $V^{\mathfrak{g}}$ is the subrepresentation of *invariants* we have considered before, and one can check directly that $\mathfrak{g}.V$ is a subrepresentation, or note that it is the image of the $\mathfrak{g}$-homomorphism $a\colon \mathfrak{g} \otimes V \to V$ given by $a(x \otimes v) = \rho(x)(v)$. See Example 2.3.4 for more details. It is the smallest subrepresentation U of V such that $\mathfrak{g}$ acts trivially on V/U.

The key to Weyl's theorem is then the following proposition, whose proof we postpone.

Proposition 6.3.4. *Let $\mathfrak{g}$ be a semisimple Lie algebra and (V, ρ) a representation of $\mathfrak{g}$. Then $V = V^{\mathfrak{g}} \oplus \mathfrak{g}.V$.*

Definition 6.3.5. If V is any vector space and $U \leq V$ is a subspace, a *projection* to U is a linear map $p\colon V \to U$ such that $p_{|U} = 1_U$ (and hence $\mathrm{im}(p) = U$). Equivalently, if $i\colon U \to V$ denotes the inclusion map, $p\colon V \to U$ is a projection to U if $p \circ i = 1_U$. If $p\colon V \to U$ is a projection then $V = U \oplus \ker(p)$. Indeed the sum is direct because if $v \in U \cap \ker(p)$ then $v = p(v) = 0$, hence by rank-nullity it must be all of V. Conversely, if $V = U \oplus W$, then if we define $p_W(v) = u$ where $v = u + w$, $u \in U, w \in W$, the map p_W is a projection to U. Thus we have a bijective correspondence:

$$\Pi_U = \{p \in \mathrm{Hom}(V, U) : p \circ i = 1_U\} \longleftrightarrow \{W \leq V : V = U \oplus W\} = \mathscr{C}_U$$

between Π_U, the set of linear projection from V to U and $\mathscr{C}_U$ the set of complementary subspaces to U in V. If V is a $\mathfrak{g}$-representation, then this bijection restricts to one between $\mathfrak{g}$-invariant projections and complementary subrepresentations.

The direct sum decomposition $V = V^{\mathfrak{g}} \oplus \mathfrak{g}.V$ of Proposition 6.3.4 therefore yields a $\mathfrak{g}$-invariant projection $\pi_0^V\colon V \to V^{\mathfrak{g}}$. Moreover, if (V, ρ) and (W, σ) are finite-dimensional $\mathfrak{g}$-representations and $\phi \in \mathrm{Hom}_{\mathfrak{g}}(V, W)$, it is easy to check that $\phi(V^{\mathfrak{g}}) \subseteq W^{\mathfrak{g}}$ and $\phi(\mathfrak{g}.V) \subseteq \mathfrak{g}.W$, thus we see

$$\pi_0^W \circ \phi = \phi \circ \pi_0^V, \quad \forall \phi \in \mathrm{Hom}_{\mathfrak{g}}(V, W) \tag{6.3.1}$$

Remark 6.3.6. The maps π_0^V are the analogues for a semisimple Lie algebra $\mathfrak{g}$ of the "averaging" operators a_V for representations of a finite group G where, for a G-representation (V, τ), the operator a_V is given by $|G|^{-1}\sum_{g\in G}\tau(g)$. The operators a_V play a crucial role in the proof of Maschke's theorem, and are compatible with G-homomorphisms in the same sense that the π_0^V are compatible with $\mathfrak{g}$-homomorphisms, that is, they satisfy $\phi \circ a_V = a_W \circ \phi$ for any homomorphism of G-representations $\phi\colon V \to W$.

Proof of Weyl's theorem:

Let $i\colon U \to V$ denote the inclusion of a subrepresentation U of V, where V is a finite-dimensional $\mathfrak{g}$-representation. By Definition 6.3.5, we must show that there is a $\mathfrak{g}$-invariant projection from $s\colon V \to U$, since then $\ker(s)$ will be a complementary subrepresentation to U. Let $H_1 = \mathrm{Hom}(V, U)$ and $H_2 = \mathrm{Hom}(U, U)$, and let $i^*\colon H_1 \to H_2$ denote the restriction map $\phi \mapsto \phi_{|U} = \phi \circ i$. Since V and U are $\mathfrak{g}$-representations, H_1 and H_2 are $\mathfrak{g}$-representations and, moreover, it follows from the fact that i is a $\mathfrak{g}$-homomorphism that i^* is a homomorphism of $\mathfrak{g}$-representations[2]. The set of projections to U is $\Pi = \{p \in H_1 : i^*(p) = 1_U\} \subseteq H_1$, and $p \in \Pi$ is $\mathfrak{g}$-homomorphism if $p \in H_1^{\mathfrak{g}}$, thus we need to show that $\Pi \cap H_1^{\mathfrak{g}}$ is nonempty.

We claim that for any $p \in \Pi$, its invariant part $\pi_0^{H_1}(p) \in \Pi \cap H_1^{\mathfrak{g}}$, so that $\ker(\pi_0^{H_1}(p))$ is a complementary subrepresentation to U as required. To see that $\pi_0^{H_1}(p) \in \Pi$ note by (6.3.1) we have

$$i^*(\pi_0^{H_1}(p)) = \pi_0^{H_2}(i^*(p)) = \pi_0^{H_2}(1_U) = 1_U.$$

where the second equality holds because p is a projection, and the third since $1_U \in \mathrm{Hom}_{\mathfrak{g}}(U, U) = H_2^{\mathfrak{g}}$.

6.3.1 Casimir operators

Lemma 6.3.7. *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. Equivalently, t_V induces a nondegenerate invariant form on $\rho(\mathfrak{g})$.*

[2] Explicitly, if $x \in \mathfrak{g}, \phi \in H_1$ then $x(i^*(\phi)) = x_{|U} \circ (\phi \circ i) - (\phi \circ i) \circ x_{|U} = (x_{|U} \circ \phi) \circ i - (\phi \circ x) \circ i = i^*(x(\phi))$

Proof. If $\rho(\mathfrak{g}) = 0$ then the Lemma holds trivially. Otherwise, we may replace $\mathfrak{g}$ by its image $\mathfrak{g}_1 = \rho(\mathfrak{g}) \neq 0$, which, since it is nonzero, is semisimple because $\mathfrak{g}$ is. Now let $\mathfrak{r} = \mathrm{rad}(t_V)$, an ideal in $\mathfrak{g}_1$. Since it $\mathfrak{g}_1$ is semisimple, it follows that $\mathfrak{r}$ is semisimple (or zero), and hence by part *i)* of Theorem 6.2.5, $D(\mathfrak{r}) = \mathfrak{r}$, that is, $\mathfrak{r}$ is perfect. But by Proposition 5.2.12, the trace form on $\mathfrak{gl}_V$ does not vanish identically on any nonzero perfect subalgebra of $\mathfrak{gl}_V$, hence we must have $\mathfrak{r} = \{0\}$ as required. □

Definition 6.3.8. Let $\mathfrak{g}$ be a semisimple Lie algebra and (V,ρ) a representation of $\mathfrak{g}$ with $\rho(\mathfrak{g}) \neq 0$. Then if we let $\mathfrak{g}_1 = \rho(\mathfrak{g})$, then by Lemma 6.3.7, t_V is nondegenerate on $\mathfrak{g}_1$, and so induces an isomorphism of $\mathfrak{g}$-representations $\theta_V\colon \mathfrak{g}_1 \to \mathfrak{g}_1^*$. Let $\tau\colon \mathfrak{g}_1^* \to \mathfrak{gl}_V$ denote the composition of θ_V^{-1} with $i\colon \mathfrak{g}_1 \to \mathfrak{gl}_V$ the inclusion map. We have a sequence of $\mathfrak{g}_1$- (and $\mathfrak{g}$-) homomorphisms

$$(\mathfrak{g}_1 \otimes \mathfrak{g}_1)^* = \mathfrak{g}_1^* \otimes \mathfrak{g}_1^* \xrightarrow{\tau\otimes\tau} \mathfrak{gl}_V \otimes \mathfrak{gl}_V \xrightarrow{c} \mathfrak{gl}_V,$$

where c is just composition of linear maps. These are both homomorphisms of $\mathfrak{g}$-representations since we have already seen that τ is, and the fact that c is follows the discussion in the proof of Lemma 5.2.4.[3] We thus obtain the *Casimir operator*,

$$C = C_V = c(\tau \circ \tau)(t_V) \in \mathfrak{gl}_V{}^{\mathfrak{g}}.$$

The fact that $C \in \mathfrak{gl}_V{}^{\mathfrak{g}}$ means that $C\rho(x) = \rho(x)C$ for all $x \in \mathfrak{g}$, that is, C_V is a $\mathfrak{g}$-endomorphism of V.

Lemma 6.3.9. *Let $\{x_1, \dots, x_n\}$ be a basis of $\mathfrak{g}_1$ and $\{\delta_1, \dots, \delta_n\}$ the corresponding dual basis of $\mathfrak{g}_1^*$. Then $t_V = \sum_{i=1}^n \delta_i \otimes \theta_V(x_i)$, and hence $C_V = \sum_{i=1}^n y_i x_i \in \mathfrak{gl}_V$, where $y_i = \theta_V^{-1}(\delta_i)$. It follows that*

i) $V^{\mathfrak{g}} \subseteq \ker(C_V)$ *and* $\mathrm{im}(C_V) \subseteq \mathfrak{g}.V$,

ii) $\mathrm{tr}(C_V) = \dim(\mathfrak{g}_1) = \dim(\rho(\mathfrak{g}))$.

Proof. Clearly we may write any element of $\mathfrak{g}^* \otimes \mathfrak{g}^*$ in the form $\sum_{i=1}^n \delta_i \otimes f_i$ for some $f_i \in \mathfrak{g}^*$.[4] But by definition, $\theta_V(x_i)(y) = t_V(x_i, y) = \sum_{j=1}^n \delta_j(x_i).f_j(y) = f_i(y)$, hence $t_V = \sum_{i=1}^n \delta_i \otimes \theta_V(x_i)$ as claimed. Thus $C_V = c \circ (\tau \otimes \tau)(t_V) = c(\sum_{i=1}^n \theta_V^{-1}(\delta_i) \otimes x_i) = \sum_{i=1}^n y_i x_i \in \mathfrak{gl}_V$. Since $x_i, y_i \in \mathfrak{g}_1$, part *i)* is immediate from the definitions, while for *ii)* since $\mathrm{tr}(y_i x_i) = t_V(y_i, x_i) = 1$ it is clear that $\mathrm{tr}(C_V) = \sum_{i=1}^n 1 = n = \dim(\mathfrak{g}_1)$. □

Example 6.3.10. Let us take $\mathfrak{g} = \mathfrak{sl}_2 \subseteq \mathfrak{gl}_2$. Then the trace form $t(x,y) = \mathrm{tr}(x.y)$ is non-degenerate and invariant, and

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

so if we let $\delta_e, \delta_f, \delta_h$ be the basis of $\mathfrak{g}^*$ dual to $\{e,f,h\}$, we see that $t = (1+\sigma)(\delta_e \otimes \delta_f + \delta_h \otimes \delta_h)$ where $\sigma \in \mathrm{End}(\mathfrak{g}^* \otimes \mathfrak{g}^*)$ is the map $\sigma(a \otimes b) = b \otimes a$. It follows that $\theta_V(e) = \delta_f$, $\theta_V(f) = \delta_e$ and $\theta_V(h) = \frac{1}{2}\delta_h$, and hence $C_V = ef + fe + \frac{1}{2}h^2$.

Proof of Proposition 6.3.4: We prove the statement by induction on $\dim(V)$, the case $\dim(V) = 0$ being trivial. If $V = V^{\mathfrak{g}}$ then certainly $\mathfrak{g}.V = \{0\}$ and the statement holds. Thus we may assume that $V \neq V^{\mathfrak{g}}$, so that $\rho(\mathfrak{g}) \neq \{0\}$, hence we have a Casimir operator $C \in \mathfrak{gl}_V$. 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 proposition holds for representations U and W it certainly holds for their direct sum $U \oplus W$, we are done by induction unless C has exactly one generalised eigenspace, *i.e.* $V = V_\lambda$. But then by part *ii)* of Lemma 6.3.9, $\dim(V).\lambda = \mathrm{tr}(C) = \dim(\rho(\mathfrak{g}))$, so that $\lambda \neq 0$[5], and hence C is invertible. The by part *i)* of Lemma 6.3.9 we have $V^{\mathfrak{g}} \subseteq \ker(C) = \{0\}$ and $\mathrm{im}(C) = V \subseteq \mathfrak{g}.V$, so that $V = \{0\} \oplus \mathfrak{g}.V = V^{\mathfrak{g}} \oplus \mathfrak{g}.V$ as required.

6.3.2 The Jordan decomposition: functoriality

Given a representation (V,ρ) of $\mathfrak{g}$, it is thus natural to ask whether $\rho(x) = \rho(s) + \rho(n)$ is again the naive Jordan decomposition of $\rho(x)$.

Theorem 6.3.11. *Let $\mathfrak{g}$ be a semisimple Lie algebra and let (V,ρ) be a representation of $\mathfrak{g}$. Then if $s \in \mathfrak{g}$ is semisimple, so is $\rho(s)$, and similarly if $n \in \mathfrak{g}$ is nilpotent, then so is $\rho(n)$. In particular, if $x \in \mathfrak{g}$ has abstract Jordan decomposition $x = s + n$, then $\rho(x) = \rho(s) + \rho(n)$ is the naive Jordan decomposition of $\rho(x)$.*

[3] It is also equivalent to the fact that, for any $a \in \mathfrak{gl}_V$, the map $\mathrm{ad}(a)$ is a derivation for the associative algebra $\mathrm{End}(V)$.

[4] Indeed if V and W are vector spaces and $B = \{e_1, \dots, e_d\}$ is a basis of V, then any element of $V \otimes W$ may be written uniquely as $\sum_{i=1}^d e_i \otimes w_i$ for $w_i \in W$, $(1 \leq i \leq d)$.

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

Proof. Let us first show that the final sentence follows from the fact that $\rho(s)$ is semisimple if s is semisimple and $\rho(n)$ is nilpotent if n is nilpotent. Indeeed since $[s,n] = 0$ and ρ is a Lie algebra bomomorphism, $[\rho(s), \rho(n)] = 0$, thus $\rho(s)$ semisimple and $\rho(n)$ nilpotent implies that the pair $(\rho(s), \rho(n))$ satisfy the characterising property of the naive Jordan decomposition established in Lemma 6.2.9.

- $\rho(n)$ *is nilpotent*: Let $\iota\colon \mathfrak{gl}_1 \to \mathfrak{g}$ be given by $\iota(t) = t.n$, so that $(\mathfrak{g}, \mathrm{ad}\circ\iota)$ and $(V, \rho\circ\iota)$ are $\mathfrak{gl}_1$-representations. Let $V = \bigoplus_{\lambda\in\Psi(n)} V_\lambda$ be the generalised weight-space decomposition of V as a representation of $\mathfrak{gl}_1$ – that is, the generalised eigenspace decomposition of V with respect to $\rho(\iota(1)) = \rho(n)$. Since $\mathrm{ad}(n)$ is nilpotent, $\mathfrak{g} = \mathfrak{g}_0$ as a representation of $\mathfrak{gl}_1$.

  Fix $\lambda \in \Psi(n)$. If $\tilde{a}\colon \mathfrak{g}_0 \otimes V_\lambda \to V$ is given by $\tilde{a}(x\otimes v) = \rho(x)(v)$, then $\tilde{a}$ is a $\mathfrak{g}$-homomorphism, and hence a homomorphism of $\mathfrak{gl}_1$-representations. Since $\mathfrak{g}_0 \otimes V_\lambda \subseteq (\mathfrak{g}\otimes V)_{\lambda+0}$ it follows that $\mathfrak{g}.V_\lambda \subseteq V_\lambda$, that is V_λ is a $\mathfrak{g}$-subrepresentation of V. But now $\mathfrak{g} = D(\mathfrak{g})$, hence if $\mathfrak{g}_\lambda$ denotes its image in $\mathfrak{gl}_{V_\lambda}$, $D(\mathfrak{g}^\lambda) = \mathfrak{g}^\lambda$, so that $\mathfrak{g}^\lambda \subseteq \mathfrak{sl}_{V_\lambda}$. But then $\mathrm{tr}_{V_\lambda}(\rho(n) = \lambda.\dim(V_\lambda) = 0$, hence $\lambda = 0$. It follows that $V = V_0$ and $\rho(n)$ is nilpotent as required.

- $\rho(s)$ *is semisimple*: Since $\mathfrak{g}$ is semisimple, Weyl's theorem ensures that V is completely reducible, and so it suffices to check that $\rho(s)$ is semisimple in the case where V is irreducible. Let $V = \bigoplus_{\lambda\in S_V\subseteq \mathsf{k}} V_\lambda$ be the generalised eigenspace decomposition of $\rho(s)$, where $S_V \subseteq \mathsf{k}$ is the set of eigenvalues of $\rho(s)$, and let $\mathfrak{g} = \bigoplus_{\alpha\in S_{\mathfrak{g}}} \mathfrak{g}^s_\alpha$ be the decomposition of $\mathfrak{g}$ into the eigenspaces of $\mathrm{ad}(s)$ (since $\mathrm{ad}(s)$ is semisimple, $\mathfrak{g}$ is the direct sum of its $\mathrm{ad}(s)$-eigenspaces). Let $V^s_\lambda \subseteq V_\lambda$ be the $\rho(s)$-eigenspace of $\rho(s)$ inside the generalised eigenspace and let $V^s = \bigoplus_{\lambda\in\Psi_V} V^s_\lambda$. We claim that V^s is a $\mathfrak{g}$-subrepresentation of V. Note that the claim establishes the semisimplicity of $\rho(s)$, since $V^s_\lambda \neq 0$ if and only if $V_\lambda \neq 0$, so we must have $0 \neq V^s \subseteq V$. But since V is irreducible, it follows $V^s = V$ as required.

  To verify the claim we may assume that $v \in V_\lambda$ and $x \in \mathfrak{g}_\alpha$. Then

  $$\begin{aligned}\rho(s)\big(\rho(x)(v)\big) &= \big(\rho([s,x]) + \rho(x)\rho(s)\big)(v)\\ &= \alpha(s)\rho(x)(v) + \lambda(s)\rho(x)(v) = (\alpha+\lambda)(s)\big(\rho(x)(v)\big),\end{aligned}$$

  that is $\rho(x)(v) \in V^s_{\alpha+\lambda} \subset V^s$, and $V^s \leq V$ is a subrepresentation of V as claimed.

□

Remark 6.3.12. Note that the proof that $\rho(n)$ is nilpotent does not require that $\mathfrak{g}$ is semisimple, it only requires that $\mathfrak{g}$ be perfect. On the other hand, the proof that $\mathrm{ad}(s)$ is semisimple uses Weyl's theorem, to reduce to the semisimple case. In fact it is the case that if $\mathfrak{g}$ is a perfect Lie algebra in characteristic zero, then every element $x \in \mathfrak{g}$ has an abstract Jordan decomposition $x = s + n$, and that decomposition yields the naive Jordan decomposition of its image $\rho(x) \in \mathfrak{gl}_V$ for any finite-dimensional representation (V, ρ) of $\mathfrak{g}$.

