Finite Dimensional Lie and Associative Algebras

Christopher Brookes

These notes, taken by Markus Himmel, will at times differ significantly from what was lectured. In particular, all errors are almost certainly my own.

Contents

CHAPTER 1

Introduction

Definition 1.1. Let k be a field. A Lie algebra $\mathfrak{L}$ over k is a k-vector space with a bilinear map $[\cdot,\cdot]\colon \mathfrak{L}\times\mathfrak{L}\to\mathfrak{L}$ satisfying

(1) $\forall x\in\mathfrak{L}\colon [x,x]=0$, and
(2) $\forall x,y,z\in\mathfrak{L}\colon [x,[y,z]]+[y,[z,x]]+[z,[x,y]]=0$. This is the Jacobi identity.

Bilinearity and (1) imply

(1') $\forall x,y\in\mathfrak{L}\colon [x,y]=-[y,x]$,

and if $\operatorname{char} k\neq 2$, then bilinearity and (1') imply (1).

Remark. Groups describe symmetries. Lie algebras describe infinitesimal symmetries.

For example, let $G=\mathrm{GL}_n(\mathbb{R})$. This is an example of a Lie group, i.e., an analytic manifold with continuous group operations. The associated Lie algebra is the tangent space T_1G at the identity.

The matrix exponential diffeomorphically (with inverse log) takes a neighborhood of 0, which is the same as T_1G, to a neighborhood of 1.

$\exp A\exp B=\exp(\mu(A,B))$ for sufficiently small A and B.

The Taylor series for μ is

$$\mu(A,B)=A+B+\frac{1}{2}[A,B]+\text{higher degree terms},$$

where $[A,B]=AB-BA$ (matrix multiplication).

This is an example of a Lie bracket. Note that $T_1G\times T_1G\to T_1G$, $(A,B)\mapsto[A,B]$ is bilinar, skew-symmetric.

The Lie algebra corresponding to G is often called $\mathfrak{g}$.

Note that

(1) The first approximation to the group product is addition in the Lie algebra T_1G.
(2) If $[g_1,g_2]=g_1g_2g_1^{-1}g_2^{-1}$ is the group commutator, then the Lie bracket is the first approximation of the commutator $[\exp A,\exp B]$ in G.
(3) The Jacobi identity arises from the associativity in G. Note that Lie algebras in general are non-associative.

As a further example, let $G=\mathrm{GL}_n(\mathbb{C})$. This is an example of an algebraic group, i.e., a complex algebraic variety with continuous group operations. We have $T_1G\cong M_n(\mathbb{C})$ as the tangent space at the identity. Similarly to before, we define a Lie bracket and end up with a complex Lie algebra.

Definition 1.2. (a) A Lie subalgebra $\mathfrak{J}$ of $\mathfrak{L}$ is a k-subspace such that $[x,y]\in\mathfrak{J}$ for $x,y\in\mathfrak{J}$.
(b) An ideal $\mathfrak{J}$ of $\mathfrak{L}$ is a k-subalgebra such that $[x,y]\in\mathfrak{J}$ for $x\in\mathfrak{J}$ and $y\in\mathfrak{L}$.

In a couple of lectures we will define a canonical ideal $R(\mathfrak{L})$.

Definition 1.3. (a) $\mathfrak{L}$ is semisimple if $R(\mathfrak{L})=0$. In general $\mathfrak{L}/R(\mathfrak{L})$ is semisimple.

(b) $\mathfrak{L}$ is simple if the only ideals are 0 and $\mathfrak{L}$.

We will see that semisimple Lie algebras are direct products of finitely many simple ones. In this course we will concentrate on the simple complex Lie algebras.

We will find that classifying these boils down to classifying finite root systems, which are collections of combinatorial data. Root systems have a symmetry group called the Weyl group and are labelled by Dynkin diagrams.

Root systems also appear in the representations of quivers (i.e., directed graphs) arising in algebraic geometry.

Definition 1.4. An associative ring R with unity is a k-algebra if there is a ring homomorphism $\phi\colon k \to R$ such that $\phi(k) \leq Z(R)$, where $Z = \{r \in R \mid \forall s \in R\colon rs = sr\}$ is the centre of R.

We can regard k as a subalgebra of R and R is a k-vector space.

Remark. If R is a k-algebra, we can define a Lie bracket $[r, s] = rs - sr$, where we use the associative product, so R is a Lie algebra.

Definition 1.5. (a) A k-subspace J of R is a left ideal if $\forall r \in R, s \in J\colon rs \in J$. Right ideals are define analogously. A (2-sided) ideal is both a left and a right ideal.

We'll see that in finite-dimensional k-algebras there is a canonical ideal, the Jacobson radical $J(R)$.

Definition 1.6. (a) R is semisimple if $J(R) = 0$, and in general $R/J(R)$ is semisimple.

(b) R is simple if the only ideals are 0 and R.

Exercise: $M_n(k)$ is a simple algebra (work out the left and the right ideals).

We will prove the Artin-Wedderburn theorem which says the finite-dimensional semisimple algebras are direct products of simple ones, where simple algebras are isomorphic to $M_n(D)$, where D is a division algebra, where $\dim_k D < \infty$.

An example of a skew field are the quaternions $\mathbb{H}$. They are an $\mathbb{R}$-algebra with a basis $1, i, j, k$ such that $ij = k, ji = -k$. The quaternions are not a $\mathbb{C}$-algebra.

Artin-Wedderburn applies in

(a) representation theory of finite groups,
(b) path algebras R of quivers, where R-modules correspond to representations of quivers.

Definition 1.7. An R-module M is indecomposable if one cannot express it as $M = M_1 \oplus M_2$ with $M_1, M_2 \neq 0$.

We will consider quivers where the path algebras only have finitely many isomorphism classes of indecomposable modules. These quivers are called quivers of finite representation type.

The classification due to Gabriel again involves root systems labelled by Dynkin diagrams.

CHAPTER 2

Elementary properties of Lie algebras

Remark. Assume that $\operatorname{char} k = 0$.

Example. $\mathfrak{gl}_n$ has Lie subalgebras:

(1) $\mathfrak{sl}_n$ is the subalgebra of trace zero matrices. It is associated with SL_n.

Example: $\mathfrak{sl}_2$ is a 3-dimensional k-vector space. It has a standard basis given by

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We notice that $[e, f] = h$, $[h, e] = 2e$, $[h, f] = -2f$.

(2) $\mathfrak{so}_n$ is the subalgebra of skew-symmetric ($A + A^T = 0$) matrices. It is associated with SO_n, the special orthogonal group (endomorphisms preserving an inner product).

Example: $\mathfrak{so}_3$ is a 3-dimensional k-vector space. It has a basis given by

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We have $[A_1, A_2] = A_3$, $[A_2, A_3] = A_1$, $[A_3, A_1] = A_2$.

(3) $\mathfrak{sp}_{2n}$ is the subalgebra of matrices A, such that $JA^TJ^{-1} + A = 0$ where J has -1s on the lower-left half of the antidiagonal and 1s on the upper-right half of the antidiagonal. It is associated with the group SP_{2n} preserving a non-degenerate skew-symmetric bilinar form (also known as a symplectic form).