Chapter 7

The structure of semisimple Lie algebras

If $(\mathfrak{g}, \mathfrak{h})$ is a Cartan pair, the amount of information captured by the Cartan decomposition $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha\in\Phi} \mathfrak{g}_\alpha$ depends on $\mathfrak{g}$. At one extreme, when $\mathfrak{g}$ is nilpotent, we have the trivial decomposition $\mathfrak{h} = \mathfrak{g}$. The semisimple case is in some sense at the opposite extreme: the Cartan subalgebra $\mathfrak{h}$ turns out to be abelian and the decomposition of $\mathfrak{g}$ is as fine as possible – $\mathfrak{g}$ is a semisimple $\mathfrak{h}$-representation.

7.1 The Killing form and the Cartan decomposition

Proposition 7.1.1. *Let $(\mathfrak{g}, \mathfrak{h})$ be a Cartan pair and let $\mathfrak{g} = \bigoplus_{\lambda\in\Phi_0} \mathfrak{g}_\lambda$ be the associated Cartan decomposition of $\mathfrak{g}$, where $\Phi_0 = \{0\} \cup \Phi$, $\mathfrak{g}_0 = \mathfrak{h}$, and let κ denote the Killing form of $\mathfrak{g}$.*

i) We have $\kappa(\mathfrak{g}_\lambda, \mathfrak{g}_\mu) = 0$ unless $\lambda + \mu = 0$.

ii) For any $\lambda \in \Phi_0$ the restriction of κ to $\mathfrak{g}_\lambda \times \mathfrak{g}_{-\lambda}$ gives a linear map $\theta_\lambda : \mathfrak{g}_\lambda \to \mathfrak{g}^*_{-\lambda}$. The Killing form κ is nondegenerate, and hence $\mathfrak{g}$ semisimple, if and only if, for every $\lambda \in \Phi$ the linear map θ_λ is an isomorphism. In particular:

a) the restriction of κ to $\mathfrak{h} = \mathfrak{g}_0$ is nondegenerate,

b) $\dim(\mathfrak{g}_\lambda) = \dim(\mathfrak{g}_{-\lambda})$ and so $\lambda \in \Phi$ if and only if $-\lambda \in \Phi$.

Proof. Let $\theta\colon \mathfrak{g} \to \mathfrak{g}^*$ be the map given by $\theta(x)(y) = \kappa(x, y)$ for $x, y \in \mathfrak{g}$. By 2.2.1, $\mathfrak{g}^* \cong \bigoplus_{\lambda\in\Phi_0} \mathfrak{g}^*_\lambda$ as an $\mathfrak{h}$-representation. Now $\mathfrak{g}_\lambda$ is the λ-generalised weight space of $\mathfrak{g}$, hence its only composition factor is k_λ. By Lemma 2.2.9 it follows that $\mathfrak{g}^*_\lambda$ has $\mathsf{k}_{-\lambda}$ as its unique composition factor, and hence $\mathfrak{g}^* = \bigoplus_{\lambda\in\Phi_0} \mathfrak{g}^*_\lambda$ gives the $\mathfrak{h}$-weight isotypical decomposition of $\mathfrak{g}^*$ where $(\mathfrak{g}^*)_\lambda \cong (\mathfrak{g}_{-\lambda})^*$.

Since κ is invariant, $\theta \in \mathrm{Hom}(\mathfrak{g}, \mathfrak{g}^*)^{\mathfrak{g}} = \mathrm{Hom}_{\mathfrak{g}}(\mathfrak{g}, \mathfrak{g}^*)$, *i.e.* it is a homomorphism of $\mathfrak{g}$-representations. In particular, it is an $\mathfrak{h}$-homomorphism, so that $\theta(\mathfrak{g}_\lambda) \subseteq (\mathfrak{g}^*)_\lambda = \mathfrak{g}^*_{-\lambda}$. It follows that $\kappa(\mathfrak{g}_\lambda, \mathfrak{g}_\mu) = \theta(\mathfrak{g}_\lambda)(\mathfrak{g}_\mu) = 0$ unless $\mu = -\lambda$ as claimed. Moreover, if $\mathfrak{g}$ is semisimple, then κ is nondegenerate, *i.e.* θ is an isomorphism. But as θ is an $\mathfrak{h}$-homomorphism this forces $\theta_\lambda = \theta_{|\mathfrak{g}_\lambda} : \mathfrak{g}_\lambda \to \mathfrak{g}^*_{-\lambda}$ to be an isomorphism for all λ as required. □

Remark 7.1.2. When $\mathfrak{g}$ is semisimple, by Proposition 7.1.1, κ induces an isomorphism $\theta_0\colon \mathfrak{h} \to \mathfrak{h}^*$. Given $\lambda \in \mathfrak{h}^*$, we will write $t_\lambda \in \mathfrak{h}$ for $\theta_0^{-1}(\lambda)$, so that t_λ is uniquely determined by the condition that $\kappa(t_\lambda, h) = \lambda(h)$ for all $h \in \mathfrak{h}$.

Lemma 7.1.3. *If $(\mathfrak{g}, \mathfrak{h})$ is a Cartan pair and κ is the Killing form of $\mathfrak{g}$, then we have*

$$\kappa(h_1, h_2) = \sum_{\alpha\in\Phi} \dim(\mathfrak{g}_\alpha)\alpha(h_1)\alpha(h_2), \quad \forall h_1, h_2 \in \mathfrak{h}. \tag{7.1.1}$$

If $\mathfrak{g}$ is semisimple, it follows that

i) $\langle\Phi\rangle_{\mathsf{k}} = \mathfrak{h}^*$

ii) $\mathfrak{h}$ is abelian.

iii) Recall $\mathfrak{h}_\alpha = [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subseteq \mathfrak{h}$. We have $\mathfrak{h}_\alpha \cap \ker(\alpha) = \{0\}$, and hence $\dim(\mathfrak{h}_\alpha) \le 1$.

Proof. Since $\mathfrak{g} = \bigoplus_{\lambda\in\Phi_0} \mathfrak{g}_\lambda$ as an $\mathfrak{h}$-representation, $\kappa(h_1,h_2) = \sum_{\lambda\in\Phi_0} t_{\mathfrak{g}_\lambda}(h_1,h_2)$. Equation (7.1.1) thus follows immediately by applying Lemma 2.2.8 to the $\mathfrak{h}$-representations $\mathfrak{g}_\lambda$.

Now suppose that $\mathfrak{g}$ is semisimple and let $S = \langle\Phi\rangle_{\mathsf{k}}$. Since $\Phi \subseteq D(\mathfrak{h})^0 \subseteq \mathfrak{h}^*$, clearly $S \subseteq D(\mathfrak{h})^0$. Let $S^0 = \Phi^0 = \{h \in \mathfrak{h} : \alpha(h) = 0, \forall\, \alpha \in \Phi\}$, where, since $\mathfrak{h}$ is finite-dimensional, we view S^0 as a subspace of $\mathfrak{h}$ via the canonical isomorphism $(\mathfrak{h}^*)^* \cong \mathfrak{h}$. It is clear from (7.1.1) that $S^0 \subseteq \mathrm{rad}(\kappa_{|\mathfrak{h}})$, but if $\mathfrak{g}$ is semisimple, part *ii*) of Proposition 7.1.1 shows that this is $\{0\}$. But then $S = (S^0)^0 = \mathfrak{h}^*$, so that Φ spans $\mathfrak{h}^*$ establishing *i*). But $S \subseteq D(\mathfrak{h})^0$ hence $D(\mathfrak{h})^0 = \mathfrak{h}^*$ and hence $D(\mathfrak{h}) = \{0\}$, establishing part *ii*). Finally, by Lemma 5.2.10 applied to the adjoint representation of $\mathfrak{g}$, if $\beta \in \Phi$, then we have $\beta_{|\mathfrak{h}_\alpha} = r_\beta.\alpha_{|\mathfrak{h}_\alpha}$ for some $r_\beta \in \mathbb{Q}$. But then if $z \in \mathfrak{h}_\alpha \cap \ker(\alpha)$ it follows $\beta(z) = r_\beta.\alpha(z) = 0$ for all $\beta \in \Phi$, and hence $z \in S^0 = \{0\}$ as required. Since $\dim(\ker(\alpha)) = \dim(\mathfrak{h}) - 1$, clearly $\dim(\mathfrak{h}_\alpha) \leq 1$. $\square$

7.1.1 The roots of a semisimple Lie algebra

Definition 7.1.4. The standard basis for $\mathfrak{sl}_2(\mathsf{k})$ is $e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. The relations obeyed by $\{e,h,f\}$ in $\mathfrak{sl}_2$ are $[e,f] = h, [h,e] = 2e$ and $[h,f] - 2f$. If $\mathfrak{g}$ is an arbitrary Lie algebra, an $\mathfrak{sl}_2$*-triple* is a triple $\{E,H,F\}$ of elements of $\mathfrak{g}$ which obey the same relations, so $[E,F] = H$, $[H,E] = 2E$ and $[H,F] = -2F$. Such a triple determines a homomorphism $\theta\colon \mathfrak{sl}_2 \to \mathfrak{g}$ where $\theta(e) = E$, $\theta(h) = H$ and $\theta(f) = F$. Since $\mathfrak{sl}_2$ is simple, such a homomorphism is determined by its image up to a scalar.

Definition 7.1.5. Recall that, as in Definition 2.2.10, if W is an $\mathfrak{h}$-representation then we write W^s for the socle of W, that is, the sum of all irreducible $\mathfrak{h}$-subrepresentations of W. If $(\mathfrak{g},\mathfrak{h})$ is a Cartan pair and $\mathfrak{g} = \bigoplus_{\lambda\in\Phi_0} \mathfrak{g}_\lambda$, is the decomposition of $\mathfrak{g}$ as an $\mathfrak{h}$-representation, then the $\mathfrak{h}$-representation $\mathfrak{g}_\lambda$ has k_λ as its only composition factor and hence $\mathfrak{g}_\lambda^s = \{x \in \mathfrak{g}_\lambda : [h,x] = \lambda(h).x, \forall\, h \in \mathfrak{h}\}$, that is, if $x \in \mathfrak{g}_\lambda^s$ and $x \neq 0$ then $\mathsf{k}.x \cong \mathsf{k}_\alpha$ as $\mathfrak{h}$-representations. For example, $\mathfrak{g}_0^s = \mathfrak{g}^\mathfrak{h}$, *i.e.* $\mathfrak{g}_0^s$ is the subrepresentation of $\mathfrak{h}$-invariants in $\mathfrak{g}$.

The following Lemma is the first step in constructing a copy of $\mathfrak{sl}_2$ for each pair $\{\pm\alpha\}$ of roots in a semisimple Lie algebra.

Lemma 7.1.6. *Let $(\mathfrak{g},\mathfrak{h})$ be a Cartan pair such that $\mathfrak{g}$ is semisimple, and let $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha\in\Phi} \mathfrak{g}_\alpha$ the associated Cartan decomposition. If $x \in \mathfrak{g}_\alpha^s$ and $y \in \mathfrak{g}_{-\alpha}$ then $[x,y] = \kappa(x,y).t_\alpha$. Moreover if $e_\alpha \in \mathfrak{g}_\alpha^s$ is nonzero, then $ad(e_\alpha)(\mathfrak{g}_{-\alpha}) = \mathsf{k}.t_\alpha = \mathfrak{h}_\alpha$ so that $\mathfrak{h} = \mathfrak{h}_\alpha \oplus \ker(\alpha)$. In particular, if $h_\alpha \in \mathfrak{h}_\alpha$ is given by $\alpha(h_\alpha) = 2$, there is an $f_\alpha \in \mathfrak{g}_{-\alpha}$ such that $[e_\alpha, f_\alpha] = h_\alpha$.*

Proof. For *i*) take any $x \in \mathfrak{g}_\alpha^s, y \in \mathfrak{g}_{-\alpha}$ and $h \in \mathfrak{h}$. Then for all $h \in \mathfrak{h}$,

$$\theta_0([x,y])(h) = \kappa(h,[x,y]) = \kappa([h,x],y) = \kappa(x,y)\alpha(h)$$

Thus $[x,y] = \theta_0^{-1}(\kappa(x,y)\alpha) = \kappa(x,y)t_\alpha$. By part *iii*) of Lemma 7.1.6, we know $\mathfrak{h}_\alpha \cap \ker(\alpha) = \{0\}$. Hence $\mathfrak{h}_\alpha =\subseteq \mathsf{k}.t_\alpha$. But since $\theta_\alpha\colon \mathfrak{g}_\alpha \to \mathfrak{g}_{-\alpha}^*$ is an isomorphism, for any non-zero $e_\alpha \in \mathfrak{g}_\alpha^s$ we may find $y \in \mathfrak{g}_{-\alpha}$ with $\kappa(e_\alpha, y) \neq 0$. Hence there is an $f_\alpha \in \mathsf{k}.y$ with $\kappa(e_\alpha, f_\alpha) = 2\alpha(t_\alpha)^{-1}$ so that $[e_\alpha, f_\alpha] = h_\alpha$ and moreover $\mathfrak{h}_\alpha \supseteq \mathrm{ad}(e_\alpha)(\mathsf{k}.f_\alpha) = \mathsf{k}.h_\alpha$ hence $\mathfrak{h}_\alpha = \mathsf{k}.h_\alpha = \mathsf{k}.t_\alpha$.

$\square$

Definition 7.1.7. Given a root $\alpha \in \Phi$, the element h_α is known as the *coroot* associated to α. We will write $\Phi^\vee = \{h_\alpha : \alpha \in \Phi\} \subset \mathfrak{h}$ for the set of coroots.

Proposition 7.1.8. *Let $\mathfrak{g}$ be a semisimple Lie algebra and $\mathfrak{h}$ a Cartan subalgebra with Cartan decomposition $\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha\in\Phi} \mathfrak{g}_\alpha$. Then*

i) The root spaces $\mathfrak{g}_\alpha$ are one-dimensional, and if $\alpha \in \Phi$, $c \in \mathbb{Z}$, then $c.\alpha \in \Phi$ if and only if $c = \pm 1$.

ii) $\mathfrak{sl}_\alpha = \mathfrak{g}_\alpha \oplus \mathfrak{h}_\alpha \oplus \mathfrak{g}_{-\alpha}$ is a subalgebra of $\mathfrak{g}$ isomorphic to $\mathfrak{sl}_2(\mathsf{k})$.

iii) $\mathfrak{g}$ is a semisimple $\mathfrak{h}$-representation, so that for all $h \in \mathfrak{h}$, $ad(h)$ is semisimple, and hence $\mathfrak{h}$ consists of semisimple elements, i.e. the Jordan decomposition of $h \in \mathfrak{h}$ is $h = h_s$.

Proof. Fix $\alpha \in \Phi$, and let $\{e_\alpha, f_\alpha, h_\alpha\}$ be as in Lemma 7.1.6, so that $\{e_\alpha, h_\alpha, f_\alpha\} \subseteq M$ where

$$M = \bigoplus_{k\leq 1} M_{k.\alpha}, \quad M_\alpha = \mathsf{k}.e_\alpha,\ M_{p.\alpha} = \mathfrak{g}_{p.\alpha}, \forall\, p \leq 0.$$

We claim that M is a subalgebra. Since $M_{p.\alpha} \subseteq \mathfrak{g}_{p.\alpha}$ for all $p \in \mathbb{Z}$, it suffices to check that $[M_{k.\alpha}, M_{l.\alpha}] \subseteq M_{(k+l).\alpha}$ for $k \leq l \leq 1$. For $k+l \leq 0$ this is clear because $M_{(k+l)\alpha} = \mathfrak{g}_{(k+l)\alpha}$ while if $k = 0$ it is equivalent to $M_{l.\alpha}$ being an $\mathfrak{h}$-subrepresentation which we have already checked unless $l = 1$, but as $e_\alpha \in \mathfrak{g}_\alpha^s$, $M_\alpha \cong \mathsf{k}_\alpha$. Finally if $k = l = 1$, then $[M_\alpha, M_\alpha] = \mathsf{k}.[e_\alpha, e_\alpha] = 0 = M_{2\alpha}$.

It follows that $\mathrm{ad}(h_\alpha) = [\mathrm{ad}(e_\alpha), \mathrm{ad}(f_\alpha)]$ acts on M with trace zero. But $M = \bigoplus_{k\leq 1} M_{k.\alpha}$ is the decomposition of M into $\mathrm{ad}(h_\alpha)$ generalised eigenspaces where $M_{k\alpha} = M_{2k,h_\alpha}$, hence

$$0 = \sum_{p\leq 1} 2p.\dim(M_{p.\alpha}) = 2 - \sum_{p\geq 1} 2p.\dim(\mathfrak{g}_{-p.\alpha}) = 2.\big(1 - \sum_{p>0} p.\dim(\mathfrak{g}_{-p.\alpha})\big)$$

If $p > 1$ then $p.\dim(\mathfrak{g}_{-p.\alpha}) > 1$ unless $\dim(\mathfrak{g}_{-p.\alpha}) = 0$ hence $\dim(\mathfrak{g}_{-p.\alpha}) = 0$ for all $p > 1$ so that $\dim(\mathfrak{g}_{-\alpha}) = 1$. Since $\dim(\mathfrak{g}_\lambda) = \dim(\mathfrak{g}_{-\lambda})$ it follows $\dim(\mathfrak{g}_{c.\alpha}) = 0$ if $|c| > 1$ and $\dim(\mathfrak{g}_\alpha) = \dim(\mathfrak{g}_{-\alpha}) = 1$, which proves part *i*). It follows that $\mathfrak{sl}_\alpha$ is three-dimensional with basis $\{e_\alpha, h_\alpha, f_\alpha\}$. But $[e_\alpha, f_\alpha] = h_\alpha$ by our choice of f_α, and as $\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}$ are 1-dimensional, $\mathfrak{g}_{\pm\alpha} \cong \mathsf{k}_{\pm\alpha}$, and hence as $\alpha(h_\alpha) = 2$ we have $[h_\alpha, e_\alpha] = 2e_\alpha$, $[h_\alpha, f_\alpha] = -2f_\alpha$. Thus $\{e_\alpha, h_\alpha, f_\alpha\}$ is an $\mathfrak{sl}_2$-triple, so that $\mathfrak{sl}_\alpha \cong \mathfrak{sl}_2$, establishing part *ii*).

Finally, since $\mathfrak{h}$ is abelian $\mathfrak{h} \cong \mathsf{k}_0^{\dim(\mathfrak{h})}$ thus the Cartan decomposition $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha\in\Phi} \mathfrak{g}_\alpha$ exhibits $\mathfrak{g}$ as a direct sum of irreducible $\mathfrak{h}$-representations, hence $\mathrm{ad}(h)$ acts diagonalisably on $\mathfrak{g}$ for all $h \in \mathfrak{h}$, which proves part *iii*). □

We can use our strategy of evaluating traces in two ways once more to obtain finer information about the set of roots associated to the Cartan decomposition of a semisimple Lie algebra. For this we need the some more terminology:

Definition 7.1.9. Suppose that α, β are two roots in $\mathfrak{g}$. Then we may consider the roots which have the form $\beta + k\alpha$ for some integer $k \in \mathbb{Z}$. Clearly, since $\mathfrak{g}$ is finite dimensional, there are integers $p, q > 0$ such that $\beta + k\alpha \in \Phi_0$ for each k with $-p \leq k \leq q$, but neither $\beta - (p+1)\alpha$ nor $\beta + (q+1)\alpha$ are in $\Phi_0 = \Phi$. We call this set of roots[1] $S_\alpha(\beta)$ the *α-string* through β.

Proposition 7.1.10. *Let $\alpha, \beta \in \Phi$ and suppose that $S_\alpha(\beta) = \{\beta - p\alpha, \dots, \beta + q\alpha\}$ is the α-string through β. Then we have*

i)

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

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

ii) If $\beta = c.\alpha$ for some $c \in \mathsf{k}$, then $c = \pm 1$.

iii) If $S = \{\beta + k\alpha \in \Phi \cup \{0\} : k \in \mathbb{Z}\}$, then $S = S_\alpha(\beta)$.

Proof. First note that if $0 \in S_\alpha(\beta)$ then $\beta = k.\alpha$ for some $k \in \mathbb{Z}$, and the claims of this proposition in that case all follow from Proposition 7.1.8. Thus we may assume $\beta \notin \mathbb{Z}.\alpha$ so that $S_\alpha(\beta) \subseteq \Phi$. Let $\gamma = \beta - p\alpha$, so that the $S_\alpha(\beta) S_\alpha(\gamma) = \{\gamma, \gamma + \alpha, \dots, \gamma + s\alpha\}$, where $s = p + q$. Consider the subspace $M_\beta = \bigoplus_{0\leq k\leq s} \mathfrak{g}_{\gamma+k\alpha}$, and let $\{e_\alpha, h_\alpha, e_{-\alpha}\}$ be an $\mathfrak{sl}_2$-triple corresponding to an isomorphism $\varphi\colon \mathfrak{sl}_2 \to \mathfrak{sl}_\alpha$ as in Remark 7.1.4. Since $\gamma - \alpha \notin \Phi \cup \{0\}$ and $\gamma + (s+1)\alpha \notin \Phi \cup \{0\}$, we see that $\mathfrak{sl}_\alpha$ preserves M_β. Hence $\mathrm{tr}_{M_\beta}(\mathrm{ad}(h_\alpha)_{|M_\beta}) = \mathrm{tr}_M([\mathrm{ad}(e_\alpha)_{|M_\beta}, \mathrm{ad}(e_{-\alpha})_{|M_\beta}]) = 0$. Let $e\colon \mathfrak{gl}_1 \to \mathfrak{h}$ be the map $e(t) = t.h_\alpha$, so that $e^\top(\lambda) = \lambda(h_\alpha)$ is the unique eigenvalue of $\mathrm{ad}(h_\alpha)_{|\mathfrak{g}_\alpha}$ where $e^\top : \mathfrak{h}^* \to \mathfrak{gl}_1{}^* = \mathfrak{gl}_1 = \mathsf{k}$ denotes the transpose of e.

But the weight spaces of M_β are $\mathfrak{g}_{\gamma+k\alpha}$ where $0 \leq k \leq s$, which are thus 1-dimensional eigenspaces for $\mathrm{ad}(h_\alpha)$ with eigenvalue $e^\top(\gamma + k\alpha) = \gamma(h_\alpha) + 2k$. Thus the spectrum of $\mathrm{ad}(h_\alpha)$ is an arithmetic progression $\{\gamma(h_\alpha) + 2k : 0 \leq k \leq s\}$ with each eigenvalue having multiplicity one. But then since $\mathrm{tr}(\mathrm{ad}(h_\alpha) = 0$, their mean value is 0, so $0 = \gamma(h_\alpha) + \frac{2}{s+1}\sum_{k=0}^{s} = \gamma(h_\alpha) + s$, and so

$$e^\top(S_\alpha(\beta)) = \{-s, -s+2, \dots, s-2, s\}. \tag{7.1.2}$$

Since $s = p + q$, it follows that $\beta(h_\alpha) = (\gamma + p\alpha)(h_\alpha) = p - q$ as required, and $\beta - (p-q)\alpha = \gamma + q.\alpha \in S_\alpha(\beta) \subseteq \Phi$.

For part *ii*), note $e^\top$ restricts to an isomorphism $\mathsf{k}.\alpha \to \mathsf{k}$, with $e^\top(a\alpha) = 2a$. Thus if $\beta = c.\alpha$, since $\beta \notin \mathbb{Z}.\alpha$ by assumption, $0 \notin e^\top(S_\alpha(\beta))$ so by (7.1.2), s must be odd, where $s + 1 = |S_\alpha(c.\alpha)|$. But then, again by (7.1.2),

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

$1 \in \{-s, -s+2, \dots s\}$ and since $1 = e^{\top}(\alpha/2)$, it follows $\alpha/2 \in S_\alpha(\beta) \subseteq \Phi$. But then $\alpha = 2(\alpha/2) \in 2.\Phi$, which is a contradiction. Hence if $\beta \notin \mathbb{Z}.\alpha$ we must have $\beta \notin \mathsf{k}.\alpha$.

Finally, note that S is clearly the disjoint union of the α-root strings it contains, and by the above, if $S_\alpha(\beta')$ is any such string, we may form the corresponding $\mathfrak{sl}_\alpha$-subrepresentation $M_{\beta'}$ of $\mathfrak{g}$. But the eigenvalues of $\mathrm{ad}(h_\alpha)$ on $M_{\beta'}$ must be $e^{\top}(S_\alpha(\beta')) = P_s = \{-s, -s+2, \dots, s-2, s\}$ where $s = |S_\alpha(\beta')|$. Since for $s \leq s'$ we have $P_s \subseteq P_{s'}$, any two such sets intersect. Since the map $\beta + k\alpha \mapsto \beta(h_\alpha) + 2k$ is injective and the roots strings in S are pairwise disjoint, it follows $S = S_\alpha(\beta)$ as claimed.

□

7.1.2 Rational form of $\mathfrak{h}$ and inner product spaces

Recall that since $\kappa_{|\mathfrak{h}}$ is non-degenerate, it gives an isomorphism $\theta \colon \mathfrak{h} \to \mathfrak{h}^*$. For $\lambda \in \mathfrak{h}^*$, we write t_λ for $\theta^{-1}(\lambda)$, 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 $\mathfrak{sl}_\alpha = \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subseteq \mathfrak{g}$ is a subalgebra isomorphic to $\mathfrak{sl}_2$. Indeed each summand is 1-dimensional, and $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = \mathsf{k}.t_\alpha \subseteq \mathfrak{h}$. We set $h_\alpha = 2\alpha(t_\alpha)^{-1}.t_\alpha$, so that $\alpha(h_\alpha) = 2$.

Definition 7.1.11. Let $\mathfrak{h}_\mathbb{Q} = \{h \in \mathfrak{h} : \alpha(h) \in \mathbb{Q}, \forall \alpha \in \Phi\}$. Clearly $\mathfrak{h}_\mathbb{Q}$ is a $\mathbb{Q}$-vector space (a subspace of $\mathfrak{h}$ viewed as a $\mathbb{Q}$-vector space[2]). Recall also that a symmetric bilinear form $(.,.)$ on a $\mathbb{Q}$-vector space V is said to be *positive definite* if $(v, v) \geq 0$ with equality holding precisely when $v = 0$. Such a form is also commonly called an *inner product*.

Lemma 7.1.12. *The $\mathbb{Q}$-vector space $\mathfrak{h}_\mathbb{Q}$ has the following properties*

i) If $\kappa_\mathbb{Q}$ denotes the restriction of the Killing form to $\mathfrak{h}_\mathbb{Q}$, then $\kappa_\mathbb{Q}$ is a $\mathbb{Q}$-valued positive definite symmetric bilinear form on $\mathfrak{h}_\mathbb{Q}$.

ii) $\mathfrak{h}_\mathbb{Q} = \langle \Phi^\vee \rangle_\mathbb{Q} = \langle \{t_\alpha : \alpha \in \Phi\} \rangle_\mathbb{Q}$, and $\dim_\mathbb{Q}(\mathfrak{h}_\mathbb{Q}) = \dim_\mathsf{k}(\mathfrak{h})$. In particular, θ identifies the dual $(\mathfrak{h}_\mathbb{Q})^*$ of $\mathfrak{h}_\mathbb{Q}$ with $\langle \Phi \rangle_\mathbb{Q}$.

Proof. For part *i*), since we have shown in Proposition 7.1.8 that $\dim(\mathfrak{g}_\alpha) = 1$ for all $\alpha \in \Phi$, we may simplify the expression for $\kappa_{|\mathfrak{h}}$ given in (7.1.1) to obtain $\kappa(h_1, h_2) = \sum_{\gamma \in \Phi} \gamma(h_1)\gamma(h_2)$. It is thus immediate from the definition of $\mathfrak{h}_\mathbb{Q}$ that the Killing form is $\mathbb{Q}$-valued on $\mathfrak{h}_\mathbb{Q}$. Moreover, if $h \in \mathfrak{h}_\mathbb{Q}$, since $\gamma(h)^2 \geq 0$ for all $\gamma \in \Phi$, we have $\kappa_\mathbb{Q}(h, h) \geq 0$, with equality if and only if $\gamma(h) = 0$ for all $\gamma \in \Phi$, and since $\langle \Phi \rangle = \mathfrak{h}^*$, this holds only if $h = 0$. Thus $\kappa_\mathbb{Q}$ is positive definite as claimed.

For part *ii*), by Proposition 7.1.10, the roots are $\mathbb{Z}$-valued on set of coroots $\Phi^\vee$, hence $\mathfrak{h}_\mathbb{Q} \supseteq \langle \Phi^\vee \rangle_\mathbb{Q}$. Moreover $h_\alpha = c_\alpha.t_\alpha$ where $2 = c_\alpha.\kappa(t_\alpha, t_\alpha)$, hence $\kappa(h_\alpha, h_\alpha) = c_\alpha^2 \kappa(t_\alpha, t_\alpha) = 2.c_\alpha \in \mathbb{Z}$. Thus $t_\alpha = 2\kappa(h_\alpha, h_\alpha)^{-1}.h_\alpha \in \mathbb{Q}.h_\alpha$, and hence clearly $\langle \{t_\alpha : \alpha \in \Phi\} \rangle_\mathbb{Q} = \langle \Phi^\vee \rangle_\mathbb{Q}$.

Now since Φ spans $\mathfrak{h}^*$, we may find a subset $B = \{\gamma_1, \dots, \gamma_l\} \subseteq \Phi$ which is a basis of $\mathfrak{h}^*$. Let $B' = \theta_0^{-1}(B) = \{t_{\gamma_i} : 1 \leq i \leq l\} \subseteq \langle \Phi^\vee \rangle_\mathbb{Q}$. Since it is a k-basis of $\mathfrak{h}$, B' is linearly independent over k and hence over $\mathbb{Q}$. It follows $\dim_\mathbb{Q}(\langle \Phi^\vee \rangle_\mathbb{Q}) \geq l$. But if $\eta \colon \mathfrak{h}_\mathbb{Q} \to \mathbb{Q}^l$ is given by $\eta(h) = (\gamma_i(h))_{i=1}^l$, then η is injective because B is a basis of $\mathfrak{h}^*$. Hence $\dim(\mathfrak{h}_\mathbb{Q}) \leq l$. It follows that $\mathfrak{h}_\mathbb{Q} = \langle \Phi^\vee \rangle_\mathbb{Q}$ and $\dim_\mathbb{Q}(\mathfrak{h}_\mathbb{Q}) = l = \dim_\mathsf{k}(\mathfrak{h})$. Since $\mathrm{span}\{t_\alpha : \alpha \in \Phi\} = \mathfrak{h}_\mathbb{Q}$, the final statement in *ii*) is now also clear. □

Definition 7.1.13. Let $(-,-)$ denote the bilinear form on $\mathfrak{h}^*$ which is obtained by identifying $\mathfrak{h}^*$ with $\mathfrak{h}$: that is

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

Clearly it is a nondegenerate symmetric bilinear form, and via the previous Lemma, for all $\alpha, \beta \in \Phi$ we have $(\alpha, \beta) = \kappa(t_\alpha, t_\beta) \in \mathbb{Q}$, so that it restricts to a $\mathbb{Q}$-valued symmetric bilinear form on $\mathfrak{h}_\mathbb{Q}^*$ which is positive definite.

***Remark 7.1.14.** The group Y generated by $\{h_\alpha : \alpha \in \Phi\}$ is a finitely generated abelian group which is a subgroup of a $\mathbb{Q}$-vector space, and so is torsion-free. It follows from the structure theorem for finitely generated abelian groups[3] that Y is therefore actually a free abelian group. Moreover, the inner product restricts to an *integer-valued* positive definite form on Y. A finitely generated free abelian group with such a form is called a *lattice*. (See the final problem sheet for some discussion of these.) Note that any basis B for Y is also a $\mathbb{Q}$-basis of $\mathfrak{h}_\mathbb{Q}$ but not conversely – this gives at least some motivation for the notion of a base we will see shortly, in that some subsets of Φ may yield only a $\mathbb{Q}$-basis of $\mathfrak{h}_\mathbb{Q}$, whereas others may yield a $\mathbb{Z}$-basis of Y.

[2] Since k has characteristic zero, it contains a canonical copy of $\mathbb{Q}$ – it is the intersection of all of the subfields of k)

[3] which you may have seen in a previous algebra course...

7.2 Abstract root systems

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

We will work with rational inner product spaces V, that is, $\mathbb{Q}$-vector spaces V equipped with a positive-definite symmetric bilinear form[4] which we will denote by $(.,.)$. Such vector spaces have, in addition to a notion of length given by, for any $v \in V$, the norm $||v|| = (v,v)^{1/2}$, a notion of angle: by the Cauchy-Schwarz inequality there is a unique $\theta \in [0,\pi]$ with

$$\cos(\theta) = \frac{(v_1,v_2)}{||v_1||.||v_2||} \in [-1,1].$$

The group of orthogonal linear transformations of V is

$$\mathrm{O}(V) = \{g \in \mathrm{GL}(V) : (g(v), g(w)) = (v,w), \forall\, v,w \in V\}.$$

Definition 7.2.1. 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 7.2.2. A pair (V, Φ) consisting a rational inner product space V and a finite subset $\Phi \subset V\backslash\{0\}$ is called an *(abstract) root system* if it satisfies the following properties:

i) Φ spans V;

ii) If $\alpha \in \Phi$, $c \in \mathbb{Q}$, then $c\alpha \in \Phi$ if and only if $c = \pm 1$;

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

iv) If $\alpha, \beta \in \Phi$ and we define

$$\langle \alpha, \beta \rangle = \frac{2(\alpha,\beta)}{(\alpha,\alpha)}, \tag{7.2.1}$$

then $\langle \alpha, \beta \rangle \in \mathbb{Z}$. We say $\langle \alpha, \beta \rangle$ is a *Cartan integer*.

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

Lemma 7.2.3. *Let $(\mathfrak{g}, \mathfrak{h})$ be a Cartan pair where $\mathfrak{g}$ is semisimple, and let $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha\in\Phi} \mathfrak{g}_\alpha$ be the associated Cartan decomposition and let $\mathfrak{h}^*_{\mathbb{Q}}$ be the $\mathbb{Q}$-span of Φ in $\mathfrak{h}^*$. Then $(\mathfrak{h}^*_{\mathbb{Q}}, \Phi)$ is an abstract root system.*

Proof. Let $(-,-)$ denote the symmetric bilinear form on $\mathfrak{h}^*_{\mathbb{Q}}$ induced by the restriction of the Killing form $\kappa_{|\mathfrak{h}}$ as in §7.1.2. Lemma 7.1.12 shows that this restriction is positive definite. Property *i*) for an abstract root system follows immediately from the definitions, and the remaining properties follow from Proposition 7.1.10: part *ii*) of that Proposition establishes property *ii*), while part *i*) establishes properties *iii*) and *iv*).

□

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

[4] Such forms only make sense over *ordered* fields, such as $\mathbb{Q}$ or $\mathbb{R}$.

Definition 7.2.4. 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 $O(V)$.

Example 7.2.5. 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 $\mathfrak{gl}_n$. The Cartan decomposition of $\mathfrak{sl}_n$ is $\mathfrak{h} \oplus \bigoplus_{1 \leq i \neq j \leq n} \mathsf{k}.E_{ij}$, where $\mathrm{ad}(h)(E_{ij}) = (h_i - h_j)E_{ij}$, where $h_k = \epsilon_k(h)$ for $k \in \{1, \dots, n\}$. Thus

$$\mathfrak{h}_{\mathbb{Q}} = \{\sum_{i=1}^n h_i E_{ii} : h_i \in \mathbb{Q}, \sum_{i=1}^n h_i = 0\} \quad \text{and} \quad \mathfrak{h}_{\mathbb{Q}}^* = \left\{\sum_{i=1}^n c_i \varepsilon_i : c_i \in \mathbb{Q}\right\} \Big/ \mathbb{Q}.(\varepsilon_1 + \dots + \varepsilon_n),$$

where the roots in $\mathfrak{h}_{\mathbb{Q}}^*$ are the (images of the) vectors $\{\varepsilon_i - \varepsilon_j : 1 \leq i, j \leq n, i \neq j\}$. Moreover, the Killing form for $\mathfrak{sl}_n$ is $\kappa(x, y) = 2n.\mathrm{tr}(xy)$, so the $\{E_{ii} : 1 \leq i \leq n\}$ are an orthogonal basis of $\mathfrak{h}$ with $\kappa(E_{ii}, E_{ii}) = 2n$. 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.

7.2.1 Positive sets and sets of simple roots

Since the set of roots Φ spans V, it certainly contains (many) subsets which form a basis of V. The key to the classification of root systems is to show that there is a special class of such bases which capture enough of the geometry of the set of roots that the entire root system can be recovered from the bases of this form.

Definition 7.2.6. Given a set of vectors X in a vector space V, we will write

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

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

Definition 7.2.7. Let (V, Φ) be a root system, and let Δ be a subset of Φ. We say that Δ is a *base* (or a set of *simple roots*) for Φ if Δ is a linearly independent and for each $\alpha \in \Phi$, exactly one of α or $-\alpha$ lies in $\mathbb{N}.\Delta$. Note that since $\langle \Phi \rangle_{\mathbb{Q}} = V$, a base Δ is in particular a basis of V. Given a base Δ of Φ we set $\Phi_\Delta^+ = \mathbb{N}\Delta \cap \Phi$ and $\Phi_\Delta^- = -\Phi_\Delta^+$, the subsets of Δ*-positive* and Δ*-negative* roots respectively.

Remark 7.2.8. One can express the conditions that a subset $\Delta \subseteq \Phi$ must satisfy to be a base in various equivalent ways. For example, one can rephrase them as follows: *i)* $\Phi \subseteq \mathbb{Z}.\Delta$ and *ii)* $\mathbb{N}.\Delta \cap -\mathbb{N}.\Delta = \emptyset$ and $\Phi \cap \mathbb{N}.\Delta$ is as large as possible, *i.e.* $|\Phi \cap \mathbb{N}.\Delta| = |\Phi|/2$.

The second condition is perhaps less natural-seeming, but it is helpful to note that if it holds, and we write $\mathfrak{n}^\pm = \bigoplus_{\alpha \in \Phi_\Delta^\pm} \mathfrak{g}_\alpha$, then $\mathfrak{n}^+$ and $\mathfrak{n}^-$ are subalgebras of $\mathfrak{g}$ and in fact $\mathfrak{g} = \mathfrak{n}^+ \oplus \mathfrak{h} \oplus \mathfrak{n}^-$ (a direct sum as a vector space – each summand is only a subalgebra, not an ideal, of $\mathfrak{g}$). Moreover, the Killing form induces an isomorphism $\theta^+ : \mathfrak{n}^+ \to (\mathfrak{n}^-)^*$.