(4) $\mathfrak{b}_n$ is the subalgebra of upper triangular matrices, also called the Borel subalgebra and is associated with the inverted upper triangular matrices.

(5) $\mathfrak{n}_n$ is the subalgebra of strictly upper triangular matrices with zeros on the leading diagonal. It is associated with the upper triangular matrices with ones on the leading diagonal.

We can also consider $\mathrm{End}_k(R)$, which are the k-linear maps $R \to R$, where R is an associative algebra. If $\dim R = n$, then $\mathrm{End}(R) = M_n(k)$. $\mathrm{End}(R)$ has a Lie subalgebra called $\mathrm{Der}(R)$ consisting of derivations.

Definition 2.1. A k-linear map $D\colon R \to R$ is called a derivation if it satisfies the Leibnitz rule:

$$D(rs) = D(r)s + rD(s),$$

where we are taking products in R.

Example. We have $\mathrm{Der}(k[X]) = \{fD \mid f \in k[X]\}$, where $D\colon k[X] \to k[X]$ is the differential (straightforward proof by induction).

$\mathrm{Der}(k[X, X^{-1}])$ is called the Witt Lie algebra, which is closely related to the Virasoro algebra (appears in geometry and physics). It is infinite-dimensional.

Geometrically, when R is a coordinate ring, then $\mathrm{Der}(R)$ corresponds to vector fields. However, R need not be commutative in the general case.

DEFINITION 2.2. An inner derivation is a k-linear map $R \to R$ of the form $s \mapsto [r, s]$ for some $r \in R$.

The inner derivations form a Lie subalgebra of $\mathrm{Der}(R)$ and in fact form a (Lie) ideal.

REMARK.
(1) If R is commutative, then $\mathrm{Innder}(R) = 0$.
(2) At the end of the commutative algebra course you may meet Hochschild cohomology (a cohomology theory for associative algebras). The first Hochschild cohomology group $HH^1(R, R)$ is the quotient $\mathrm{Der}(R)/\mathrm{Innder}(R)$, which is a Lie algebra.
(3) Lie algebras appear as derivations of other algebraic structures. For example for the octonians one gets the Lie algebra G_2.

1. Representations

DEFINITION 2.3.
(a) A morphism of Lie algebras $\rho\colon \mathfrak{L}_1 \to \mathfrak{L}_2$ is a k-linear map such that $\rho([x, y]) = [\rho(x), \rho(y)]$.
(b) A representation of $\mathfrak{L}$ is a morphism of Lie algebras $\rho_V : \mathfrak{L} \to \mathrm{End}\, V$, where V is a vector space. If $\dim V < \infty$, we call ρ_V a linear representation.

If $U \leq V$ and $\rho_V(\mathfrak{L})(U) \subseteq U$, then there is a subrepresentation $\rho_U : \mathfrak{L} \to \mathrm{End}\, U$ where $\rho_U(x)(u) \coloneqq \rho_V(x)(u)$ for $x \in \mathfrak{L}, u \in U$.

An irreducible representation is one that does not admit any proper subrepresentations.

EXAMPLE.
(1) The adjoint representation $\mathrm{ad}_{\mathfrak{L}}\colon \mathfrak{L} \to \mathrm{End}\, \mathfrak{L}$ is given by $x \mapsto (y \mapsto [x, y])$.

It is indeed a homomorphism: if $x, y, z \in \mathfrak{L}$, then we may calculate

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

where we have used the Jacobi identity.

DEFINITION 2.4. The centre of $\mathfrak{L}$ is defined to be

$$\ker \mathrm{ad}_{\mathfrak{L}} = \{x \in \mathfrak{L} \mid \forall y \in \mathfrak{L}\colon [x, y] = 0\}.$$

Note that if the centre is 0 then the adjoint representation is injective and we can regard $\mathcal{L}$ as a subalgebra of $\mathrm{End}\, \mathfrak{L}$. If $\mathfrak{L}$ is finite-dimensional, then $\mathfrak{L}$ is a subalgebra of $\mathfrak{gl}_n \cong \mathrm{End}\, \mathfrak{L}$, where $n = \dim \mathfrak{L}$.

REMARK. There is a difficult result called Ado's theorem which states that if $\mathrm{char}\, k = 0$ and $\mathfrak{L}$ is finite-dimensional then there is an injective morphism of Lie algebras $\mathfrak{L} \to \mathfrak{gl}_n$ for some n.

Iwasawa then extended this to characteristic $p > 0$ (quite hard).

EXAMPLE. Let $k = \mathbb{R}$. $\mathbb{R}^3$ is a Lie algebra under the cross product (have to check the Jacobi identity). If e_1, e_2, e_3 form the standard basis, then we find that

$$e_1 \times e_2 = e_3, \qquad e_2 \times e_3 = e_1, \qquad e_3 \times e_1 = e_2.$$

Using this, we quickly calculate that the adjoint map $\mathrm{ad}_{\mathbb{R}^3} : \mathbb{R}^3 \to \mathrm{End}\, \mathbb{R}^3 \cong M_3(\mathbb{R})$ sends e_i to the basis element $A_i \in \mathfrak{so}_3(\mathbb{R})$.

In particular $\ker \mathrm{ad}_{\mathbb{R}^3} = 0$, $\mathrm{im}\, \mathrm{ad}_{\mathbb{R}^3} = \mathfrak{so}_3$. Thus $\mathbb{R}^3$ with the vector product is isomorphic to $\mathfrak{so}_3(\mathbb{R})$ as a Lie algebra.

EXAMPLE. We define a morphism

$$\begin{aligned}\rho\colon \mathfrak{sl}_2 &\to \mathrm{Der}(k[X,Y]) \subseteq \mathrm{End}(k[X,Y])\\ e &\mapsto X\frac{\partial}{\partial Y}\\ f &\mapsto Y\frac{\partial}{\partial X}\\ h &\mapsto X\frac{\partial}{\partial X} - Y\frac{\partial}{\partial Y}\end{aligned}$$

An easy but somewhat lengthy calculation shows that this is a morphism (notably, we use the symmetry of second partial derivatives). Note that the images of e, f, h map V_n, the span of the monomials of total degree n ($\dim V_n = n+1$; for example, V_1 has basis elements X, Y, while V_2 has basis elements X^2, XY, Y^2) to itself. So we have subrepresentations $\mathfrak{sl}_2 \to \mathrm{End}\, V_n$. Exercise: think about the cases $n = 1$ and $n = 2$ and show that they are irreducible.

LEMMA 2.5. The subrepresentations $\rho_n\colon \mathfrak{sl}_2 \to \mathrm{End}(V_n)$ are irreducible.