The following definition gives us a natural way of decomposing the roots Φ into "positive" and "negative" subsets.

Definition 7.2.9. Let V be a $\mathbb{Q}$-vector space. A *positive set* $\mathscr{P}$ in V is a subset $\mathscr{P} \subseteq V \backslash \{0\}$ such that

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

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

Example 7.2.10. If V is a $\mathbb{Q}$-vector space and $\vec{B} = (e_1, \dots, e_n)$ is an ordered basis of V, and let $\{\delta_1, \dots, \delta_l\}$ be the dual basis of V^* so that for any $v \in V$ we have $v = \sum_{i=1}^l \delta_i(v).e_i$. Let $s \colon V\backslash\{0\} \to \{1, \dots, l\}$ be given by $s_{\vec{B}}(v) = \min\{k : 1 \leq k \leq l, \delta_k(v) \neq 0\}$, and, if $v \in V\backslash\{0\}$ has $s_{\vec{B}}(v) = k$, let $p_{\vec{B}}(v) = \delta_k(v)$. Then

$$\mathscr{P}(\vec{B}) = \{v \in V\backslash\{0\} : p_{\vec{B}}(v) > 0\} = \left\{\sum_{i=k}^l \lambda_i e_i : k \in \{1, 2, \dots, l\}, \lambda_k > 0\right\}$$

is a positive system.

Definition 7.2.11. Let (V, Φ) be a root system and fix a positive set $\mathscr{P}$ in V. Let $\Phi^+_{\mathscr{P}} = \Phi \cap \mathscr{P}$ and $\Phi^-_{\mathscr{P}} = \Phi \cap (-\mathscr{P})$ (where the positive set $\mathscr{P}$ is understood from context, we will simply write Φ^+). We say that Φ^+ is a set of *positive roots* if there is a positive set $\mathscr{P}$ in V for which $\Phi^+ = \Phi^+_{\mathscr{P}} = \Phi \cap \mathscr{P}$.

Given a set of positive roots $\Phi^+_{\mathscr{P}}$, we say that $\alpha \in \Phi^+_{\mathscr{P}}$ is *decomposable* if $\alpha = \beta + \gamma$ for some $\beta, \gamma \in \Phi^+_{\mathscr{P}}$. A root is *indecomposable* if it is not decomposable. Let $\Pi_{\mathscr{P}}$ be the set of indecomposable roots in $\Phi^+_{\mathscr{P}}$.

Proposition 7.2.12. *Let (V, Φ) be an abstract root system.*

i) Let (V, Φ) be a root system and suppose that $\mathscr{P}$ is a positive system. Then the set of indecomposable roots $\Pi = \Pi_{\mathscr{P}}$ is linearly independent and $\Phi^+_{\mathscr{P}} = \mathbb{N}.\Pi_{\mathscr{P}} \cap \Phi$, so that Π is a base of (V, Φ) and the map $\Phi^+_{\mathscr{P}} \mapsto \Pi_{\mathscr{P}}$ is bijective with inverse $\Pi_{\mathscr{P}} \mapsto \mathbb{N}.\Pi_{\mathscr{P}} \cap \Phi$. In particular, any root system has a base.

ii) Every base of (V, Φ) is of the form $\Pi_{\mathscr{P}}$ for some positive system $\mathscr{P}$, thus the map $\Delta \to \Phi^+_\Delta$ gives a bijection between set of all bases of (V, Φ) and the set of all sets of positive roots $\Phi^+_{\mathscr{P}}$ since for any positive system $\mathscr{P}$ with $\Phi_{\mathscr{P}} = \Phi^+_\Delta$ we have $\Pi_{\mathscr{P}} = \Delta$.

Proof. We establish part *i)* in three steps:

Step 1: We claim that $\Phi^+_{\mathscr{P}} \subseteq \mathbb{N}.\Pi$. The set Φ^+ is totally ordered by the order given by $\mathscr{P}$, so if Φ^+ is not contained in $\mathbb{N}.\Pi$, we may consider $\alpha \in \Phi^+$, the minimal element of Φ^+ not contained on $\mathbb{N}.\Pi$. Clearly $\alpha \notin \Pi$, hence α is decomposable and there are $\gamma_1, \gamma_2 \in \Phi^+$ with $\alpha = \gamma_1 + \gamma_2$. But then $\gamma_1, \gamma_2 < \alpha$, so by minimality of α, $\gamma_1, \gamma_2 \in \mathbb{N}.\Pi$. But $\mathbb{N}.\Pi$ is closed under addition, so $\alpha = \gamma_1 + \gamma_2 \in \mathbb{N}.\Pi$, giving a contradiction. Hence $\Phi^+ \subseteq \mathbb{N}.\Pi$ as required, and $\Phi^+_{\mathscr{P}} \mapsto \Pi_{\mathscr{P}}$ is bijective.

Step 2: Next we claim that if $\alpha, \beta \in \Phi^+_{\mathscr{P}}$ are distinct roots such that $(\alpha, \beta) > 0$, then at least one of them must be decomposable. To see this, first note that $\langle \alpha, \beta \rangle, \langle \beta, \alpha \rangle > 0$. Since $\alpha, \beta \in \Phi^+$ are distinct, they are linearly independent, and hence by Cauchy-Schwarz, $\langle \alpha, \beta \rangle . \langle \beta, \alpha \rangle \in \{1, 2, 3\}$. It follows one of $\langle \alpha, \beta \rangle$ or $\langle \beta, \alpha \rangle = 1$. By symmetry we may assume $\langle \alpha, \beta \rangle = 1$, and hence $s_\alpha(\beta) = \beta - \alpha \in \Phi$. But then one of $\alpha - \beta$ or $\beta - \alpha$ lies in Φ^+. But as $\alpha = (\alpha - \beta) + \beta$ and $\beta = (\beta - \alpha) + \alpha$, it follows that one of α or β must be is decomposable as required.

Step 3 Let S be any subset of $\mathscr{P}$ with the property that if $s, t \in S$ are distinct, then $(s, t) \leq 0$. We claim that S is linearly independent. Since by step 2, $\Pi_{\mathscr{P}}$ has this property, and by step 1 $\Phi^+_{\mathscr{P}} \subseteq \mathbb{N}.\Pi_{\mathscr{P}}$, it follows from the claim that $\Pi_{\mathscr{P}}$ is a base. Suppose that $\sum_{s \in T} c_s.s = 0$ is a linear dependence, where $T \subseteq S$. Then let $T^\pm = \{s \in T : \pm c_s > 0\}$, and let $z = \sum_{t \in T^+} c_t.t = \sum_{s \in T^-} (-c_s).s$. Now $\sum_{t \in T^+} c_t.t \in \mathscr{P}$ unless $T^+ = \emptyset$ since $T^+ \subseteq \mathscr{P}$ and $c_t > 0$ for all $t \in T^+$. But

$$(z, z) = (\sum_{t \in T^+} c_t.t, \sum_{s \in T^-} (-c_s).s) = \sum_{s,t} c_t(-c_s)(s, t) \leq 0$$

hence $z = 0$. But then $T^+ = \emptyset$, and $\sum_{s \in T^-} (-c_s).s = 0$. Since $(-c_s) \geq 0$ and $s \in \mathscr{P}$ for all $s \in T^-$ this implies $c_s = 0$ for all $s \in T^- = T$, so S is linearly independent as required.

For part *ii)*, given a base Δ, pick an arbitrary ordering $\Delta^{ord} = \{\alpha_1, \dots, \alpha_l\}$. Then Δ^{ord} is an ordered basis of V and hence gives a positive set $\mathscr{P} = \mathscr{P}_{\Delta^{ord}}$ as in Example 7.2.10. Now $\Phi^+_\Delta = \mathbb{N}.\Delta \cap \Phi \subseteq \mathscr{P}$ since $\Delta \subseteq \mathscr{P}$, and hence $\Phi^+_\Delta \subseteq \Phi \cap \mathscr{P} = \Phi^+_{\mathscr{P}}$, and since Φ^+_Δ and $\Phi^+_{\mathscr{P}}$ both have $\frac{1}{2}|\Phi|$ elements, $\Phi^+_\Delta = \Phi^+_{\mathscr{P}}$ (thus the set of positive roots $\Phi^+_{\mathscr{P}}$ is independent of ordering of Δ which we chose to obtain $\mathscr{P}$).

Let $\rho \in V^*$ be given by $\rho(\alpha) = 1$ for all $\alpha \in \Delta$. If $\mathscr{P}$ is any positive set for which $\Phi^+_{\mathscr{P}} = \Phi^+_\Delta = \mathbb{N}.\Delta \cap \Phi$ we see that $\delta(\alpha) \in \mathbb{N}$ for all $\alpha \in \Phi^+_{\mathscr{P}}$ so that if α is decomposable we must have $\delta(\alpha) \geq 2$, while $\delta(\alpha) = 1$ if and only if $\alpha \in \Delta$. It follows that $\Delta \subseteq \Pi_{\mathscr{P}}$. Since we have just seen that $\Pi_{\mathscr{P}}$ is a base of V it follows $\Pi_{\mathscr{P}} = \Delta$. Part *ii)* follows immediately.

□

Lemma 7.2.13. *Let (V, Φ) be an abstract root system, and suppose that $\alpha \in \Phi$.*

1. *There is a positive set $\mathscr{P}$ for which α is the minimal element of $\Phi^+_{\mathscr{P}}$, In particular, $\alpha \in \Phi^+_{\mathscr{P}}$ is indecomposable, and so belongs to $\Pi_{\mathscr{P}}$, a base of (V, Φ).*
2. *If Δ is a base and $\alpha \in \Delta$, then if $\beta \in \Phi^+_\Delta$ and $\beta \neq \alpha$, then $s_\alpha(\beta) \in \Phi^+_{\mathscr{P}}$. Hence if we write $\Phi^+(\alpha) = \Phi^+ \backslash \{\alpha\}$ then we have $s_\alpha(\Phi^+(\alpha)) = \Phi^+(\alpha)$ whenever $\alpha \in \Delta$.*

Proof. Pick a basis $B \subset \Phi$ of V containing α and let $\vec{B} = \{\gamma_1, \dots, \gamma_l\}$ be an ordering of it in which $\alpha = \gamma_l$. Let $D = \{\delta_1, \dots, \delta_l\}$ Let $\mathscr{P} = \mathscr{P}_{\vec{B}}$. Then if $\beta \in \Phi^+_{\mathscr{P}}$ has $\beta \leq \alpha$, since $\beta \in \mathscr{P}$ and $s_{\vec{B}} = l$, we must have $s_{\vec{B}}(\beta) = l$ also, and hence $\beta = c.\alpha$ for some $c > 0$. But since $\alpha, \beta \in \Phi$ this implies $\beta = \alpha$ as required. Since β is minimal, it must be indecomposable, and hence $\beta \in \Pi_{\mathscr{P}}$ which we have seen is always a base.

Next if Δ is a base containing α, we may take $B = \Delta$, so that $\mathscr{P} = \mathscr{P}_{\vec{\Delta}}$. If $\vec{\Delta_1} = s_\alpha(\vec{\Delta})$, then $\mathscr{P}_{\vec{\Delta}_1} = s_\alpha(\mathscr{P})$. But $s_\alpha(\gamma_i) = \gamma_i - \langle \gamma_l, \gamma_i \rangle.\gamma_l$, hence the dual basis to $\Delta_1 = s_\alpha(\Delta)$ is $D_1 = s_\alpha^\top(D) = \{\delta_1, \ldots, \delta_{l-1}, -\delta_l - \sum_{i<l}\langle \gamma_l, \gamma_i\rangle \delta_i\}$. It follows that $p_\Delta(v) = p_{\Delta_1}(v)$ for all $v \in V \backslash \mathbb{Q}.\alpha$, hence $s_\alpha(\mathscr{P}) = \mathscr{P}_{\vec{\Delta}_1}$ and $\mathscr{P} \backslash \mathbb{Q}.\alpha = \mathscr{P}_{\vec{\Delta}_1} \backslash \mathbb{Q}.\alpha = s_\alpha(\mathscr{P}) \backslash \mathbb{Q}.\alpha$, so that $(\Phi \backslash \{\pm\alpha\}) \cap \mathscr{P} = (\Phi \backslash \{\pm \alpha\} \cap \mathscr{P}_1$ as required.

□

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

Proposition 7.2.14. *Suppose that Δ_1 is any base of (V, Φ) and let Φ_1^+ be the corresponding set of positive roots. Then there is some $w \in W_0$ such that $w(\Phi_1^+) = \Phi_0^+$, and hence $w(\Delta_1) = \Delta_0$.*

Proof. We prove this by induction on $d = |\Phi_0^+ \cap \Phi_1^-|$. If this $d = 0$ then $\Phi_0^+ = \Phi_1^+$ and hence $\Delta_0 = \Delta_1$ (hence we may take $w = e$ the identity element of W_0). Next suppose that $d > 0$. Let $\mathscr{P}_1$ be a positive set such that $\Phi_1^+ = \Phi \cap \mathscr{P}_1$. If $\Delta_0 \subseteq \Phi_1^+$, then since any element of Φ_0^+ is a positive integer combination of Δ_0, it follows $\Phi_0^+ \subseteq \mathscr{P}_1 \cap \Phi = \Phi_1^+$ and hence $\Phi_0^+ \cap \Phi_1^- = \emptyset$, which contradicts the assumption that $d > 0$. Thus there is some $\alpha \in \Delta_0$ such that $\alpha \in \Phi_1^-$. But then using the notation of Lemma 7.2.13 we see that

$$|\Phi_0^+ \cap s_\alpha(\Phi_1^-)| = |s_\alpha(\Phi_0^+) \cap \Phi_1^-| = |\big(\{-\alpha\} \cup \Phi_0^+(\alpha)\big) \cap \Phi_1^-| = |\Phi_0^+(\alpha) \cap \Phi_1^-| = d - 1$$

where the first equality holds because $s_\alpha^2 = 1_V$, the second equality follows from Lemma 7.2.13, and the third from the fact that Φ_1^- contains α and hence not $-\alpha$. But then by induction there is a $w \in W_0$ with $ws_\alpha(\Phi_1)^+ = \Phi_0^+$. Since $ws_\alpha \in W_0$ we are done. □

Corollary 7.2.15. *Suppose that $\beta \in \Phi$. Then there is a $w \in W_0$ and an $\alpha \in \Delta_0$ such that $w(\beta) = \alpha$. In particular, W is generated by the reflections $\{s_\gamma : \gamma \in \Delta_0\}$, that is, $W = W_0$.*

Proof. For the first claim follows from the fact that every root lies in a base for (V, Φ), shown in part i) of Lemma 7.2.13, together with Proposition 7.2.14.

For the final claim, note that if $\beta \in \Phi$ then we have just shown that there is a $w \in W_0$ such that $w(\beta) = \gamma$ for some $\gamma \in \Delta_0$. But then clearly $s_\beta = w^{-1}s_\gamma w \in W_0$, and so since $W = \langle s_\beta : \beta \in \Phi \rangle$ it follows that $W \leq W_0$. Since $W_0 \leq W$ by definition, it follows $W = W_0$ as required.

□

Remark 7.2.16. In fact W acts *simply transitively* on the bases of (V, Φ), that is, the action is transitive and, if Δ is a base and $w \in W$ is such that $w(\Delta) = \Delta$, then $w = 1$. The proof (which we will not give) consists of examining the minimal length expression for w in terms of these generators $\{s_\alpha : \alpha \in \Delta_0\}$.

7.2.2 Cartan matrices and isomorphisms of root systems

First let us formulate the notion of an isomorphism of root systems:

Definition 7.2.17. If (V, Φ) and (V', Φ') are root systems, we say that a linear map $\phi \colon V \to V'$ is an *isomorphism of root systems* if it is a isomorphism of vector spaces such that $\phi(\Phi) = \Phi'$ and

$$\langle \phi(\alpha), \phi(\beta) \rangle = \langle \alpha, \beta \rangle, \quad \forall\, \alpha, \beta \in \Phi.$$

Note that ϕ need *not* be an isometry: if $0 < c < 1$, then $(V, c.\Phi)$ is a root system which is not isometric to (V, Φ), but $\phi(x) = c.x$ is an isomorphism from (V, Φ) to $(V, c.\Phi)$.

Definition 7.2.18. Let (V, Φ) be a root system. The *Cartan matrix* associated to (V, Φ) is the matrix

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

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 the set of bases of Φ, the Cartan matrix is independent[5] of the choice of base (though clearly determined only up to orderings of the base Δ).

Theorem 7.2.19. *Let (V, Φ) be a root system. Then (V, Φ) is determined up to isomorphism by the Cartan matrix 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. 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 $W = W(V, \Phi)$ and $W' = W(V', \Phi')$ 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 by Corollary 7.2.15 we have

$$\phi(\Phi) = \phi(W.\Delta) = (\phi W \phi^{-1})(\phi(\Delta)) = W'\phi(\Delta) = W'.\Delta' = \Phi'$$

Finally, fixing $\alpha \in \Delta$, clearly the linear functionals given by $v \mapsto \langle \alpha, v \rangle$ and $v \mapsto \langle \phi(\alpha), \phi(v) \rangle$ $(v \in V)$ agree if $v \in \Delta$, hence by linearity they are equal. Hence $\langle \alpha, \beta \rangle = \langle \phi(\alpha), \phi(\beta) \rangle$ if $\alpha \in \Delta$ and $\beta \in \Phi$. But since W and W' act by isometries

$$\langle w(\alpha), w(\beta) \rangle = \langle \alpha, \beta \rangle = \langle \phi(\alpha), \phi(\beta) \rangle = \langle w'(\phi(\alpha)), w'(\phi(\beta)) \rangle = \langle \phi(w(\alpha)), \phi(w(\beta)) \rangle,$$

so that since $W.\Delta = \Phi$, it follows that $\langle \alpha, \beta \rangle = \langle \phi(\alpha), \phi(\beta) \rangle$ for all $\alpha, \beta \in \Phi$.

□

Thus to classify root systems up to isomorphism it is enough to classify Cartan matrices. It turns out that there is a more combinatorial way to encode the information given by a Cartan matrix, because the possible entries of the Cartan matrix are heavily constrained, as the next Lemma shows:

Lemma 7.2.20. *Let (V, Φ) be a root system and let $\alpha, \beta \in \Phi$ be such that $\alpha \neq \pm\beta$. Then the Cartan integer $\langle \alpha, \beta \rangle \in \{0, \pm 1, \pm 2, \pm 3\}$ Moreover, the angle between α and β and the ratio $\|\alpha\|^2/\|\beta\|^2$ are determined by the pair $\langle \alpha, \beta \rangle$, $\langle \beta, \alpha \rangle$, as the table below shows:*

$\langle \alpha, \beta \rangle$	$\langle \beta, \alpha \rangle$	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

Proof. By assumption, we know that both $\langle \alpha, \beta \rangle$ and $\langle \beta, \alpha \rangle$ are integers with the same sign. By the Cauchy-Schwarz inequality, if θ denotes the angle between α and β, then:

$$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4\frac{(\alpha, \beta)^2}{\|\alpha\|^2.\|\beta\|^2} = 4\cos(\theta)^2 < 4. \tag{7.2.2}$$

Noting that $\cos(\theta)^2$ determines the angle between the two vectors (or rather the one which is less than π) and (if $(\alpha, \beta) \neq 0$) $\langle \beta, \alpha \rangle / \langle \alpha, \beta \rangle = \|\alpha\|^2/\|\beta\|^2$ (where we write $\|v\|^2 = (v, v)$), it is then easy to verify the table given above by a case-by-case check. □

[5]This might appear to overlook something: While it is true that W acts transitively on the set of bases, so we may pick an arbitrary base in order to compute the Cartan matrix, if $W_\Delta = \{w \in W : w(\Delta) = \Delta\}$ then W_Δ acts on C_Δ, and thus could yield constraints on the possible structure of the Cartan matrix. In fact W_Δ is trivial, so we are not missing anything!

Definition 7.2.21. As the previous Lemma shows, if $C = (c_{ij})$ is a Cartan matrix, its entries c_{ij} are highly constrained: indeed $c_{ii} = 2$ and if $i \neq j$, $c_{ij} \in \{0, -1, -2, -3\}$ and $\{c_{ij}, c_{ji}\} = \{-1, -c_{ij}c_{ji}\}$ so that c_{ij} is determined by the product $c_{ij}.c_{ji}$ and the relative lengths of the two roots (set out in the table above). 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*.

Definition 7.2.22. We say that a root system (V, Φ) is *reducible* if there is a partition of the roots into two non-empty subsets $\Phi_1 \sqcup \Phi_2$ such that $(\alpha, \beta) = 0$ for all $\alpha \in \Phi_1, \beta \in \Phi_2$. Then if we set $V_1 = \text{span}(\Phi_1)$ and $V_2 = \text{span}(\Phi_2)$, clearly $V = V_1 \oplus V_2$ and we say (V, Φ) is the sum of the root systems (V_1, Φ_1) and (V_2, Φ_2). This allows one to reduce the classification of root systems to the classification of *irreducible* root systems, *i,e.* root systems which are not reducible. It is straight-forward to check that a root system is irreducible if and only if its associated Dynkin diagram is connected.

Definition 7.2.23. (*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 $(\ell \geq 1)$:

$$\bullet \text{---} \bullet \cdots \bullet \text{---} \bullet$$

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

- Type B_ℓ $(\ell \geq 2)$:

$$\bullet \text{---} \bullet \cdots \bullet \Longrightarrow \bullet$$

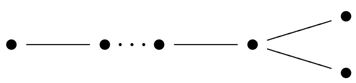
$$V = \mathbb{Q}^\ell, \Phi = \{\pm\varepsilon_i \pm \varepsilon_j : 1 \leq i, j \leq \ell, i \neq j\} \cup \{\pm\varepsilon_i : 1 \leq i \leq \ell\},$$
$$\Delta = \{\varepsilon_1, \varepsilon_{i+1} - \varepsilon_i : 1 \leq i \leq \ell - 1\}$$

- Type C_ℓ $(\ell \geq 3)$:

$$\bullet \text{---} \bullet \cdots \bullet \Longleftarrow \bullet$$

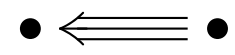
$$V = \mathbb{Q}^\ell, \Phi = \{\pm\varepsilon_i \pm \varepsilon_j : 1 \leq i, j \leq \ell, i \neq j\} \cup \{2\varepsilon_i : 1 \leq i \leq \ell\},$$
$$\Delta = \{2\varepsilon_1, \varepsilon_{i+1} - \varepsilon_i : 1 \leq i \leq \ell - 1\}$$

- Type D_ℓ $(\ell \geq 4)$:



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

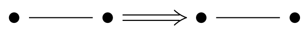
- Type G_2.



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

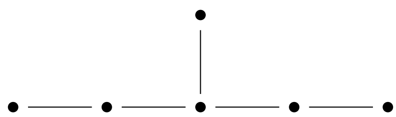
$$\begin{aligned} V &= \{v \in \mathbb{Q}^3 : (v, e) = 0\}, \Phi = \{\varepsilon_i - \varepsilon_j : i \neq j\} \cup \{\pm(3\varepsilon_i - e) : 1 \leq i \leq 3\} \\ \Delta &= \{\varepsilon_1 - \varepsilon_2, e - 3\varepsilon_1\} \end{aligned}$$

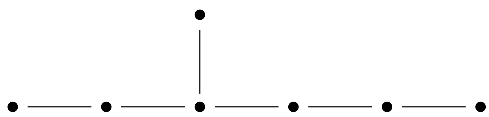
- Type F_4:

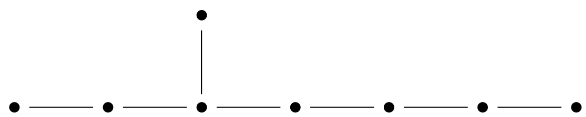


$$\begin{aligned} V &= \mathbb{Q}^4, \\ \Phi &= \{\pm\varepsilon_i : 1 \leq i \leq 4\} \cup \{\pm\varepsilon_i \pm e_j : i \neq j\} \cup \{\frac{1}{2}(\pm\varepsilon_1 \pm \varepsilon_2 \pm \varepsilon_3 \pm \varepsilon_4)\} \\ \Delta &= \{\varepsilon_2 - \varepsilon_3, \varepsilon_3 - \varepsilon_4,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

- Type E_n ($n = 6, 7, 8$).







  These can all be constructed inside E_8 by taking the span of the appropriate subset of a base, so we just give the root system for E_8.

$$V = \mathbb{Q}^8, \Phi = \{\pm\varepsilon_i \pm \varepsilon_j : i \neq j\} \cup \{\frac{1}{2}\sum_{i=1}^{8}(-1)^{a_i}\varepsilon_i : \sum_{i=1}^{8} a_i \in 2\mathbb{Z}\},$$

$$\Delta = \{\varepsilon_1 + \varepsilon_2, \varepsilon_{i+1} - \varepsilon_i, \frac{1}{2}(\varepsilon_1 + \varepsilon_8 - (\varepsilon_2 + \varepsilon_3 + \ldots + \varepsilon_7)) : 1 \leq i \leq 6\}.$$

Note that the Weyl groups of type B_ℓ and C_ℓ are equal. The reason for the restriction on ℓ in the types B,C,D is to avoid repetition, *e.g.* B_2 and C_2 are the same up to relabelling the vertices.

Remark 7.2.24. I certainly don't expect you to remember the root systems of the exceptional types, but you should be familiar with the ones for type A, B, C and D. The ones of rank two (*i.e.* A_2, B_2 and G_2) are also worth knowing (because for example you can draw them!)

7.3 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. By the conjugacy of Cartan subalgebras (see Remark 5.1.4) this gives a well-defined map from simple Lie algebras to irreducible root systems. Then the following theorem shows that its image classifies simple Lie algebras up to isomorphism.

Theorem 7.3.1. *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 $\mathfrak{h}_2$.*

Thus to obtain a classification of simple Lie algebras, it remains to determine which irreducible root systems are associated to a simple Lie algebra. In fact all of them are!

Theorem 7.3.2. *There exists a simple Lie algebra corresponding to each irreducible root system.*

Thus Theorem 7.3.1 says each irreducible root system is associated to at most one simple Lie algebra (up to isomorphism) and so is a kind of *uniqueness theorem*, while Theorem 7.3.2 shows each irreducible root system comes from a simple Lie algebra, so is an *existence theorem*. Theorem 7.3.1 is not difficult given the machinery we have developed in this course. The existence statement is the more substantial result, but we developed enough machinery to see the existence in most cases: the four infinite families A, B, C, D correspond to the classical Lie algebras $\mathfrak{sl}_{\ell+1}$, $\mathfrak{so}_{2\ell}$, $\mathfrak{sp}_{2\ell}$ and $\mathfrak{so}_{2\ell+1}$ – their root systems can be computed directly (as you say in the Problem Sheets). 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. Appendix III gives an outline of this approach to the existence theorem, describing all the necessary constructions, but omitting some of the details of the proofs.

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 $\mathrm{Aut}(\mathfrak{g}) \cap \mathrm{GL}_n(\mathsf{k})$. 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.[6]

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

Appendices

I (Multi)-linear algebra

In this appendix, k *denotes an arbitrary field unless further conditions are explicitly stated.*

I.1 Primary Decomposition

Definition I.1. Let k be an algebraically closed field and V a k-vector space. If $x \in \mathrm{End}_{\mathsf{k}}(V)$ and $\lambda \in \mathsf{k}$, the *generalized eigenspace* for x with eigenvalue λ is

$$V_{\lambda,x} = \{v \in V : \exists n \geq 0, (x-\lambda)^n(v) = 0\},$$

Thus $V_{\lambda,x} \neq \{0\}$ if and only if x has an eigenvector $v \in V \backslash \{0\}$ with eigenvalue λ. The set of eigenvalues of x is called the *spectrum* of x, denoted $\mathrm{Sp}(x) = \{\lambda \in \mathsf{k} : \dim(V_{\lambda,x}) > 0\}$. The subspaces $V_{\lambda,x}$ are clearly invariant under the action of x, that is $x(V_\lambda) \subseteq V_\lambda$.

The following proposition, used in a number of places in this course, is a standard result in Linear Algebra. We provide a proof for the sake of completeness.

Proposition I.2. *Let* k *be an algebraically closed field and V a* k*-vector space, and let $x\colon V \to V$ be a linear map. There is a canonical direct sum decomposition*

$$V = \bigoplus_{\lambda \in \mathsf{k}} V_{\lambda,x},$$

of V into the generalized eigenspaces of x.

Proof. Let $m_x \in \mathsf{k}[t]$ be the minimal polynomial of x. Then if $\phi\colon \mathsf{k}[t] \to \mathrm{End}(V)$ given by $t \mapsto x$ denotes the natural map, we have $\mathsf{k}[t]/(m_x) \cong \mathrm{im}(\phi) \subseteq \mathrm{End}(V)$. If $m_x = \prod_{i=1}^k (t-\lambda_i)^{n_i}$ where the $S(x) = \{\lambda_i : 1 \leq i \leq k\}$ is the spectrum of x, then the Chinese Remainder Theorem and the first isomorphism theorem shows that

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

It follows that we may write $1 \in \mathsf{k}[t]/(m_x)$ as $1 = e_1 + \ldots + e_k$ according to the above decomposition. Now clearly $e_i e_j = 0$ if $i \neq j$ and $e_i^2 = e_i$, so that if $U_i = \mathrm{im}(e_i)$, then we have $V = \bigoplus_{1 \leq i \leq k} U_i$. Moreover, each e_i can be written as polynomials in x by picking any representative in $\mathsf{k}[t]$ of e_i (thought of as an element of $\mathsf{k}[t]/(m_x)$). Note in particular this means that each U_i is invariant under $\mathrm{im}(\phi)$.

Now the characteristic polynomial of $x_{|V_{\lambda_i}}$ is clearly just $(t-\lambda_i)^{d_i}$ where $d_i = \dim(V_{\lambda_i,x})$, and evidently this divides $\chi_x(t)$ the characteristic polynomial of $x \in \mathrm{End}(V)$. But since $V = \bigoplus_{i=1}^k U_i$ we must have $\chi_x(t) = \prod_{i=1}^k (t-\lambda_i)^{m_i}$, where $m_i = \dim(U_i)$ and hence $d_i \leq m_i$. Since $U_i \subseteq V_{\lambda_i,x}$ we also have $m_i \leq d_i$, and hence they must be equal, so $V_{\lambda_i} = U_i$ as required. □

The next Lemma is included for completeness – it readily implies that the coefficients of the characteristic polynomial χ_a of an element $a \in A$ of a subspace $A \subseteq \mathfrak{gl}_V$ are polynomial functions of the coordinates of a given by taking a basis of the subspace. This is used in the proof of the existence of Cartan subalgebras.

Lemma I.3. *Suppose that V and A are finite dimensional vector spaces, $\varphi\colon A \to End(V)$ is a linear map, and $\{a_1, a_2, \dots, a_k\}$ is a basis of A. Let*

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

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

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

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

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

I.2 Tensor Products

I.2.1 Definition and construction

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

Definition I.4. If $V_1, V_2, \dots, V_k$ and U are vector spaces over a field k, let

$$\mathscr{M}(V_1, \dots V_k, U) = \{\theta\colon V_1 \times \dots \times V_k \to U : \theta \text{ is } k\text{-linear}\}$$

be the vector space of all k-(multi-)linear maps on $V_1 \times \dots \times V_k$ taking values in a vector space U. Here we say that a function $\theta\colon V_1 \times V_2 \times \dots \times V_k \to U$ is a k-linear if it is linear in each component separately, that is, if for any k-tuples of vectors $(v_i)_{1\leq i\leq k}, (u_j)_{1\leq j\leq k} \in V_1 \times \dots \times V_k$ and any $\lambda \in \mathsf{k}$, we have for each $i \in \{1, 2, \dots, k\}$,

$$\theta(v_1, \dots, \lambda.v_i + u_i, \dots v_k) = \lambda.\theta(v_1, \dots, v_i, \dots v_k) + \theta(v_1, \dots, u_i, \dots v_k),$$

Pick a basis B_i of V_i for each i $(1 \leq i \leq k)$, and let B_i^* denote the corresponding dual basis of V_i^*. If $b \in B_i$, let δ_b denote the corresponding element of the dual basis B_i^*, so that $B_i^* = \{\delta_b : b \in B_i\}$. Let $\mathbf{B} = B_1 \times B_2 \times \dots \times B_k$

Proposition I.5. *In the notation given above, the restriction to $\mathbf{B}$ gives an isomorphism*

$$r_{\mathbf{B}}\colon \mathscr{M}(V_1, \dots, V_k; U) \to U^{\mathbf{B}} = \{f : \mathbf{B} \to U\}$$

from the space of all k-multilinear maps taking values in U to the space of all U-valued functions on $\mathbf{B}$. Indeed $r_{\mathbf{B}}$ has inverse given explicitly by

$$\mathscr{F}_{\mathbf{B}}(f)(v_1, \dots, v_k) = \sum_{\mathbf{b}=(b_1,\dots,b_k)\in\mathbf{B}} \delta_{b_1}(v_1)\dots\delta_{b_k}(v_k)f(\mathbf{b}).$$

Proof. Note that if we pick $\mathbf{b} = (b_1, \dots, b_k) \in \mathbf{B}$ then the product $\delta_{\mathbf{b}} = \delta_{b_1}.\delta_{b_2}\dots\delta_{b_k}$ is a k-linear map (since multiplication distributes over addition). Since it is easy to see that $\delta_{\mathbf{b}}(\mathbf{b}') = \delta_{\mathbf{b},\mathbf{b}'}$ (that is, is zero unless $\mathbf{b} = \mathbf{b}'$ in which case it is equal to 1), it is immediate that $r_{\mathbf{B}}(\mathscr{F}_{\mathbf{B}}(f)) = f$, so we must show that $\mathscr{F}_{\mathbf{B}}(r_{\mathbf{B}}(\theta)) = \theta$ for any $\theta \in \mathscr{M}(V_1, \dots, V_k; U)$. Explicitly, we must show that

$$\theta = \sum_{\mathbf{b}\in B_1\times\dots\times B_k} \delta_{\mathbf{b}}\theta(\mathbf{b}) \tag{I.1}$$

Indeed applying θ to a k-tuple $\mathbf{b} \in B_1 \times \dots B_k$, we see that the coefficients on the right-hand side are uniquely determined, so it remains to show the products $\delta_{\mathbf{b}}$ of dual basis vectors do indeed span.

The case $k = 1$ is simply the standard argument that the functions $\{\delta_{b_1}\}_{b_1\in B_1}$ are indeed a basis of V_1^*: if $v_1 \in V_1$ then we may write $v_1 = \sum_{b_1\in B_1} \lambda_{b_1} b_1$ for unique scalars $\lambda_{b_1} \in \mathsf{k}$. By the definition of the functions δ_{b_1}, it then follows that $\delta_{b_1}(v_1) = \lambda_{b_1}$, so that $v_1 = \sum_{b_1 \in B_1} \delta_{b_1}(v_1).b_1$. Applying θ gives $\theta(v_1) = \sum_{b_1 \in B_1} \delta_{b_1}(v_1).\theta(b_1)$. But as this holds for all $v_1 \in V_1$, it follows that $\theta = \sum_{b_1 \in B_1} \theta(b_1).\delta_{b_1}$, as required.

The general case then follows by an easy induction: Indeed for any k-tuple of vectors $(v_i)_{1\leq i\leq k}$ with $v_i \in V_i$, using the case $k = 1$, we may write $v_1 = \sum_{b_1 \in B_1} \delta_{b_1}(v_1).b_1$. But then if θ is k-linear we have

$$\theta(v_1, \dots, v_k) = \theta(\sum_{b_1 \in B_1} \delta_{b_1}(v_1).b_1, v_2, \dots, v_k) = \sum_{b_1 \in B_1} \delta_{b_1}(v_1).\theta(b_1, v_2, \dots, v_k).$$

But for each $b_1 \in B_1$, the map $(v_i)_{2\leq i\leq k} \mapsto \theta(b_1, v_2, \dots, v_k)$ is a $(k-1)$-linear map from $V_2 \times \dots V_k$ to k, hence the result follows by induction. □

Remark I.6. Note that, for $k = 1$, this says that a linear map is uniquely determined by its values on a basis of V_1, and the statement should be thought of as saying that a k-linear map is similarly determined "by its values on bases" where the statement of the question gives the precise meaning to the vague phrase in quotation marks.

The previous Proposition gives one way of constructing the tensor product: If V and W are k-vector spaces and we pick bases B_V and B_W of V and W respectively, then by the Proposition, if we set $B = B_V \times B_W$, then for any vector space U, we have

$$\mathscr{M}(V, W; U) \cong U^B \cong \mathrm{Hom}_{\mathsf{k}}(S(B), U), \tag{I.2}$$

where $S(B)$ is the vector space with basis B, that is, the space of finite formal linear combinations of elements of B. The first isomorphism above is a direct consequence of the Proposition where we take $k = 2$ and $V_1 = V, V_2 = W$, while the second is the case $k = 1$ of the proposition with $V_1 = S(B)$. Now taking $U = S(B)$ in (I.2), the identity linear map from $S(B)$ to itself corresponds to a bilinear map $t\colon V \times W \to S(B)$.

Lemma I.7. *The bilinear map $t\colon V \times W \to S(B)$ has the universal property, so that the pair $(S(B), t)$ is the tensor product of V and W.*

Proof. This is essentially established in the previous paragraph: if $\theta\colon V \times W \to U$ is bilinear, then since $\theta_{|B}\colon B \to U$ extends to a linear map $\tilde{\theta}\colon S(B) \to U$. Tracking how the isomorphism of Proposition I.5 identifies θ with the linear map $\tilde{\theta}$ it is easy to see that this can be expressed by means of the bilinear map $t\colon V \times W \to S(B)$ as $\theta = \tilde{\theta} \circ t$. □

Remark I.8. Note that there is a natural isomorphism $\sigma\colon V \otimes W \cong W \otimes V$ given by $v \otimes w \mapsto w \otimes v$. If $V \neq W$, we will normally abuse notation and identify these two spaces and thus write $V \otimes W = W \otimes V$. If $V = W$ however, $\sigma\colon V \otimes V \to V \otimes V$ is an involution on $V \otimes V$, and more generally, $V^{\otimes n} = V \otimes \dots \otimes V$, the tensor product of V with itself n times, has an action of S_n the symmetric group, which permutes the tensor factors: if $\tau \in S_n$ then $\tau(v_1 \otimes \dots \otimes v_n) := v_{\tau(1)} \otimes \dots \otimes \tau_{\tau(n)}$.