PROOF. Suppose $\rho_n(\mathfrak{sl}_2)(U) \subseteq U$ for a subspace U. Then if $U \neq 0$ there exists $f \in U$, where $\sum_{i+j=n} \lambda_{ij} X^i Y^j$ where not all λ_{ij} are zero. Then

$$\rho_n(e)(f) = XD_Y(f) = \sum j\lambda_{ij}X^{i+1}Y^{j-1} \in U.$$

Repeatedly applying $\rho_n(e)$ yields a nonzero scalar multiple of X^n, so $X^n \in U$. Now apply $\rho_n(f)$ repeatedly to get nonzero scalar multiples of all monomials in V_n. So if U is nonzero, then $U = V_n$ as required. $\square$

REMARK. Note that $\bigoplus V_n = k[X,Y]$.

A note about terminology: Strictly speaking, the representation is the map $\mathfrak{L} \to \mathrm{End}(V)$. Often, V is also called the representation. This is an abuse of notation. In this course, we will use the term "module" for V, for example "V is a module for $\mathfrak{sl}_2$" or "V is a $\mathfrak{sl}_2$-module." Similarly, we'll sometimes use the term "simple module" to refer to irreducible representations.

We'll see later that the V_n are precisely the simple finite-dimensional $\mathfrak{sl}_2$-modules up to isomorphism.

Also any finite-dimensional $\mathfrak{sl}_2$-module is a direct sum of copies of the V_n.

However, there are infinite-dimensional $\mathfrak{sl}_2$-modules that aren't such direct sums. There will be an example on the example sheet.

DEFINITION 2.6. A Lie algebra is called abelian if $\forall x, y \in \mathfrak{L}, [x,y] = 0$.

For example, all 1-dimensional Lie algebras are abelian.

DEFINITION 2.7. The derived series of $\mathfrak{L}$ is defined inductively: $\mathfrak{L}^{(0)} := \mathfrak{L}$, $\mathfrak{L}^{(n+1)} := [\mathfrak{L}^{(n)}, \mathfrak{L}^{(n)}]$, where $[\mathfrak{L}, \mathfrak{L}]$ is the span (!) of the elements of the form $[x,y]$, $x, y \in \mathfrak{L}$.

We call $\mathfrak{L}^{(1)}$ the derived subalgebra of $\mathfrak{L}$.

Note that $\mathfrak{L}^{(i)}$ is a Lie ideal of $\mathfrak{L}$: this follows from induction and the Jacobi identity.

DEFINITION 2.8. The Lie algebra $\mathfrak{L}$ is called soluble if $\mathfrak{L}^{(r)} = 0$ for some r. The derived length of $\mathfrak{L}$ is the least such r.

For example, being a non-zero abelian Lie algebra is equivalent to the derived length being 1.

REMARK. If J is an ideal of $\mathfrak{L}$, then $\mathfrak{L}/J$ is a Lie algebra via $[x+J, y+J] := [x,y] + J$.

LEMMA 2.9. (1) Subalgebras and quotients of soluble Lie algebras are soluble.
(2) If J is an ideal such that J and $\mathfrak{L}/J$ are soluble, then $\mathfrak{L}$ is soluble.

PROOF. For (1), it suffices to notice that for a subalgebra J we have $J^{(i)} \subseteq J \cap \mathfrak{L}^{(i)}$, and the for an ideal L we have $(\mathfrak{L}/J)^{(i)} \subseteq \mathfrak{L}^{(i)} + J$. Both statements are readily proved by induction.

For (2), we first notice that $\mathfrak{L}^{(i)} + J \subseteq (\mathfrak{L}/J)^{(i)}$. But since $\mathfrak{L}/J$ is soluble, this implies that $\mathfrak{L}^{(n)} \subseteq J$ for some n. Since J is soluble, we conclude that there is some m such that $\mathfrak{L}^{(m+n)} = (\mathfrak{L}^{(n)})^{(m)} \subseteq J^{(m)} = 0$. □

EXAMPLE. Let $\mathfrak{L}$ be a 2-dimensional Lie algebra. Either $\mathfrak{L}$ is abelian or there are x, y such that $[x, y] \neq 0$, so $\mathfrak{L}^{(1)} \neq 0$.

However, x and y form a basis of $\mathfrak{L}$, $\mathfrak{L}^{(1)}$ is equal to the span of $[x, y]$. Therefore, the derived series of $\mathfrak{L}$ looks like

$$\mathfrak{L} \supsetneq \mathfrak{L}^{(1)} \supsetneq 0.$$

So in the first case, where $\mathfrak{L}$ is abelian, the derived length is 1, and otherwise the derived length is 2.

Annoying exercise: classify three-dimensional Lie algebras. It is done in Jacobson's book.

DEFINITION 2.10. The lower central series is defined inductively: $\mathfrak{L}_{(1)} := \mathfrak{L}$, $\mathfrak{L}_{(n+1)} := [\mathfrak{L}_{(n)}, \mathfrak{L}]$. Recall that we are taking spans here.

Note $\mathfrak{L}_{(i)}$ are ideals of $\mathfrak{L}$.

We say that $\mathfrak{L}$ is nilpotent if $\mathfrak{L}_{(c+1)} = 0$ for some c. The nilpotency class of $\mathfrak{L}$ is the smallest such c.

Note that if $\mathfrak{L}$ is nilpotent, then $\mathfrak{L}$ is soluble.

EXAMPLE. Recall that $\mathfrak{n}_n$ is the Lie algebra of strictly upper triangular matrices. Exercise: this is nilpotent for every n.

For example, $\mathfrak{n}_3$ is called the Heisenberg Lie algebra. It has dimension 3. There is an obvious basis

$$x = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad z = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We can calculate that $[x, y] = z$, $[x, z] = 0$, $[y, z] = 0$, so $\mathfrak{n}_3$ is nonabelian and of nilpotency class 2. In general, we can show that $\mathfrak{n}_n$ is of nilpotency class $n - 1$.

EXAMPLE. Recall $\mathfrak{b}_n$ consists of the upper triangular matrices. We have $\mathfrak{b}_n^{(1)} = \mathfrak{n}_n$. $\mathfrak{b}_n$ is soluble but not nilpotent for $n \geq 2$.

LEMMA 2.11. If $\mathfrak{L}$ is a Lie algebra and $n \in \mathbb{N}$, then $\mathfrak{L}^{(n)} \subseteq \mathfrak{L}_{(2^n)}$.

PROOF. We will first show that for natural numbers i and j we have $[\mathfrak{L}_{(i)}, \mathfrak{L}_{(j)}] \subseteq \mathfrak{L}_{(i+j)}$.

We do induction on j. The case $j = 1$ is true by definition.

Now assume that for some $j \in \mathbb{N}$ and all $i \in \mathbb{N}$ we have $[\mathfrak{L}_{(i)} + \mathfrak{L}_{(j)}] \subseteq \mathfrak{L}_{(i+j)}$. Let $i \in \mathbb{N}$. We need to show that $[\mathfrak{L}_{(i)}, \mathfrak{L}_{(j+1)}] \subseteq \mathfrak{L}_{(i+j+1)}$. We will check this on generators, so let $x \in \mathfrak{L}_{(i)}$, $y \in \mathfrak{L}_{(j)}$ and $z \in \mathfrak{L}$. We need to show that $[x, [y, z]] \in \mathfrak{L}_{(i+j+1)}$.