Example I.9. If $V = \mathsf{k}$ and W is an arbitrary k-vector space , then if $s\colon \mathsf{k} \times W \to W$ is scalar multiplication map given by $s(\lambda, w) = \lambda.w$, it is clearly bilinear and it is straight-forward to check that it has the universal property so that $\mathsf{k} \otimes W \cong W$.

The following Lemma may help give a better sense for what a "general" element of $V \otimes W$ looks like.

Lemma I.10. *Suppose that V and W are k-vector spaces, and let $x \in V \otimes W$. Then*

i) If $x = \sum_{i=1}^n v_i \otimes w_i$ and $\{w_1, \dots, w_m\}$ is a maximal linearly independent subset[7] of $\{w_1, \dots, w_n\}$, we may write $x = \sum_{i=1}^m v_i' \otimes w_i$. Moreover, if $\{v_1, \dots, v_n\}$ is linearly independent, then $\{v_1', \dots, v_m'\}$ is linearly independent.

ii) Any $x \in V \otimes W$ may be written as a sum $\sum_{i=1}^m v_i \otimes w_i$ where $\{v_1, \dots, v_m\}$ and $\{w_1, \dots, w_m\}$ are linearly independent sets.

Proof. By assumption $\{w_1, \dots, w_m\}$ is a basis of $Y = \langle w_1, \dots, w_n \rangle_{\mathsf{k}}$, so that if $k > m$ then $w_k = \sum_{i=1}^m \lambda_i^k w_i$ for some $\lambda_i^k \in \mathsf{k}$. Hence $\sum_{k=m+1}^n v_k \otimes w_k = \sum_{1\leq i\leq m<k\leq n} \lambda_i^k v_k \otimes w_i$ and

$$x = \sum_{i=1}^n v_i \otimes w_i = \sum_{i=1}^m v_i \otimes w_i + \sum_{1\leq i\leq m<k\leq n} \lambda_i^k v_k \otimes w_i = \sum_{i=1}^m \left(v_i + \sum_{k=m+1}^n \lambda_i^k v_k \right) \otimes w_i.$$

[7]This can always be arranged by permuting the w_i appropriately

Thus if we set $v_i' = v_i + \sum_{k=m+1}^{n} \lambda_i^k v_k \in V$ we have $x = \sum_{i=1}^{m} v_i' \otimes w_i$. Now if $\{v_1, \dots, v_n\}$ are linearly independent, then clearly $\{v_1', \dots, v_m'\}$ will also be linearly independent, hence *i*) follows.

For *ii*) if $x = \sum_{i=1}^{n} v_i \otimes w_i$ as in the statement of *i*), then applying *i*) we obtain $x = \sum_{i=1}^{m} v_i' \otimes w_i$. If $\sigma\colon V \otimes W \to W \otimes V$ is the isomorphism $\sigma(v \otimes w) = w \otimes v$ then $\sigma(x) = \sum_{i=1}^{m} w_i \otimes v_i'$, and applying *i*) to $\sigma(x) \in W \otimes V$ (so that V and W are interchanged) we see that after reordering so that $\{v_1', \dots, v_l'\}$ is a maximal linearly independent subset of $\{v_1', \dots, v_m'\}$, $\sigma(x) = \sum_{i=1}^{l} w_i' \otimes v_i'$, where the set $\{w_1', \dots, w_l'\}$ is also linearly independent. It follows that $x = \sigma^{-1}(\sigma(x)) = \sum_{i=1}^{l} v_i' \otimes w_i'$ gives an expression for x in the required form. □

The construction of the tensor product allows us to replace bilinear maps with linear ones, but one can also relate bilinear maps to "linear maps to spaces of linear maps" – which is really just the process of taking a function of two variables and holding one variable fixed in order to obtain a function of one variable. Formally, we can state:

Lemma I.11. *Let V, W and U be vector spaces over a field* k. *Then we have natural isomorphisms*

$$Hom(V, Hom(W, U)) \cong \mathscr{M}(V, W; U) = Hom(V \otimes W, U) \cong Hom(W, Hom(V, U))$$

In particular, if $U = \mathsf{k}$, this shows that $Hom(V, W^*) \cong \mathscr{M}(V, W; \mathsf{k})$.

Proof. Since there is an obvious identification between $\mathscr{M}(V, W; U) \cong \mathscr{M}(W, V; U)$ it suffices to establish the first isomorphism. But if $\theta \in \mathrm{Hom}(V, \mathrm{Hom}(W, U))$, then we let $\Psi(\theta)(v, w) = \theta(v)(w)$. The fact that θ is linear shows that $\Psi(\theta)$ is linear in V, while the fact that $\theta(v)$ lies in $\mathrm{Hom}(W, U)$ shows that $\Psi(\theta)$ is linear in W. Conversely, if $b \in \mathscr{M}(V, W; U)$, then we may define $\Upsilon(b)(v) = [w \mapsto b(v, w)]$. The map $\Upsilon(b) \in \mathrm{Hom}(W, U)$ because b is linear in W, while the map Υ is linear because b is linear in V. It is clear that Ψ and Υ are inverse to each other, thus the first isomorphism is established.

The final claim follows immediately. □

Definition I.12. Let U, V, W be k-vector spaces. If $\phi \in \mathrm{Hom}(U, V)$ and $\psi \in \mathrm{Hom}(V, W)$, then the map $(\psi, \phi) \mapsto \psi \circ \phi$ given by composition of functions induces a linear map

$$c\colon \mathrm{Hom}(V, W) \otimes \mathrm{Hom}(U, V) \to \mathrm{Hom}(U, W).$$

In particular, taking $V = \mathsf{k}$, we obtain a morphism $\vartheta\colon U^* \otimes V \to \mathrm{Hom}(U, V)$, and if $U = V$ then one also has the composition in the opposite direction, $\iota\colon V^* \otimes V \to \mathsf{k}$, where $\iota(f \otimes v) = f(v)$.

Lemma I.13. *Let V and W be vector spaces. The natural map $\vartheta\colon V^* \otimes W \to Hom_{\mathsf{k}}(V, W)$ is injective with image the space $Hom^{fr}(V, W) = \{\alpha\colon V \to W : \dim(im(\alpha)) < \infty\}$ of linear maps of finite rank. Moreover, when V is finite-dimensional, if $\iota\colon V^* \otimes V \to \mathsf{k}$ is the contraction map and $\alpha \in Hom_{\mathsf{k}}(V, V)$, then $(\iota \circ \theta^{-1})(\alpha) = tr(\alpha)$.*

Proof. To see that the map $\theta\colon V^* \otimes W \to \mathrm{Hom}(V, W)$ is injective, suppose that $t = \sum_{i=1}^{n} \delta_i \otimes w_i \in \ker(\theta)$. By part *i*) of Lemma I.10 we may assume that $\{w_1, \dots, w_n\}$ is linearly independent. But then $\vartheta(t) = 0$ implies that for all $v \in V$ we have $\vartheta(t)(v) = \sum_{i=1}^{n} \delta_i(v).w_i = 0$ and so, since the w_i are linearly independent, we see that for all $v \in V$ and each i, $(1 \leq i \leq n)$ we have $\delta_i(v) = 0$, that is, $\delta_i = 0$. But then clearly $t = \sum_{i=1}^{n} 0 \otimes w_i = 0$ as required.

To see that $\mathrm{im}(\vartheta) = \mathrm{Hom}^{fr}(U, W)$ note that $\theta(f \otimes w) = f.w$ has image contained in $\mathsf{k}.w$, hence $\mathrm{rank}(\vartheta(f \otimes w)) \leq 1$. But then a finite sum $\sum_{k=1}^{m} f_i \otimes w_i$ can have rank at most m, so that $\mathrm{im}(\vartheta) \subseteq \mathrm{Hom}^{fr}(V, W)$. To see that $\mathrm{Hom}^{fr}(V, W) \subseteq \mathrm{im}(\vartheta)$, suppose that $\alpha\colon U \to W$ is finite rank so that $\mathrm{im}(\alpha) = W_1$ is finite-dimensional. Pick a basis $\{w_1, \dots, w_n\}$ for W_1, and let $\{\delta_1, \dots, \delta_n\}$ be the corresponding dual basis of W_1^* so that if $w \in W_1$ then $w = \sum_{i=1}^{n} \delta_i(w).w_i$. But then if $v \in V$, $\alpha(v) \in W_1$ so that $\alpha(v) = \sum_{i=1}^{n} \delta_i(\alpha(v)).w_i$. Thus $\alpha = \vartheta(\sum_{i=1}^{n} \alpha^{\top}(\delta_i) \otimes w_i)$ lies in the image of ϑ as claimed.

Finally we consider the contraction map $\iota\colon V^* \times V \to \mathsf{k}$. This is again composition, but now in the opposite order, so that $v\colon \mathsf{k} \to V$ and $f\colon V \to \mathsf{k}$ compose to give $f(v) \in \mathsf{k}$. If $\{e_1, \dots, e_n\}$ is a basis of V and $\{\delta_1, \dots, \delta_n\}$ the dual basis of V^*, then $1_V = \sum_{i=1}^{n} \delta_i.e_i$ (since they agree on the basis $\{e_1, \dots, e_n\}$) and so $\alpha = \alpha \circ 1_V = \alpha\big(\sum_{i=1}^{n} \delta_i.e_i\big) = \sum_{i=1}^{n} \delta_i.\alpha(e_i)$. Thus $\vartheta^{-1}(\alpha) = \sum_{i=1}^{n} \delta_i \otimes \alpha(e_i)$, and we have

$$\mathrm{tr}(\alpha) = \sum_{i=1}^{n} \delta_i(\alpha(e_i)) = \iota(\sum_{i=1}^{n} \delta_i \otimes \alpha(e_i)) = \iota \circ \theta^{-1}(\alpha),$$

where the first equality is simply the definition of tr, the second follows from the the definition of ι and the third by using the formula we just obtained for $\vartheta^{-1}(\alpha)$. □

Remark I.14. Since we only use the cases where V and W are finite dimensional, the reader is welcome to ignore the generality the result is stated in and assume throughout that all vector spaces are finite dimensional. Here one can be a bit more concrete: if $\{e_1, \dots, e_n\}$ is a basis of V and $\{f_1, \dots, f_m\}$ is a basis of W, then taking the dual basis $\{\delta_1, \dots, \delta_n\}$ of V^* it is easy to see that the images of $\delta_i \otimes f_j$ under ϑ correspond to the elementary matrices E_{ij} under the identification of $\mathrm{Hom}_{\mathsf{k}}(V, W)$ given by the choice of bases for V and W, hence ϑ is an isomorphism.

Remark I.15. We will usually abuse notation somewhat and write $\iota\colon V \otimes V^* \to \mathsf{k}$ rather than $\iota \circ \sigma$ where $\sigma\colon V \otimes V^* \to V^* \otimes V$ interchanges the tensor factors.

I.2.2 Linear maps between tensor products.

Let $\alpha\colon V_1 \to V_2$ and $\beta\colon W_1 \to W_2$ be linear maps. If $v \in V_1, w \in W_1$, the map $(v, w) \mapsto \alpha(v) \otimes \beta(w)$ from $V_1 \times W_2 \to V_2 \otimes W_2$ is bilinear, and so induces a linear map $\mathrm{Hom}(V_1 \otimes W_1, V_2 \otimes W_2)$, which we denote by $\alpha \otimes \beta$. In fact, the map $(\alpha, \beta) \mapsto \alpha \otimes \beta$ is itself bilinear, and so we even obtain a map

$$\mathrm{Hom}(V_1, W_1) \otimes \mathrm{Hom}(V_2, W_2) \to \mathrm{Hom}(V_1 \otimes V_2, W_1 \otimes W_2). \tag{I.3}$$

Moreover, it follows immediately from the definitions that (I.3) also respects composition. In more detail, if $\alpha_2\colon V_2 \to V_3$ and $\beta_2\colon W_2 \to W_3$ are linear maps to any vector spaces V_3 and W_3, then $(\alpha_2 \otimes \beta_2) \circ (\alpha_1 \otimes \beta_1) = (\alpha_2 \circ \alpha_1) \otimes (\beta_2 \circ \beta_1)$. Indeed, if $v \in V_1, w \in W_1$, then

$$\begin{aligned}(\alpha_2 \otimes \beta_2) \circ (\alpha_1 \otimes \beta_1)(v \otimes w) &= (\alpha_2 \otimes \beta_2)(\alpha_1(v) \otimes \beta_1(w)) \\ &= (\alpha_2 \circ \alpha_1)(v) \otimes (\beta_2 \circ \beta_1)(w) \\ &= (\alpha_2 \circ \alpha_1) \otimes (\beta_2 \circ \beta_1)(v \otimes w).\end{aligned}$$

When all the vector spaces V_1, V_2, W_1, W_2 are finite dimensional, the map (I.3) is actually an isomorphism, indeed using Lemma I.13 you can check that

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

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

Example I.16. The map $\iota\colon V^* \otimes V \to \mathsf{k}$ also describes the composition of linear maps: Suppose we have three vector spaces V, W and U. The composition gives a bilinear map from $\mathrm{Hom}(U, V) \times \mathrm{Hom}(V, W)$ to $\mathrm{Hom}(U, W)$, thus it is equivalent to a linear map $\tilde{m}\colon \mathrm{Hom}(U, V) \otimes \mathrm{Hom}(V, W) \to \mathrm{Hom}(U, W)$.

$$\begin{tikzcd}
\mathrm{Hom}(U,V)\otimes\mathrm{Hom}(V,W) \arrow[rr] & & (U^*\otimes V)\otimes(V^*\otimes W) \arrow[dll] \\
U^*\otimes(V\otimes V^*)\otimes W \arrow[r] & U^*\otimes \mathsf{k}\otimes W \arrow[r] & U^*\otimes W
\end{tikzcd}$$

where the first arrow is the induced by the isomorphisms provided by Lemma I.13, the second from the associativity of tensor products, and the third arrow is $1_V \otimes \iota \otimes 1_{V^*}$. By Example I.9 scalar multiplication gives a natural isomorphism $s\colon \mathsf{k} \otimes (V \otimes W) \to (V \otimes W)$, and the final arrow swaps the first two factors and then applies s. Identifying the term $U^* \otimes W$ with $\mathrm{Hom}(U, W)$ this becomes the composition of linear maps. This can be checked by considering the composition of rank-one maps: if $f \otimes v \in U^* \otimes V$ then it corresponds to the rank-one map $u \mapsto f(u).v$ $(u \in U)$. Thus if we take $f.v \in \mathrm{Hom}(U, V)$ and $g.w \in \mathrm{Hom}(V, W)$ (where $f \in U^*, v \in V, g \in V^*, w \in W$) then $(g.w) \circ (f.v)(u) = (g.w)(f(u).v) = g(f(u).v).w = f(u).g(v).w = g(v).(f.w)(u)$. On the other hand $(f \otimes v) \otimes (g \otimes w) \mapsto f \otimes g(v) \otimes w \mapsto g(v).(f \otimes w)$.

Remark I.17. It is sometimes useful to have the following notational convention: Given a tensor product of more than two vector spaces, such as $U^* \otimes V \otimes V^* \otimes W$, then it can be convenient to write ι_{32} for the map which acts via ι on the third and second factors (that is swapping the second and third factors, applying ι and the repeating the swap) and by the identity on the remaining tensor factors.

I.2.3 Tensor products and duality

Suppose that V and W are finite dimensional vector spaces. We wish to understand the relationship between the tensor product of the dual spaces $V^* \otimes W^*$ and the dual space of the tensor product $(V \otimes W)^*$. If $\eta \in V^*$ and $\nu \in W^*$ then $\eta.\nu\colon V \times W \to \mathsf{k}$ given by $(v, w) \mapsto \eta(v).\nu(w)$ is a bilinear map, hence it induces a linear map which by abuse of notation we will also denote as $\eta.\nu\colon V \otimes W \to \mathsf{k}$. Thus $(\eta, \nu) \mapsto \eta.\nu$ is a map from $V^* \times W^*$ to $(V \otimes W)^*$. Since it is also bilinear it induces a linear map $d_{V,W}\colon V^* \otimes W^* \to (V \otimes W)^*$. The map $d_{V,W}$ is injective since, if $t \in V^* \otimes W^*$ is in $\ker(d_{V,W})$ then by part *i*) of Lemma I.10 we may write $t = \sum_{i=1}^m \delta_i \otimes \eta_i$ where $\{\eta_1, \dots, \eta_m\}$ is linearly independent in W^*. But then $d_{V,W}(t) = \sum_{i=1}^m \delta_i.\eta_i = 0$, and so in particular, for all $v \in V$ we must have $\sum_{i=1}^m \delta_i(v).\eta_i = 0$, so by the linearly independence of the η_i we have $\delta_i(v) = 0$ for all $v \in V$, that is $\delta_i = 0$, and hence $t = 0$ as required.

Another way to view $d_{V,W}$ is as follows: Let $c_{V,W}\colon V^* \otimes W^* \otimes V \otimes W \to \mathsf{k}$ be the linear map given by $c_{V,W} = \iota_{13}^V \otimes \iota_{24}^W$, where ι_{13}^V denotes the contraction map ι^V acting on the first and third tensor factors, and ι_{24}^W similarly denotes the ccontraction ι^W acting on the second and fourth factor, that is

$$c_{V,W}\left(\eta \otimes \nu \otimes v \otimes w\right) = \iota^V(\eta \otimes v).\iota^W(\nu \otimes w) = \eta(v).\nu(w).$$

Now $c_{V,W}$ yields a bilinear map $t_{V,W}\colon (V^* \otimes W^*) \times (V \otimes W) \to \mathsf{k}$ which induces, by Lemma I.11, a linear map $(V^* \otimes W^*) \to (V \otimes W)^*$, and this is just the map $d_{V,W}$ constructed above. Moreover, the linear functional $c_{V,W}$, by permuting the tensor factors, can be viewed as a linear functional

$$c_{V,W} = \iota_{14}^V \otimes \iota_{23}^W\colon (V^* \otimes W) \otimes (W^* \otimes V) \to \mathsf{k}.$$

Now if we assume V and W are finite-dimensional, then $(V^* \otimes W) \cong \mathrm{Hom}(V, W)$ and $W^* \otimes V \cong \mathrm{Hom}(W, V)$, so that $c_{V,W}$ gives a linear functional on $\mathrm{Hom}(V, W) \otimes \mathrm{Hom}(W, V)$. Now $1_{V^*} \otimes \iota_{23}^W \otimes 1_V$ corresponds to composition of linear maps $(a, b) \mapsto a \circ b$ and $\iota_{14}^V \otimes \iota_{23}^W = \iota^V \circ (1_{V^*} \otimes \iota_{23}^W \otimes 1_V)$, and $\iota^V\colon V^* \otimes V \cong \mathrm{Hom}(V, V) \to \mathsf{k}$ gives the trace map. Thus $c_{V,W}$ viewed as a linear map $\mathrm{Hom}(V, W) \otimes \mathrm{Hom}(W, V) \to \mathsf{k}$ is just the trace form $(a \otimes b) \mapsto \mathrm{tr}(ab)$. As noted in the proof of Lemma 5.2.4, this description of the trace form also makes the symmetry property $\mathrm{tr}(ab) = \mathrm{tr}(ba)$ is evident.

I.3 Bilinear forms

Definition I.18. Let V be a k-vector space. A *bilinear form* on V is a bilinear map $B\colon V \times V \to \mathsf{k}$, that is, an element of $\mathscr{M}(V, V; \mathsf{k})$. We will denote the vector space of all bilinear forms on V as $\mathrm{Bil}(V)$. From the universal property of tensor products, $\mathrm{Bil}(V) \cong (V \otimes V)^*$.

Let $S_2 = \{e, \sigma\}$ denote the symmetric group on two letters. There is a natural linear action of S_2 on $\mathrm{Bil}(V)$ given by $\sigma(B)(v, w) = B(w, v)$ (for any $v, w \in V$). A form B is said to be *symmetric* if $B = \sigma(B)$, that is if $B(v, w) = B(w, v)$ for all $v, w \in V$, and *skew-symmetric* if $\sigma(B) = -B$.

If $B \in \mathrm{Bil}(V)$ satisfies $B(v, v) = 0$ we say that B is *alternating*. Since $0 = B(v + w, v + w) = B(v, v) + B(v, w) + B(w, v) + B(w, w) = B(v, w) + B(w, v)$, any alternating bilinear form is skew-symmetric. Conversely, if B is skew-symmetric, then $B(v, v) = -B(v, v)$ so that $2.B(v, v) = 0$. Thus, *provided that* $\mathrm{char}(\mathsf{k}) \neq 2$, the alternating and skew-symmetry properties coincide. Moreover, if we write $S^2(V)$ for the space of symmetric bilinear forms on V and $\Lambda^2(V)$ for the space of alternating bilinear forms on V, then we have $\mathrm{Bil}(V) = S^2(V) \oplus \Lambda^2(V)$, as they are the $+1$ and -1 eigenspace of the involution σ. More concretely, we have $B = B^+ + B^-$, where $B^+ = (B + \sigma(B))/2$ and $B^- = (B - \sigma(B))/2$ where $\sigma(B^\pm) = \pm B^\pm$.

We may deal with the symmetric and skew-symmetric cases uniformly (to some extent) by working with a form B which has the property that $B(v, w) = \epsilon.B(w, v)$ for all $v, w \in V$, where $\epsilon \in \{\pm 1\}$.

Remark I.19. Lemma I.11 gives a natural isomorphism

$$\Theta\colon \mathrm{Bil}(V) = \mathscr{M}(V, V; \mathsf{k}) \to \mathrm{Hom}(V, \mathrm{Hom}(V, \mathsf{k})) = \mathrm{Hom}(V, V^*)$$

It follows that giving a bilinear form on V is equivalent to giving a linear map from V to V^*. Note that the action of $\sigma \in S_2$ gives a second isomorphism $\Theta_1\colon \mathrm{Bil}(V) \to \mathrm{Hom}(V, V^*)$, where $\Theta_1 = \Theta \circ \sigma$, that is, $\Theta_1(B)(v)(w) = B(w, v)$. For symmetric bilinear forms the two maps agree, but for arbitrary bilinear forms they yield different isomorphisms.

Definition I.20. Given a bilinear form B, we set

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

(here the subscript "L" denotes "left"). Similarly, we set

$$\mathrm{rad}_R(B) = \ker(\Theta_1(B)) = \{v \in V : B(w, v) = 0, \forall\, w \in V\}.$$

If B is symmetric or alternating, then $\mathrm{rad}_L(B) = \mathrm{rad}_R(B)$, but this need not be true otherwise. We say that B is *non-degenerate* if $\mathrm{rad}_L(B) = \{0\}$. Note that, even though in general $\mathrm{rad}_L(B) \neq \mathrm{rad}_R(B)$, it is still the case that $\mathrm{rad}_L(B) = \{0\}$ if and only if $\mathrm{rad}_R(B) = \{0\}$.

From now on we will only work with symmetric and alternating bilinear forms. Fix such a $B \in \mathrm{Bil}(V)$ so that $\sigma(B) = \epsilon.B$ for some $\epsilon \in \{\pm 1\}$. Then if U is a subspace of V, we define

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

When B is nondegenerate, so that $\Theta(B)$ is an isomorphism, this shows that $\dim(U^{\perp}) = \dim(U^0) = \dim(V) - \dim(U)$. The next Lemma shows that this can be refined slightly.

Lemma I.21. *Let V be a finite-dimensional k-vector space equipped with a symmetric (or alternating) bilinear form B. Then for any subspace U of V we have the following:*

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

ii) The restriction of B to U is nondegenerate if and only if $V = U \oplus U^{\perp}$.

Proof. Let $\phi\colon V \to U^*$ be given by $\phi(v)(u) = B(v, u)$, that is $\phi(v) = (\Theta(B)(v))_{|U}$. Clearly $\ker(\phi) = U^{\perp}$, while $\mathrm{im}(\phi) \leq U^*$ and hence $\dim(\mathrm{im}(\phi)) \leq \dim(U)$. The inequality in *i*) now follows from rank-nullity.

For the second part, note that B is non-degenerate on U if and only if $U \cap U^{\perp} = \{0\}$. But then the inequality in *i*) shows that we must have $U \oplus U^{\perp} = V$ for dimension reasons. □

I.4 Classification of symmetric bilinear forms

There is a natural linear action of $\mathrm{GL}(V)$ on the space $\mathrm{Bil}(V)$: if $g \in \mathrm{GL}(V)$ and $B \in \mathrm{Bil}(V)$ then we set $g(B)$ to be the bilinear form given by

$$g(B)(v, w) = B(g^{-1}(v), g^{-1}(w)), \quad (v, w \in V),$$

where the inverses ensure that the above equation defines a left action. It is clear the action preserves the subspace of symmetric bilinear forms.

Since we can find a invertible map taking any basis of a vector space to any other basis, the next lemma says that over an algebraically closed field there is only one nondegenerate symmetric bilinear form up to the action of $\mathrm{GL}(V)$, that is, when k is algebraically closed the nondegenerate symmetric bilinear forms are a single orbit for the action of $\mathrm{GL}(V)$.

Lemma I.22. *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[8]

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

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. □

[8] Note that this identity holds unless $\mathrm{char}(\mathsf{k}) = 2$. It might be useful to remember this identity when understanding the Proposition which is the key to the proof of the Cartan Criterion: it claims that if $\mathfrak{g} = D\mathfrak{g}$ then there is an element $x \in \mathfrak{g}$ with $\kappa(x, x) \neq 0$. Noting the above identity, we see this is equivalent to asserting that κ is nonzero.

Remark I.23. 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").

One can also classify alternating forms using essentially the same strategy, except that if B is a non-zero alternating form on a vector space V, one shows that it contains a two-dimensional space H on which B is nondegenerate. Then we can choose a basis $\{e, f\}$ of V with $B(e, f) = 1 = -B(f, e)$, and then since $V = H \oplus H^\perp$ one can apply induction. Moreover, in the alternating case, the classification holds over any field k where $\mathrm{char}(\mathsf{k}) \neq 2$.

II Reminder on Representation theory

We recall here some basics of representation theory used in the course, all of which is covered (in much more detail than we need) in the Part B course on Representation theory. Let $\mathfrak{g}$ be a Lie algebra. The main body of the notes proves all that is needed in the course, but the material here might help clarify some arguments. We will always assume our representations are finite dimensional unless we explicitly say otherwise.

II.1 Basic notions

Definition II.1. A representation is *irreducible* if it has no proper nonzero subrepresentations. A representation (V, ρ) is said to be *indecomposable* if it cannot be written as a direct sum of two proper subrepresentations. A representation is said to be *completely reducible* if is a direct sum of irreducible representations.

Clearly an irreducible representation is indecomposable, but the converse is not in general true. For example k^2 is naturally a representation for the nilpotent Lie algebra of strictly upper triangular matrices $\mathfrak{n}_2 \subset \mathfrak{gl}_2(\mathsf{k})$ and it is not hard to see that it has a unique 1-dimensional sub representation, hence it is indecomposable, but not irreducible.

A basic observation about irreducible representations is Schur's Lemma:

Lemma II.2. *Let $\mathfrak{g}$ be a Lie algebra and let (V, ρ), (W, σ) be irreducible representations of $\mathfrak{g}$. Then any $\mathfrak{g}$-homomorphism $\phi\colon V \to W$ is either zero or an isomorphism. In particular, if k is algebraically closed, then $\mathrm{Hom}_{\mathfrak{g}}(V, W)$ is one-dimensional.*

Proof. The proof is exactly the same as the proof for finite groups. If ϕ is nonzero, then $\ker(\phi)$ is a proper subrepresentation of V, hence as V is irreducible it must be zero. It follows V is isomorphic to $\phi(V)$, which is thus a nonzero subrepresentation of W. But then since W is irreducible we must have $W = \phi(V)$ and ϕ is an isomorphism as claimed.

Thus if $\mathrm{Hom}_{\mathsf{k}}(V, W)$ is nonzero, we may fix some $\phi\colon V \to W$ an isomorphism from V to W. Then given any $\mathfrak{g}$-homomorphism $\alpha\colon V \to W$, composing with ϕ^{-1} gives a $\mathfrak{g}$-homomorphism from V to V, thus it is enough to assume $W = V$. But then if $\alpha\colon V \to V$ is a $\mathfrak{g}$-endomorphism of V, since k is algebraically closed, it has an eigenvalue λ and so $\ker(\alpha - \lambda)$ is a nonzero subrepresentation, which must therefore be all of V, that is $\alpha = \lambda.\mathrm{id}_V$, so that $\mathrm{Hom}_{\mathfrak{g}}(V, V)$ is one-dimensional as claimed. □

II.2 Exact sequences of representations

Parallel to the notion for Lie algebras, there is also a notion of an exact sequence for representations. Let $\mathfrak{g}$ be a Lie algebra.

Definition II.3. A sequence of maps of $\mathfrak{g}$-representations

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

is said to be *exact at V* if $\mathrm{im}(\alpha) = \ker(\beta)$. A sequence of maps

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

is called a *short exact sequence* if it is exact at each of U, V and W, so that α is injective and β is surjective and $\mathrm{im}(\alpha) = \ker(\beta)$. If V is the middle term of such a short exact sequence, it contains a subrepresentation isomorphic to U, such that the corresponding quotient representation is isomorphic to W, and hence, roughly speaking, V is built by gluing together U and W. Just as for Lie algebras, an exact sequence

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

is said to be *split* if β admits a right inverse $s\colon W \to V$, that is, a $\mathfrak{g}$-homomorphism s such that $\beta \circ s = \mathrm{id}_W$.

The next Lemma shows that the situation for representations is simpler than it is for Lie algebras[9]:

Lemma II.4. *Suppose that $\mathfrak{g}$ is a Lie algebra and*

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

is a short exact sequence of $\mathfrak{g}$-representations. Then the sequence is split if and only if $\iota(U)$ has a complementary subrepresentation W', that is, $V = \iota(U) \oplus W'$, and if W' exists then necessarily $q_{|W'}: W' \to W$ is an isomorphism.

Proof. First suppose that $s: W \to V$ be a splitting map and let $W' = s(W)$. Since $q \circ s = 1_W$, it follows that $\{0\} = \ker(q_{|W'}) = \ker(q) \cap W' = \iota(U) \cap W'$. But $\iota(U) + W' = V$ since for any $v \in V$ we have $v = (v - s \circ \beta(v)) + s \circ \beta(v)$ where certainly $s \circ \beta(v) \in s(W)$, and since $q \circ s = \mathrm{id}_W$ we have

$$q(v - s \circ q(v)) = q(v) - q \circ s \circ q(v) = q(v) - \mathrm{id}_W \circ q(v) = q(v) - q(v) = 0,$$

hence $v - s \circ q(v) \in \ker(q) = \iota(U)$ and $V = \iota(U) \oplus W'$ as required.

For the converse, note that if $V = \iota(U) + W'$ then $q(V) = q(W') = W$, and $\ker(q_{|W'} = \ker(q) \cap W' = \iota(U) \cap W' = \{0\}$, so that $q_{|W'}: W' \to W$ is an isomorphism, hence $s = (q_{|W'})^{-1}$ is a splitting map. □

Remark II.5. If the short exact sequence of Lemma II.4 is split and we have a splitting map $s: W \to V$, it yields an isomorphism $\tilde{s}: U \oplus W \to V$ given by $\tilde{s}(u, w) = \iota(u) + s(w)$, since the proof of the Lemma shows that $\tilde{s}$ is a bijective homomorphism of representations. If π_1, π_2 denote the projection maps from $U \oplus W$ onto the factors U and W, then $q = \pi_2 \circ \tilde{s}^{-1}$, and if we set $p = \pi_1 \circ \tilde{s}^{-1}$, then $\tilde{s}(v) = \big(p(v), q(v - \iota(p(v))\big)$ so that the splitting is also determined by $p: V \to U$ a left-inverse to ι.

Our study of the representations of a nilpotent Lie algebra $\mathfrak{n}$ can be interpreted as calculating, for a given 1-dimensional representation k_α of $\mathfrak{n}$ which one-dimensional representations k_β can form non-split extensions of it.

Lemma II.6. *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. Using the above Lemma and induction, one can actually recover the theorem that any representation V of a nilpotent Lie algebra $\mathfrak{n}$ decomposes into a direct sum of its isotypical subrepresentations V_α (where $\alpha \in D(\mathfrak{n})^0 \subseteq \mathfrak{n}^*$).

Example II.7. To see a non-split extension, let $\mathfrak{g} = \mathfrak{n}_2$ be the one-dimensional Lie algebra, thought of as the (nilpotent) Lie algebra of 2×2 strictly upper triangular matrices. Then its natural 2-dimensional representation on k^2 given by the inclusion $\mathfrak{n}_2 \to \mathfrak{gl}_2(\mathsf{k})$ gives a non-split extension

$$0 \longrightarrow \mathsf{k}_0 \xrightarrow{i} \mathsf{k}^2 \longrightarrow \mathsf{k}_0 \longrightarrow 0$$

where k_0 is the trivial representation, and $i: \mathsf{k}_0 \to \mathsf{k}^2$ is the inclusion $t \mapsto (t, 0)$. The extension cannot be trivial, because the image of $\mathfrak{n}_2$ is non-zero. It is fact it's easy to see using linear algebra that for $\mathfrak{gl}_1(\mathsf{k}) = \mathfrak{n}_2$, an extension of one-dimensional representations k_α and k_β automatically splits if $\alpha \neq \beta$ while there is, up to isomorphism, one non-split extension of k_α with itself ($\alpha, \beta \in (\mathfrak{gl}_1(\mathsf{k}))^*$). The splitting statement is a special case of the following more general result, a special case of Theorem 4.3.13.

The notion of a composition series has an analogue for representations of a given Lie algebra $\mathfrak{g}$.

Definition II.8. Let V be a $\mathfrak{g}$-representation. A nested sequence of subrepresentations $\mathscr{C} = (V = F_0 \supset F_1 \supset \ldots \supset F_d = 0)$ is said to be a *composition series* for V if the subquotients F_{i-1}/F_i are irreducible (for each $i \in \{1, \ldots, d\}$). The isomorphism classes of the irreducibles which arise in this way are known as the *composition factors* of the series $\mathscr{C}$. For V finite-dimensional, it is clear that V must contain proper subrepresentations which are maximal with respect to containment among proper subrepresentations (since one can choose one of maximal dimension). From this an inductionon $\dim(V)$ it follows easily that any finite-dimensional representation has a composition series.

For an irreducible representation S, the multiplicity $[S : \mathscr{C}]$ with which S occurs as a composition factor if $\mathscr{C}$ is known as its composition multiplicity. Thus

$$[S : \mathscr{C}] = \left|\{j : 1 \leq j \leq d, F_{j-1}/F_j \cong S\}\right|.$$

Let us also define $[S : V] = \min\{[S : \mathscr{C}] : \mathscr{C}$ a composition series for $V\}$.

[9] In the sense that there are no non-trivial semi-direct products.

Remark II.9. A composition series $\mathscr{C} = (V = F_0 > \dots > F_d = \{0\})$ can also be viewed as the vestige of how the representation V was built up from its composition factor $S_i = F_i/F_{i+1}$. Indeed for each $k \in \{1, \dots, d\}$ we have

$$0 \longrightarrow S_k = F_k/F_{k+1} \xrightarrow{\alpha} V/F_{k+1} \xrightarrow{\beta} V/F_k \longrightarrow 0$$

Thus starting with $S_0 = F_0/F_1$, one obtains F_0/F_2 by extending it by S_1. Continuing in this way, extending F_0/F_2 by S_2 one obtains F_0/F_3 and so on, until finally we get V by extending F_0/F_{d-1} by $S_{d-1} = F_{d-1}$ to obtain V itself!

A composition series for a representation V naturally induces a composition series for any subrepresentation of V and for the image of V any homomorphism $\phi\colon V \to W$.

Proposition II.10. *Let V be a representation and $W \le V$ a subrepresentation. Then if $\mathscr{C} = (F_k)_{k=0}^d$ is a composition series for V, then $\mathscr{C}$ induces composition series $\mathscr{C}_W$ and $\mathscr{C}_{V/W}$ for W and V/W respectively. Moreover, if S is an irreducible representation, then $[S : \mathscr{C}] = [W : \mathscr{C}_W] + [S : \mathscr{C}_{V/W}]$.*

Proof. Let $W_k = W \cap F_k$, so that $(W_k)_{k=0}^d$ is a descending filtration of W by subrepresentations. Using the second isomorphism theorem we see that

$$W_k/W_{k+1} = W_k/W_k \cap F_{k+1} \cong (W_k + F_{k+1})/F_{k+1} \subseteq F_k/F_{k+1}. \tag{II.1}$$

Next, let $q\colon V \to V/W$ be the quotient map, and let $Q_k = q(F_k)$, so that $(Q_k)_{k=0}^d$ is a descending filtration of V/W. Now $\ker(q_{|F_k}) = W \cap F_k = W_k$, hence by the first isomorphism theorem $F_k/W_k \cong q(F_k)$. Under this isomorphism $q(F_{k+1}) \le q(F_k)$ is identified with $(F_{k+1} + W_k)/W_k$, hence by the 3rd isomorphism theorem

$$Q_k/Q_{k+1} = q(F_k)/q(F_{k+1}) \cong F_k/W_k\big/(F_{k+1} + W_k)/W_k \cong F_k/(F_{k+1} + W_k) \tag{II.2}$$

Now $F_{k+1} \le F_{k+1} + W_k \le F_k$, hence as F_k/F_{k+1} is irreducible, $F_{k+1} + W_k$ must be one of F_k or F_{k+1}. But it follows from (II.1) and (II.2) that in the former case $W_k/W_{k+1} \cong F_k/F_{k+1}$ and $Q_k = Q_{k+1}$, while in the latter, $W_k = W_{k+1}$ and $Q_k/Q_{k+1} \cong F_k/F_{k+1}$.