Indeed, $[x, y] \in \mathfrak{L}_{(i+j)}$ by our inductive hypothesis, so $\alpha := [z, [x, y]] \in \mathfrak{L}_{(i+j+1)}$ by definition. Furthermore, $[z, x]$ in $\mathfrak{L}_{(i+1)}$ by definition, so $\beta := [y, [z, x]] \in \mathfrak{L}_{(i+j+1)}$ by inductive hypothesis. Therefore $[x, [y, z]] = -\alpha - \beta \in \mathfrak{L}_{(i+j+1)}$ as required, completing the proof of the lemma.

Now we will proceed to prove the claim, again by induction. The base case is again trivial, and for $n \in \mathbb{N}$ we have

$$\mathfrak{L}^{(n+1)} = [\mathfrak{L}^{(n)}, \mathfrak{L}^{(n)}] \subseteq [\mathfrak{L}_{(2^n)}, \mathfrak{L}_{(2^n)}] \subseteq \mathfrak{L}_{(2^{n+1})},$$

using the inductive hypothesis and our lemma. This completes the proof. □

Remark. Out next aim is to prove some theorems.

Theorem 2.12 (Engel). Suppose $\mathfrak{L} \subseteq \operatorname{End} V$ is a subalgebra with $\dim V < \infty$ and every $x \in L$ is a nilpotent endomorphism.

Then there is some $v \in V$ such that $v \neq 0$, but $\forall x \in L\colon x(v) = 0$.

Proof. We proceed by induction on $\dim \mathfrak{L}$.

Assume first that $\dim \mathfrak{L} = 1$, i.e., $\mathfrak{L} = \langle x \rangle$. Since x is nilpotent, then x has eigenvalue 0, so there is $v \neq 0$ such that $x(v) = 0$. Since x spans $\mathfrak{L}$, we have $\mathfrak{L}(v) = 0$.

Next, assume that $\dim \mathfrak{L} > 1$. We will first show that $\mathfrak{L}$ satisfies the idealiser condition. Let $A \subsetneq \mathfrak{L}$ be a proper Lie subalgebra. Consider $\rho\colon A \to \operatorname{End} \mathfrak{L}$ given by $a \mapsto \operatorname{ad}(a) = (x \mapsto [a, x])$, the restriction of the adjoint representation of $\mathfrak{L}$ to A. Since A is a subalgebra, there is a representation $\overline{\rho}\colon A \to \operatorname{End}(L/A)$ given by $a \mapsto \overline{\operatorname{ad}(a)} = (x + A \mapsto [a, x] + A)$. This is indeed a representation, because A is a subalgebra.

By (2.17) we know that if a is nilpotent, then so is $\operatorname{ad}(a)$, which implies that $\overline{\operatorname{ad}(a)}$ is also nilpotent. Note that $\dim \overline{\rho}(A) \leq \dim A < \dim \mathfrak{L}$.

By the inductive hypothesis, we find $0 \neq x' \in L/A$ such that $\forall f \in \overline{\rho}(A)\colon f(x') = 0$. In other words, we find $x \in L \setminus A$ such that for all $a \in A$ we have

$$\overline{\rho}(a)(x + A) = A.$$

By definition of $\overline{\rho}$, this just means that $[a, x] \in A$ for all $a \in A$, which implies that $[x, a] \in A$ for $a \in A$. Therefore, $x \in \operatorname{Id}_L(A) \setminus A$ and the idealiser condition is indeed satisfied.

Now, if M is a maximal proper subalgebra of $\mathfrak{L}$, then $\operatorname{Id}_{\mathfrak{L}}(M) = \mathfrak{L}$ by maximality of M. This just means that M is an ideal of $\mathfrak{L}$. This means that $\mathfrak{L}/M$ is a Lie algebra and the maximality of M forces $\dim(\mathfrak{L}/M) = 1$, because every Lie algebra has subalgebras of dimension 1 (indeed, the span of any nonzero element is one) and these can be pulled back to Lie subalgebras in between M and $\mathfrak{L}$.

This means that $\mathfrak{L} = \langle M, x \rangle$ for some $x \in \mathfrak{L}$.

Consider $U := \{u \in V \mid M(u) = 0\}$. By the inductive hypothesis, since $\dim M < \dim \mathfrak{L}$, we know that $U \neq 0$.

Let $u \in U$ and $m \in M$. Then $m(x(u)) = ([m, x] + x \circ m)(u) = 0$, since $m \in M$ and $[m, x] \in M$ as M is an ideal. So $x(u) \in U$ for all $u \in U$. This means that x restricts to a nilpotent endomorphism of U and so has an eigenvector $0 \neq v \in U$ with $x(v) = 0$ (every eigenvector of a nilpotent endomorphism must be zero). But $v \in U$ and so $M(v) = 0$. As $\mathfrak{L}$ is the span of M and x, it follows that $\mathfrak{L}(v) = 0$ as required. □

Theorem 2.13 (Lie). Assume that k is algebraically closed of characteristic 0. Again, let $\mathfrak{L} \subseteq \operatorname{End} V$ be a subalgebra with $\dim V < \infty$. Suppose that $\mathfrak{L}$ is soluble. Then there is some $v \in V$ such that $v \neq 0$ and for all $x \in L$ there is $\lambda_x \in k$ such that $x(v) = \lambda_x v$.

In words: all x have a common eigenvector.

Proof. Again, we use induction on $\dim \mathfrak{L}$.

If $\dim \mathfrak{L} = 1$, then we can use the fact that k is algebraically closed to find an eigenvector of x such that $\mathfrak{L} = \langle x \rangle$, and we are done.

Next, assume that $\dim \mathfrak{L} > 1$ and suppose the the theorem is true for all soluble Lie subalgebras of $\operatorname{End} W$ of smaller dimension.

Since $\mathfrak{L} \neq 0$ and $\mathfrak{L}$ is soluble, we have $\mathfrak{L}^{(1)} \subsetneq \mathfrak{L}$. Let M be a maximal Lie subalgebra containing $\mathfrak{L}^{(1)}$. Then M is an ideal of $\mathfrak{L}$ (since $[x, y] \subseteq [\mathfrak{L}, \mathfrak{L}] \subseteq M$) and $\dim L/M = 1$ (as seen in the proof of Engel's theorem). Again, pick $x \in \mathfrak{L}$ such that $\mathfrak{L}$ is the span of M and x. By induction, we find $0 \neq u \in V$ such that $\forall m \in M \colon m(u) = \lambda_m u$. Notice that the map $\lambda \colon M \to k$ given by $m \mapsto \lambda_m$ is linear.

Let $u_0 \coloneqq u$ and inductively set $u_{i+1} \coloneqq x(u_i)$. Define $U_i \coloneqq \langle u_0, \ldots, u_i \rangle$. Let n be the smallest natural number such that $u_0, \ldots, u_n$ are linearly dependent.

We will now prove that if $m \in M$ and $i < n$, then $m(u_i) \equiv \lambda_m u_i \pmod{U_{i-1}}$. Note that this implies $M(U_i) \subseteq U_i$.

We prove this by induction on i. It is true for $i = 0$ by definition.

Next, assume it is true for $i > 0$ and $M(U_i) \subseteq U_i$. If $m(u_i) \equiv \lambda_m u_i \pmod{U_{i-1}}$, then $x(m(u_i)) \equiv \lambda_m x(u_i) = \lambda_m u_{i+1} \pmod{U_i}$ (just write out the previous relation and apply x to both sides).

Therefore,

$$m(u_{i+1}) = m(x(u_i)) = ([m, x] + x \circ m)(u_i) \equiv \lambda_m u_{i+1} \pmod{U_1},$$

using the previous calculation and the fact that $[m, x] \in M$ (since M is an ideal) and $M(U_i) \subseteq U_i$. This completes the proof of the claim.

Using the claim, we see that $M(U_{n-1}) \subseteq U_{n-1}$. On the other hand, $x(U_{n-1}) \subseteq U_{n-1}$. This means that $\mathfrak{L}(U_{n-1} \subseteq U_{n-1}$, but we halso have $x(U_{n-1} \subseteq U_{n-1}$ (by linear dependence of $u_0, \ldots, u_n$). Moreover, with respect to the basis $u_0, \ldots, u_{n-1}$, the action of M is represented by upper triangular matrices (since $M(U_i) \subseteq U_i$ with diagonal entries λ_m (by the formula modulo U_{i-1}. In particular, this is true for $m \in \mathfrak{L}^{(1)} \subseteq M$.

But matrices representing elements of $\mathfrak{L}^{(1)}$ must have trace 0 (since $\operatorname{tr} XY = \operatorname{tr} YX$). So $n\lambda_m = 0$ for $m \in \mathfrak{L}^{(1)}$. Since $\operatorname{char} k = 0$, we conclude that $\lambda_m = 0$ for $m \in \mathfrak{L}^{(1)}$.

We now claim that for $i < n$ and $m \in M$ we actually have $m(u_i) = \lambda_m u_i$ (compare this to the previous claim).

We will prove this again by induction (again the base case is trivial). For the inductive step, assume that $m(u_i) = \lambda_m u_i$ for all $m \in M$.

Then

$$m(u_{i+1}) = m(x(u_i)) = ([m, x] + x \circ m)(u_i) = x(m(u_i)) = \lambda_m u_{i+1}$$

because λ is linear and $\lambda_{[m,x]} = 0$, finishing the proof of the claim.

So now we know that $m(w) = \lambda_m w$ for all $m \in M$ and $w \in U_{n-1}$. On the other hand, $x(U_{n-1}) \subseteq U_{n-1}$ (by linear dependence). Choose an eigenvector $0 \neq v \in U_{n-1}$ of the restriction of x to U_{n-1}, say $x(v) = \lambda_x v$. Thus v is a common eigenvector for M (see beginning of this paragraph) and x, and therefore for all of $\mathfrak{L}$, since $\mathfrak{L}$ is spanned by M and x. This completes the proof. □

Corollary 2.14 (Corollary of Engel and Lie).

(a) If $\mathfrak{L}$ satisfies the condition of Engel, then we can pick a basis that defines an isomorphism $\operatorname{End} V \to M_n(k)$ such that $\mathfrak{L}$ maps to a Lie subalgebra of $\mathfrak{n}_n$.

(b) If $\mathfrak{L}$ satisfies the condition of Lie, then we can pick a basis that defines an isomorphism $\operatorname{End} V \to M_n(k)$ such that $\mathfrak{L}$ maps to a Lie subalgebra of $\mathfrak{b}_n$.

Proof. We will prove both parts at the same time by induction on $\dim V$.

By (2.12) and (2.13) we can pick a common eigenvector v_1 of $\mathfrak{L}$.

Then $\mathfrak{L}(\langle v_1 \rangle) \subseteq \langle v_1 \rangle$. Define $V_1 \coloneqq \langle v_1 \rangle$. Define $\overline{\mathfrak{L}} \coloneqq \{\overline{x} \mid x \in \mathfrak{L}\} \subseteq \operatorname{End}(V/V_1)$ where $\overline{x}(v + V_1) = x(v) + V$ for $x \in \mathfrak{L}, v \in V$. This definition makes sense because V_1 is invariant under the action of $\mathfrak{L}$.

$\overline{L}$ inherits the properties of $\mathfrak{L}$. By the inductive hypothesis, $\overline{\mathfrak{L}}$ is represented by (strictly) upper triangular matrices with regard to the basis $v_2 + V_1, \ldots, v_n + V_2$ of V/V_1. Then $v_1, \ldots, v_n$ is a basis of V with respect to which $\mathfrak{L}$ is represented by (strictly) upper triangular matrices . □

COROLLARY 2.15. If $\mathfrak{L}$ satisfies the condition of Engel, then $\mathfrak{L}$ is nilpotent as a Lie algebra.

DEFINITION 2.16. (a) The idealiser of a subset S of $\mathfrak{L}$ is

$$\mathrm{Id}_{\mathfrak{L}}(S) = \{y \in \mathfrak{L} \mid [y, S] \subseteq S\}.$$

If S is a Lie subalgebra of $\mathfrak{L}$, then $\mathrm{Id}_L(S)$ is also a Lie subalgebra. Furthermore, we have $S \subseteq \mathrm{Id}_L(S)$.

(b) We say that $\mathfrak{L}$ satisfies the idealiser condition if every proper Lie subalgebra of $\mathfrak{L}$ is properly contained in its idealiser.

REMARK. A note on terminology: some people, for example Serre, use the term normaliser instead of idealiser.

LEMMA 2.17. If $x \in \mathfrak{L} \subseteq \operatorname{End} V$ and $x^m = 0$, then $(\mathrm{ad}(x))^{2m} = 0$ in $\operatorname{End} \mathfrak{L}$.

PROOF. We may assume that $\mathfrak{L} = \operatorname{End} V$. Let $\theta\colon \operatorname{End} V \to \operatorname{End} V$ denote premultiplication my x, i.e., $y \mapsto x \circ y$. Similarly, let ϕ denote postmultiplication, i.e., $y \mapsto y \circ x$. Notice that $\mathrm{ad}(x) = \theta - \phi$. The maps θ and φ commute, and $\theta^m = 0 = \phi^m$. Therefore,

$$(\mathrm{ad}(x))^2 m = (\theta - \varphi)^{2m} = 0$$

by the binomial theorem. □

REMARK. Given such a basis, define $V_i \coloneqq \langle v_1, \ldots, v_i \rangle$. This gives a chain

$$0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V$$

where $n = \dim V$. Note that $\dim V_i = i$.

DEFINITION 2.18. Such a chain of subspaces of an n-dimensional vector space V is called a maximal flag.

Dropping the condition that $\dim V_i = i$ and allowing fewer terms in the chain, gives the definition of flag.

LEMMA 2.19. The sum of two soluble ideals of $\mathfrak{L}$ is soluble.