Thus if we let $J = \{k : 0 \le k \le d, F_{k+1} + W_k = F_k\}$ and $K = \{k : 0 \le k \le d, F_{k+1} + W_k = F_{k+1}\}$, and set $\mathscr{C}_W = (W_j)_{j\in J}$ and $\mathscr{C}_{V/W} = (Q_k)_{k\in K}$ (ordered so as to form a descending chain) it follows that $\mathscr{C}_W$ is a composition series for W and $\mathscr{C}_{V/W}$ is a composition series for V/W. Moreover if $j \in J$ then the composition factor F_j/F_{j+1} of $\mathscr{C}$ corresponds to a composition factor of $\mathscr{C}_W$, while if $j \in K$ it corresponds to a composition factor of $\mathscr{C}_{V/W}$, which readily implies the multiplicity equation. $\square$

This result allows one to give a quick proof of the *Jordan-Hölder theorem* for $\mathfrak{g}$-representations.[10]

Corollary II.11. *For any finite-dimensional representation V of a Lie algebra $\mathfrak{g}$ and any irreducible representation S, the multiplicity with which S occurs in a composition series for V is independent of the choice of composition series for V, and hence equals $[S : V]$.*

Proof. For *i*) we use induction on the minimal length $n(V)$ of a composition series for V. If $n(V) = 1$ then V is irreducible and $(V > 0)$ is its unique composition series. If $n = n(V) > 1$ then take a composition series $\mathscr{M} = (M_i)_{i=0}^n$ of V with length n and set $U = M_1$. Since $(M_{i+1})_{i=0}^{n-1}$ is a composition series for U, we have $n(U) \le n-1$. Now if $\mathscr{C} = (F_i)_{i=0}^d$ is any composition series for V, by Proposition II.10, it induces composition series $\mathscr{C}_U$ and $\mathscr{C}_{V/U}$ of U and V/U respectively. Thus if S is irreducible, by the final sentence of Proposition II.10 we have

$$[S : \mathscr{C}] = [S : \mathscr{C}_U] + [S : \mathscr{C}_{V/U}] = [S : U] + [S : V/U]$$

where the second equality follows by induction since $n(V/U) = 1$ and $n(U) \le n-1$. Thus $[S : \mathscr{C}] = [S : V]$ is independent of $\mathscr{C}$. Part *ii*) now follows immediately from the final sentence of Proposition II.10. $\square$

Remark II.12. Note that Proposition II.10 and Corollary II.11 together show that $[S : V] = [S : W] + [S : V/W]$ for any subrepresentation $W \le V$.

[10] The same proof works for representations of groups or finite-dimensional algebras.

II.3 Semisimplicity and complete reducibility

Definition II.13. A representation (V, ρ) is said to be *semisimple* if any subrepresentation U has a complement, that is, there is a subrepresentation W such that $V = U \oplus W$. A representation is said to be completely reducible if it is a direct sum of irreducible representations. Note that Lemma II.4 shows that V is semisimple if and only if every short exact sequence

$$0 \longrightarrow U \longrightarrow V \longrightarrow W \longrightarrow 0$$

splits. Indeed this follows from Lemma II.4: the image of a splitting map $s\colon W \to V$ gives a complement to the image of U, and s is determined by its image.

The following simple Lemma is one reason why exact sequences of representations are easier to work with than exact sequences of Lie algebras.

Lemma II.14. *Let $\mathfrak{g}$ be a Lie algebra and suppose that we have a short exact sequence of $\mathfrak{g}$-representations:*

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

Then dualizing we obtain a sequence

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

which is again a short exact sequence. It follows that a short exact sequence splits if and only if there is a map $t\colon V \to U$ such that $t \circ i = 1_U$.

Proof. The injectivity of $q^\top$ follows from the surjectivity of q. To see that the $i^\top$ is surjective, any functional δ on U extends to one on V. But this is easy – simply pick a complement T to U and define $\delta(T) = 0$. Finally we must show that $\mathrm{im}(q^\top) = \ker(i^\top)$. Since $q \circ i = 0$, we have $i^\top \circ q^\top = 0$, so that $\mathrm{im}(q^\top) \subseteq \ker(i^\top)$. The equality then follows by considering dimensions.

For the final sentence, note that is $p\colon V \to U$ is a map satisfying $p \circ i = 1_U$, then $i^\top \circ p^\top = 1_{U^*}$, and hence $p^\top$ is a splitting of the dual short exact sequence. But then $V^* \cong U^* \oplus W^*$, and hence taking duals and using the canonical isomorphism it follows that $V = U \oplus W$ so that our original sequence was split. ☐

Lemma II.15. *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, then $q^{-1}(W_1) = V_1$ is a subrepresentation of V, which has a complement V_2. Then we claim that $W_2 = q(V_2)$ is a complement to W_1 in W: indeed since q is surjective clearly $W = W_1 + W_2$, and if $w \in W_1 \cap W_2$ then there exist $v_2 \in V_2$ with $q(v_2) = w \in W_1$. But then $v_2 \in q^{-1}(W_1) = V_1$ hence $v_2 \in V_1 \cap V_2 = \{0\}$ and $w = q(v_2) = 0$ as required.

Next, if U is a subrepresentation of V, then picking a complement U' to U, so that $V = U \oplus U'$, the corresponding projection map $\pi\colon V \to U$ with kernel U' shows that U is isomorphic to a quotient of V, and hence is also semisimple. ☐

Lemma II.16. *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. To see *i*) implies *ii*), use induction on dimension: if U is a non-zero subrepresentation of V of minimal dimension, U must be simple. If $U = V$ then we are done, otherwise U has a non-zero complement W with $\dim(W) = \dim(V) - \dim(U) < \dim(V)$. By induction $W = \bigoplus_{k=1}^{m} S_k$ where each S_k is simple, and thus setting $U = S_{m+1}$ we see $V = \bigoplus_{k=1}^{m+1} S_k$.

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[11] $(X \oplus W) \cap U = \{0\}$, which contradicts the maximality of W, so we are done. ☐

Remark II.17. If $\mathfrak{g}$ is nilpotent, and V a $\mathfrak{g}$-representation, then for any one-dimensional representation λ, the isotypical subrepresentation (or λ-weight space) V_λ is a direct summand of V, *i.e.* it has a complementary subrepresentation. Since this is true for all λ we obtain a direct sum decomposition $V = \bigoplus_{\lambda \in (\mathfrak{g}/D(\mathfrak{g}))^*} V_\lambda$.

If $\mathfrak{g}$ is semisimple, then the representations of $\mathfrak{g}$ are semisimple, so *any* subrepresentation has a complement. It follows that if V is a $\mathfrak{g}$ representation and χ is an irreducible representation, the isotypical subrepresentation V_χ has a complement, so $V = V_\chi \oplus U$. Inductively it therefore follows that $V = \bigoplus_\chi V_\chi$ where the sum is over the irreducible representations of $\mathfrak{g}$ that occur in V. Since any semisimple representation is completely reducible, the subrepresentations V_χ are just a direct sum of copies of χ, that is, $V_\chi = V_\chi^s$, *i.e.* V_χ is equal to its socle.

In particular, the isotypic summand of V corresponding to the trivial representation χ_0 is $V_{\chi_0} = V^{\mathfrak{g}}$, the invariants of V. A consequence of the complete reducibility is that $V^{\mathfrak{g}}$ should be a direct summand of V. In fact in the proof of Weyl's theorem, we showed this by proving that $V = V^{\mathfrak{g}} \oplus \mathfrak{g}.V$, and then deduced semisimplicity from this. Note that for any Lie algebra $\mathfrak{g}$ and $\mathfrak{g}$-representation V, the quotient $V/\mathfrak{g}.V$ is the largest quotient of V on which $\mathfrak{g}$ acts trivially. If V is semisimple, and $V = \bigoplus_\chi V_\chi$ is its decomposition into isotypical summands, then if $U \leq V$ is any subrepresentation, we can similarly decompose $U = \bigoplus_\chi U_\chi$, and hence $V/U \cong \bigoplus_\chi V_\chi/U_\chi$. It follows that V/U is invariant for the action of $\mathfrak{g}$ if and only if $U_\chi = V_\chi$ for all nontrivial χ, and hence $\mathfrak{g}.V = \bigoplus_{\chi \neq 0} V_\chi$. Thus if V is semisimple, we must have $V = V^{\mathfrak{g}} \oplus \mathfrak{g}.V$, that is, the subrepresentation we used as the candidate for a complement to $V^{\mathfrak{g}}$ in our proof of Weyl's theorem was in fact forced on us.

[11] Since both $X \cap (W \oplus U) = \{0\}$ and $(X \oplus W) \cap U = \{0\}$ are both equivalent to the sum $X + W + U$ being direct.

III *On the construction of simple Lie algebras

The classification of semisimple Lie algebras, as discussed in §7.3, relies on two key results: an *Isomorphism theorem*, and an *Existence theorem*: the former ensures that the root system captures enough information to determine the Lie algebra up to isomorphism, while the latter ensures that every abstract root system arises as the root system of some semisimple Lie algebra.

This section outlines one approach to the existence theorem. Clearly it is enough to construct a simple Lie algebra for each indecomposable root system, so we will assume throughout the remainder of this section that (V, Φ) is indecomposable. We will establish the existence theorem in two steps. In the first step we consider the case where all the roots in Φ have the same length, and in the second step deduce from this the general case. An alternative elementary approach is described in [Gec17].

III.1 The simply-laced case

Definition III.1. Let (V, Φ) be an (indecomposable) root system. We say that (V, Φ) is *simply-laced* if all the roots in Φ have the same length.

If Δ is a set of simple roots for Φ, since $\Phi = W.\Delta$ (where W is the Weyl group) it is equivalent to the condition that all the roots in Δ have the same length. Since (V, Φ) is indecomposable, this in turn is equivalent to the condition that $\langle \alpha, \beta \rangle = \langle \beta, \alpha \rangle$ for all $\alpha, \beta \in \Delta$, that is, the Cartan matrix is symmetric. By Lemma 7.2.20, this is equivalent to the condition that for all $\alpha, \beta \in \Phi$ the Cartan integer $\langle \alpha, \beta \rangle \in \{0, -1\}$. If we normalize the inner product on V so that the roots have length $\sqrt{2}$, then the Cartan integers are precisely the values of the inner product on pairs of simple roots.

From the classification of abstract root systems, one can check that the simply-laced indecomposable root systems are those of types A,D and E.

To construct a Lie algebra from such a root system, we need one additional ingredient: Let $\epsilon\colon Q \times Q \to \{\pm 1\}$ be a *bimultiplicative function*, that is, for all $\alpha, \beta, \gamma \in Q$,

$$\begin{aligned} \epsilon(\alpha + \beta, \gamma) &= \epsilon(\alpha, \gamma).\epsilon(\beta, \gamma), \\ \epsilon(\alpha, \beta + \gamma) &= \epsilon(\alpha, \beta).\epsilon(\alpha, \gamma). \end{aligned}$$

and suppose also that it satisfies

$$\epsilon(\alpha, \alpha) = (-1)^{(\alpha,\alpha)/2}, \quad \forall\, \alpha \in Q \tag{III.1}$$

(note that since $(\alpha, \alpha) = 2$ for all roots $\alpha \in \Phi$, we must have $(\beta, \beta) \in 2\mathbb{Z}$ for any $\beta \in Q$). Such a function is called an *asymmetric function*. Since $(\alpha, \beta) \in \mathbb{Z}$ for any $\alpha, \beta \in Q$ we can replace α by $\alpha + \beta$ in the second condition (III.1) for an asymmetric function to obtain:

$$\epsilon(\alpha, \beta)\epsilon(\beta, \alpha) = (-1)^{(\alpha,\beta)}. \tag{III.2}$$

Note that the bimultiplicativity property means it is determined by its values on a base Δ and moreover the second condition (III.1) requires $\epsilon(\alpha_i, \alpha_i) = -1$ for any $\alpha_i \in \Delta$. To construct such a function on the rest of $\Delta \times \Delta$, orient the edges of the Dynkin diagram, whose vertices are labelled by the base $\Delta = \{\alpha_1, \dots, \alpha_l\}$, arbitrarily, and then define for $\alpha_i \neq \alpha_j$

$$\epsilon(\alpha_i, \alpha_j) = \begin{cases} -1 & \text{if there is an edge going from } \alpha_i \text{ to } \alpha_j, \\ +1 & \text{otherwise.} \end{cases}$$

It the follows from this definition that Equation (III.2) holds for all roots in our base, and thus extending this ϵ bimultiplicatively, we obtain an asymmetric function on all of Q.

We can now give a construction of the Lie algebra $\mathfrak{g}_Q$ associated to our root system: Let $\mathfrak{h}^*$ denote the extension of scalars from $\mathbb{Q}$ to our field k of V, and similarly we can extend our inner product to a symmetric bilinear form on $\mathfrak{h}^*$. Let $\mathfrak{h}$ be the dual of $\mathfrak{h}$.

Definition III.2. Let $\mathfrak{g}_Q = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Phi} \mathsf{k}.e_\alpha$ as a vector space, and let h_α be the image of α under the isomorphism between $\mathfrak{h}$ and $\mathfrak{h}^*$ given by the nondegenerate symmetric bilinear form on $\mathfrak{h}^*$ induced from the inner product on V.

We define

$$[h, h'] = 0, \forall h, h' \in \mathfrak{h};$$
$$[h, e_\alpha] = \alpha(h)e_\alpha;$$
$$[e_\alpha, e_\beta] = \begin{cases} -h_\alpha, & \text{if } \alpha + \beta = 0; \\ \epsilon(\alpha, \beta).e_{\alpha+\beta} & \text{if } \alpha + \beta \in \Phi; \\ 0 & \text{otherwise.} \end{cases}$$

We also extend the symmetric bilinear form on $\mathfrak{h}$ (obtained by identifying it with $\mathfrak{h}^*$) to all of $\mathfrak{g}_Q$ by setting $(e_\alpha, e_\beta) = -\delta_{\alpha,-\beta}$, and making $\mathfrak{h}$ orthogonal to $\bigoplus_{\alpha\in\Phi} \mathsf{k}.e_\alpha$. (Note the minus signs in the definition of the invariant form and in the bracket $[e_\alpha, e_{-\alpha}]$ are consistent.)

Proposition III.3. *The definition above gives a Lie algebra which has $\mathfrak{h}$ as a Cartan subalgebra and root system Φ, and the form on $\mathfrak{g}_Q$ is invariant.*

Proof. (*Sketch*): We must show that $\mathfrak{g}_Q$ is a Lie algebra, that is, we must check that the bilinear map $[,]$ defined above is a Lie bracket. To see that it is alternating, note that if $\{\alpha, \beta, \alpha + \beta\} \subseteq \Phi$ then, since the root system is simply-laced, $(\alpha, \beta) = -1$, and hence (III.2) shows that $\epsilon(\alpha, \beta) = -\epsilon(\beta, \alpha)$. It remains to check that $[.]$ satisfies the Jacobi identity. It is enough to check this on three basis elements, x, y and z. If at least one of our basis elements is in $\mathfrak{h}$ this is easy (the properties of the bimultiplicative function beyond the one already used for the alternating property are not involved). For example, if $x = h \in \mathfrak{h}, y = e_\alpha, z = e_\beta$ then (setting $e_{\alpha+\beta} = 0$ if $\alpha + \beta \notin \Phi$)

$$\begin{aligned} &[h, [e_\alpha, e_\beta]] + [e_\alpha, [e_\beta, h]] + [e_\beta, [h, e_\alpha]] \\ &\qquad = \epsilon(\alpha, \beta)\left((\alpha + \beta)(h)e_{\alpha+\beta} - \beta(h)e_{\alpha+\beta} - \alpha(h)e_{\alpha+\beta}\right) \\ &\qquad = 0. \end{aligned}$$

If x, y, z are of the form $e_\alpha, e_\beta, e_\gamma$ then there are a number of cases to check. Firstly, if none of $\alpha + \beta, \alpha + \gamma, \beta + \gamma$ lie in $\Phi \cup \{0\}$, then the Jacobi identity holds trivially. Thus let us suppose that $\alpha + \beta \in \Phi \cup \{0\}$. Note that $\alpha \pm \beta \in \Phi$ if and only if $(\alpha, \beta) = \mp 1$. Moreover, it follows that $\epsilon(\alpha, \alpha) = -1$ and $\epsilon(\alpha, \beta)\epsilon(\beta, \alpha) = -1$.

There are four cases: 1) $\alpha \pm \gamma \notin \Phi \cup \{0\}$; 2) either $\alpha + \gamma$ or $\alpha - \gamma = 0$; 3) $\alpha + \gamma \in \Phi$ and; 4) $\alpha - \gamma \in \Phi$. Cases 1) and 2) are easy to check, case 3) follows from

$$\epsilon(\gamma, \alpha)\epsilon(\gamma + \alpha, -\alpha) = (\alpha, \alpha).$$

In this fashion one can reduce to the case where $\alpha + \beta$, $\alpha + \gamma$ and $\beta + \gamma$ all lie in Φ. But then $(\alpha, \beta) = (\alpha, \gamma) = (\beta, \gamma) = -1$ and so $(\alpha + \beta + \gamma, \alpha + \beta + \gamma) = 0$ so that $\alpha + \beta + \gamma = 0$. In this case the Jacobi identity

$$[e_\alpha, [e_\beta, e_\gamma]] + [e_\beta, [e_\gamma, e_\alpha]] + [e_\gamma, [e_\alpha, e_\beta]] = 0$$

reduces to

$$\epsilon(\beta, \gamma)\epsilon(\alpha, \beta + \gamma) + \epsilon(\gamma, \alpha)\epsilon(\beta, \alpha + \gamma) + \epsilon(\alpha, \beta)\epsilon(\gamma, \alpha + \beta) = 0$$

which can be checked using the properties of ϵ.

It is similar, though more straight-forward, to check that the symmetric bilinear form we have defined is invariant. □

III.2 The non-simply-laced cases

One can also use the construction of the simply-laced Lie simple Lie algebras to give a construction of *all* simple Lie algebras: We do this as follows: Given a simply-laced Dynkin diagram D, a *admissible diagram automorphism* is a graph automorphism $\sigma: D \to D$ with the property that the orbit of a vertex is discrete, that is, there is no edge between a vertex i and $\sigma^k(i)$ for any $k \in \mathbb{Z}$.

Given such an automorphism, we claim that σ induces an automorphism of $\mathfrak{g}_Q$ the associated simple Lie algebra. To see this, note that we can pick the orientation of our Dynkin diagram so that it is invariant under the diagram automorphism (we will check this shortly for the automorphisms we need). Clearly σ induced an isometry of V to itself preserving the roots Φ (it clearly preserves Q and hence Φ since Φ is the set of norm 2 vectors in Q). Moreover, it preserves the bimultiplicative function ϵ because it preserved the orientation of our Dynkin diagram (by our choice of orientation).

Defining σ on $\mathfrak{g}_Q$ by letting $\sigma(e_\alpha) = e_{\sigma(\alpha)}$ and letting it act on $\mathfrak{h}$ by extension of scalars of its action on V, it is then clear that σ is a Lie algebra homomorphism. It follows that its fixed point set is a sub-Lie algebra.

Theorem III.4. *The Lie algebra $\mathfrak{g}_Q^\sigma$ is a simple Lie algebra with Dynkin diagram D^σ given as follows: the vertices of D^σ are the orbits of σ on the vertex set of D, and, for any two orbits, they are joined if there were edges joining a vertex in one orbit to a vertex in the other, etc..*

Bibliography

[Gec17] Meinolf Geck. On the construction of semisimple lie algebras and chevalley groups. *Proc. Amer. Math. Soc.*, 145(8):3233–3247, 2017.