PROOF. Let J_1 and J_2 be soluble ideals. Then $J_1 + J_2$ is an ideal (TODO: check this) of $\mathfrak{L}$. So $(J_1 + J_2)/J_1$ is an ideal of $\mathfrak{L}/J_1$ and is the image of J_2 under the canonical map $\mathfrak{L} \to \mathfrak{L}/J_1$. So $(J_1 + J_2)/J_1$ is soluble. Now use 2.9(ii) to see $J_1 + J_2$ is soluble. □

DEFINITION 2.20. The radical $R(\mathfrak{L})$ of $\mathfrak{L}$ is the maximal soluble ideal of $\mathfrak{L}$. By the previous lemma, it is the sum of all soluble ideals of $\mathfrak{L}$.

REMARK. Recall that we call $\mathfrak{L}$ seimisimple if $R(\mathfrak{L}) = 0$. Note that $R(\mathfrak{L}/R(\mathfrak{L})) = 0$, since a soluble ideal of $\mathfrak{L}/R(\mathfrak{L})$ would pull back to give an ideal $R(\mathfrak{L}) \subsetneq J$ for which $J/R(\mathfrak{L})$, so by 2.9 J would a soluble, a contradiction. Thus, $\mathfrak{L}/R(\mathfrak{L})$ is semisimple.

THEOREM 2.21 (Levi). If $\mathrm{char}\, k = 0$ and $\mathfrak{L}$ is finite-dimensional, then there is a Lie subalgebra $\mathfrak{L}_1$ such that $\mathfrak{L}_1 \cap R(\mathfrak{L}) = 0$ and $\mathfrak{L} = \mathfrak{L}_1 + R(\mathfrak{L})$.

Thus $\mathfrak{L}_1 \cong \mathfrak{L}/R(\mathfrak{L})$ is semisimple

NOT PROVED IN THIS COURSE. □

Definition 2.22. This process of splitting a Lie algebra in a soluble part and a semisimple part is called Levi decomposition. The subalgebra $\mathfrak{L}_1$ is called the Levi subalgebra or the Levi factor of $\mathfrak{L}$.

Example.

(1) $\mathfrak{L} = \mathfrak{gl}_2$. Then $R(\mathfrak{L}) = Z(L)$, where $Z(L)$ are the matrices of the form λI. Indeed, $\mathfrak{L}/R(\mathfrak{L}) \cong \mathfrak{sl}_2$ is semisimple (TODO: why?)

    By Levi's theorem, we find that $\mathfrak{L} = \mathfrak{sl}_2 + Z(\mathfrak{L})$, and $\mathfrak{sl}_2$ is the Levi subalgebra of $\mathfrak{gl}_2$.

(2) Let $\mathfrak{L}$ be the subalgebra of $\mathfrak{gl}_4$ consisting of matrices of the form

$$\begin{pmatrix} \mathfrak{sl}_2 & \star \\ 0 & \mathfrak{sl}_2 \end{pmatrix}.$$

    Then $R(\mathfrak{L})$ consists of matrices of the form

$$\begin{pmatrix} 0 & \star \\ 0 & 0 \end{pmatrix}.$$

    This is soluble, and in fact nilpotent. The Levi subalgebra consists of matrices of the form

$$\begin{pmatrix} \mathfrak{sl}_2 & 0 \\ 0 & \mathfrak{sl}_2 \end{pmatrix}.$$

    So $\mathfrak{L}_1 \cong \mathfrak{sl}_2 \times \mathfrak{sl}_2$.

(3) Let $\mathfrak{L}$ be the subalgebra of $\mathfrak{gl}_4$ consisting of matrices of the form

$$\begin{pmatrix} \mathfrak{gl}_2 & \star \\ 0 & \mathfrak{gl}_2 \end{pmatrix}.$$

    Then $R(\mathfrak{L})$ consists of matrices of the form

$$\begin{pmatrix} \lambda I & \star \\ 0 & \mu I \end{pmatrix},$$

    which is soluble but not nilpotent.

    Now we have $\mathfrak{L}/R(\mathfrak{L}) \cong \mathfrak{gl}_2/\{\lambda I\} \times \mathfrak{gl}_2/\{\mu I\} \cong \mathfrak{sl}_2 \times \mathfrak{sl}_2$. So the Levi subablegra is the same as in the previous example.

CHAPTER 3

Invariant forms and the Cartan-Killing criteria

Definition 3.1. A symmetric bilinear form $\langle\ ,\ \rangle\colon \mathfrak{L}\times\mathfrak{L}\to k$ is invariant if $\langle [x,y],z\rangle = \langle x,[y,z]\rangle$.

Definition 3.2. (a) If $\rho\colon \mathfrak{L}\to \operatorname{End} V$ for $\dim V<\infty$ is a Lie algebra representation, then

$$\langle x,y\rangle_p = \operatorname{tr}(\rho(x)\circ\rho(y))$$

is called the trace form of ρ.

(b) The trace form of the adjoint representation of $\mathfrak{L}$ for $\dim\mathfrak{L}<\infty$ is called the Killing form.

Lemma 3.3. (i) The trace form of a representation is an invariant symmetric bilinear form.

(ii) If J is a Lie ideal of $\mathfrak{L}$, then $J^\perp = \{x \mid \forall y\in J\colon \langle x,y\rangle = 0\}$ is an ideal of $\mathfrak{L}$ for any invariant form $\langle\ ,\ \rangle$.

In particular, $\mathfrak{L}^\perp$ is an ideal of $\mathfrak{L}$.

Proof. Symmetry follows from $\operatorname{tr} x\circ y = \operatorname{tr} y\circ x$. Bilinearity is immediate. For $x,y,z\in\mathfrak{L}$, we have

$$\begin{aligned}
\langle [x,y],z\rangle &= \operatorname{tr}(\rho([x,y])\circ\rho(z))\\
&= \operatorname{tr}([\rho(x),\rho(y)]\circ\rho(z))\\
&= \operatorname{tr}(\rho(x)\circ\rho(y)\circ\rho(z)) - \operatorname{tr}(\rho(y)\circ\rho(x)\circ\rho(z))\\
&= \operatorname{tr}(\rho(x)\circ\rho(y)\circ\rho(z)) - \operatorname{tr}(\rho(x)\circ\rho(z)\circ\rho(y))\\
&= \operatorname{tr}(\rho(x)\circ[\rho(y),\rho(z)])\\
&= \operatorname{tr}(\rho(x)\circ\rho([y,z]))\\
&= \langle x,[y,z]\rangle,
\end{aligned}$$

so the trace form is invariant[1]. This completes the proof of (i).

Next, let J be a Lie ideal. Let $x\in J^\perp$, $y\in\mathfrak{L}$. We will show that $[x,y]\in J^\perp$. Indeed, let $z\in J$. Then $[y,z] = -[z,y]\in J$ since J is a Lie ideal. But then $\langle [x,y],z\rangle = \langle x,[y,z]\rangle = 0$ since $x\in J^\perp$ and we are done. $\square$

Remark. There may be invariant forms on $\mathfrak{L}$ which are not the trace form of any representation.

Theorem 3.4 (Cartan's criterion for solubility). Assume that $\operatorname{char} k = 0$ and $\mathfrak{L}$ is a Lie subalgebra of $\operatorname{End} V$. Let $\langle\ ,\ \rangle$ be the trace form of the inclusion $\mathfrak{L}\to\operatorname{End} V$.

Then $\mathfrak{L}$ is soluble if and only if $\langle x,y\rangle = 0$ for all $x\in\mathfrak{L}$, $y\in\mathfrak{L}^{(1)}$, i.e., $\mathfrak{L}^{(1)}\subseteq\mathfrak{L}^\perp$.

Proof. We will only do the case $k=\mathbb{C}$. In general, we can embed any k of characteristic zero into an algebraically closed field and obtain the result from that (with some work).

[1] Note that we even have $\langle [x,y],z\rangle = 0 = \langle x,[y,z]\rangle$.

Assume first that L is soluble. By the corollary of Lie, there is a basis of V with regard to which L is represeted by upper triangular matrices, i.e., $L \subseteq \mathfrak{b}_n$. Hence, $L^{(1)} \subseteq \mathfrak{n}_n$. Hence, $\operatorname{tr}(xy) = 0$ for all $x \in L$, $y \in L^{(1)}$ since xy is triangular with 0s on the diagonal.

Conversely, it suffices to show that $L^{(1)}$ is nilpotent, hence soluble. By Engel (and its corollary), it will suffice to show that all elements in $L^{(1)}$ are nilpotent. Define $A = L^{(1)}$, $B = L$ and apply lemma 3.12. We have $T = \{t \in \operatorname{End} V \mid [t, L] \subseteq L^{(1)}\}$. Note that $L^{(1)} \subseteq L \subseteq T$. $L^{(1)}$ is spanned by $[x, z]$, $x, z \in L$. Let $t \in T$. Then

$$tr([x, z] \circ t) = \operatorname{tr}(x \circ [z, t]),$$

where $[z, t] \in L^{(1)}$ by definition of T, hence $\operatorname{tr}([x, z] \circ t) = 0$. Thus, $\operatorname{tr}(wt) = 0$ for all $w \in L^{(1)}, t \in T$. But $L^{(1)} \subseteq T$, so by the lemma every element in $L^{(1)}$ is nilpotent. □

THEOREM 3.5 (Cartan-Killing criterion for semisimplicity). Let $\operatorname{char} k = 0$. The following are equivalent for a finite-dimensional Lie algebra $\mathfrak{L}$:

(1) $\mathfrak{L}$ is semisimple,
(2) The Killing form $\langle\ ,\ \rangle_{\mathrm{ad}}$ is non-degenerate.

PROOF. We have

$$\mathfrak{L}^{\perp} = \{x \mid \forall y \in \mathfrak{L}\colon \operatorname{tr}(\operatorname{ad}(x) \circ \operatorname{ad}(y)) = 0\}.$$

Suppose J is an abelian ideal of $\mathfrak{L}$. Then $x \in \mathfrak{L}$, $y \in J$. Then $\operatorname{ad}(y)(\mathfrak{L}) \subseteq J$, so $\operatorname{ad}(x) \circ \operatorname{ad}(y)(\mathfrak{L}) \subseteq J$. Both times, we use that J is an ideal.

Since J is abelian, $\operatorname{ad}(y)(J) = 0$, hence $(\operatorname{ad}(x) \circ \operatorname{ad}(y))^2(\mathfrak{L}) = 0$. This means that $\operatorname{ad}(x) \circ \operatorname{ad}(y)$ is nilpotent in $\operatorname{End} \mathfrak{L}$ and therefore has zero trace[2]. But if $x \in \mathfrak{L}$, $y \in J$, then

$$\langle x, y\rangle_{\mathrm{ad}} = \operatorname{tr}(\operatorname{ad}(x) \circ \operatorname{ad}(y)) = 0,$$

so $y \in \mathfrak{L}^{\perp}$. Hence $J \subseteq \mathfrak{L}^{\perp}$.

Now, if $R(\mathfrak{L}) \neq 0$, then it contains a nonzero abelian ideal of $\mathfrak{L}$, for example the last nonzero term of the derived series of $R(\mathfrak{L})$.

Hence, if the Killing form is nondegenerate (this is the same as saying that $\mathfrak{L}^{\perp} = 0$), then $\mathfrak{L}$ must be semisimple, since otherwise we would have $R(\mathfrak{L}) \neq 0$, so we find a nonzero abelian ideal J which by what we have seen above is contained in $\mathfrak{L}^{\perp} = 0$, a contradiction.

Conversely, suppose $\mathfrak{L}$ is semisimple. Then $R(\mathfrak{L}) = 0$ and $J = L^{\perp}$ an ideal of $\mathfrak{L}$. Consider $\operatorname{ad}_{\mathfrak{L}}\colon \mathfrak{L} \to \operatorname{End} \mathfrak{L}$ and the image $\operatorname{ad}(J) \subseteq \operatorname{End} \mathfrak{L}$. By definition of J, we have $\operatorname{tr}(\operatorname{ad}(x) \circ \operatorname{ad}(y) = 0$ for all $x \in J, y \in \mathfrak{L}$.

In particular, $\operatorname{tr}(\operatorname{ad}(x) \circ \operatorname{ad}(y)) = 0$ for $x, y \in J$. By Cartan's solubility criterion, $\operatorname{ad}_{\mathfrak{L}}(J)$ is a soluble subalgebra of $\operatorname{End} \mathfrak{L}$.

On the other hand, $\ker \operatorname{ad}_{\mathfrak{L}} = Z(\mathfrak{L})$ is the centre of $\mathfrak{L}$ and an abelian ideal of $\mathfrak{L}$, hence soluble, so 2.9(ii) gives that J is soluble. Therefore, $J \subseteq R(\mathfrak{L}) = 0$, so $J = 0$. But since $J = \mathfrak{L}^{\perp}$, the Killing form is nondegenerate. □

DEFINITION 3.6. A derivation of a Lie algebra is a k-linear map $D\colon \mathfrak{L} \to \mathfrak{L}$ such that $D([x, y]) = [x, D(y)] + [D(x), y]$.

An inner derviation is of the form $y \mapsto [x, y]$. In other words, it is ad_x for some x.

The derivations of $\mathfrak{L}$ form a Lie subalgebra $\operatorname{Der} \mathfrak{L} \subseteq \operatorname{End} \mathfrak{L}$, and $\operatorname{ad}(\mathfrak{L})$ is a Lie ideal of $\operatorname{Der} \mathfrak{L}$.

[2] Any eigenvalue must be zero, and we can put the matrix in Jordan normal form.

THEOREM 3.7. If $\operatorname{char} k = 0$ and $\mathfrak{L}$ is a finite-dimensional semisimple Lie algebra, then $\operatorname{Der}\mathfrak{L} = \operatorname{ad}_{\mathfrak{L}}$.

Since $\mathfrak{L}$ is semisimple and the kernel of the map $\mathfrak{L} \to \operatorname{ad}_{\mathfrak{L}}$ is an abelian ideal, it must be zero (since it is trivially soluble), so we additionally get $\operatorname{ad}_{\mathfrak{L}} \cong \mathfrak{L}$.

PROOF. Let D be a derivation of $\mathfrak{L}$ and $x \in \mathfrak{L}$. Then for every $y \in \mathfrak{L}$ we have

$$\begin{aligned}[D, \operatorname{ad}_{\mathfrak{L}}(x)](y) &= (D \circ \operatorname{ad}_{\mathfrak{L}}(x) - \operatorname{ad}_{\mathfrak{L}}(x) \circ D)(y) \\ &= D([x, y]) - [x, D(y)] \\ &= [D(x), y] + [x, D(y)] - [x, D(y)] \\ &= [D(x), y] \\ &= \operatorname{ad}_{\mathfrak{L}}(D(x))(y),\end{aligned}$$

so we conclude that

$$(\star) \qquad [D, \operatorname{ad}(x)] = \operatorname{ad}(D(x)).$$

The centre $Z(\mathfrak{L})$ of $\mathfrak{L}$ is an abelian ideal, hence zero (since $\mathfrak{L}$ is semisimple)

Since $\mathfrak{L}$ is semisimple and the kernel of the map $\mathfrak{L} \to \operatorname{ad}_{\mathfrak{L}}$ is an abelian ideal, it must be zero (since it is trivially soluble), hence $\mathfrak{L} \cong \operatorname{ad}(L)$.

Let $\langle\ ,\ \rangle$ denote the Killing form on $\operatorname{Der}\mathfrak{L}$. By question 13 from the example sheet, the restriction of $\langle\ ,\ \rangle$ to $\operatorname{ad}(\mathfrak{L})$ is the Killing form on $\operatorname{ad}(\mathfrak{L})$.

Let J be the orthogonal space to $\operatorname{ad}(\mathfrak{L})$ inside $\operatorname{Der}(\mathfrak{L})$ with respect to $\langle\ ,\ \rangle$. By 3.3(ii) J is an ideal of $\operatorname{Der}\mathfrak{L}$. Now, since $\mathfrak{L}$ is semisimple, so is $\operatorname{ad}(\mathfrak{L})$, and by the Cartan-Killing criterion, $\langle\ ,\ \rangle$ restricted to $\operatorname{ad}(\mathfrak{L})$ is non-degenerate. Hence $\operatorname{ad}(\mathfrak{L}) \cap J = 0$ and $[\operatorname{ad}(\mathfrak{L}), J] \subseteq \operatorname{ad}(\mathfrak{L}) \cap J = 0$, since both are ideals.

Thus if $D \in J$, then for all $x \in \mathfrak{L}$ we have $\operatorname{ad}(D(x)) = 0$ by $(\star)$. Thus, $D(x) \in Z(\mathfrak{L}) = 0$, since $\mathfrak{L}$ is semisimple, so D is the zero derivation, and we conclude $J = 0$. This can only happen if $\operatorname{Der}(\mathfrak{L}) = \operatorname{ad}(\mathfrak{L})$ (by linear algebra) and so we are done. □

REMARK.
1. $\operatorname{Der}\mathfrak{L} = \operatorname{ad}_{\mathfrak{L}}$ is the same as saying that the first Lie algebra cohomology group of $\mathfrak{L}$, which is isomorphic to $\operatorname{Der}\mathfrak{L}/\operatorname{ad}(\mathfrak{L})$ vanishes when $\mathfrak{L}$ is semisimple.
2. If $\mathfrak{L}$ is nonzero and nilpotent, then $\operatorname{Der}\mathfrak{L}/\operatorname{ad}(\mathfrak{L})$. This is question 17 on the example sheet.
3. There are some soluble non-nilpotent $\mathfrak{L}$ where $\operatorname{Der}\mathfrak{L}/\operatorname{ad}(\mathfrak{L}) = 0$. This is question 16 on the example sheet.

EXERCISE. For a general finite-dimensional Lie algebra $\mathfrak{L}$ with an invariant form, we have

$$[R(\mathfrak{L}), R(\mathfrak{L})] \subseteq L^{\perp} \subseteq R(\mathfrak{L}),$$

but $R(\mathfrak{L})$ and $\mathfrak{L}^{\perp}$ need not be equal.

SOLUTION. TODO. □

DEFINITION 3.8. An endomorphism $x \in \operatorname{End} V$ is called semisimple if it is diagonalisable, which is equivalent to the minimal polynomial being the product of distinct linear factors.

REMARK.
1. If an endomorphism x is semisimple and W is a subspace such that $x(W) \subseteq W$ then $x|_W \colon W \to W$ is semisimple, since the minimal polynomial divides the minimal polynomial of w.
2. If x, y are semisimple endomorphisms and $x \circ y = y \circ x$, then x, y can be simultaneously diagonalised, and so $x \pm y$ is semisimple.

LEMMA 3.9 (Jordan decomposition of an endomorphism). Let x be an endomorphism.

(i) There are unique endomorphisms x_s and x_n such that x_s is semisimple, x_n is nilpotent, x_s and x_n commute and $x = x_s + x_n$.
(ii) There are unique polynomials p, q with zero constant term such that $x_s = p(x)$, $x_n = q(x)$. Hence x_s, x_n commute with all endomorphisms that commute with x.
(iii) If $U \subseteq V \subseteq X$ such that $x(W) \subseteq U$, then $x_s(W) \subseteq U$ and $x_n(W) \subseteq U$.

Proof. (iii) is an immediate consequence of (ii).

Let $\prod(t-\lambda_i)_i^m$ be the characteristic polynomial of x.

Define $V_i \coloneqq \ker(x - \lambda_i\iota)^{m_i}$ to be the generalized eigenspace, where ι is the identity. By linear algebra, we have $V = \bigoplus V_i$. The characteristic polynomial of $x|_{V_i}$ is $(t-\lambda_i)^{m_i}$.

Our goal is to find a polynomial p such that $p \equiv 0 \pmod t$ and $p \equiv \lambda_i \pmod{(t-\lambda_i)^{m_i}}$ for each i. By the Chinese Remainder Theorem, such a polynomial exists. Define $q(t) = t - p(t)$. Now set $x_s \coloneqq p(x)$, $x_n \coloneqq q(x)$.

For each i, we have

$$x_s - \lambda_i\iota = p(x) - \lambda_i\iota = r(x)(x-\lambda_i)^{m_i} + \lambda_i\iota - \lambda_i\iota = r(x)(x-\lambda_i)^{m_i},$$

hence $(x_s - \lambda_i\iota)|_{V_i} = 0$, so $x_s|_{V_i} = (\lambda_i\iota)|_{V_i}$, and so x_s is diagonalizable.

Now $(x_n)|_{V_i} = (x - x_s)|_{V_i} = (x - \lambda_i\iota)|_{V_i}$, so by definition of V_i, $x_n|_{V_i}$ is nilpotent for each i. Therefore, x_n is nilpotent.

It remains to show uniqueness of x_s and x_n. If $x = s + n$ with s semisimple and n nilpotent and s and n commute. Then n, s commute with x and with x_s and x_n, which are just polynomials in x. So $n - x_n = s - x_s$ is semisimple by the previous remark and nilpotent. But an endomorphism that is both semisimple and nilpotent must be zero. □

Definition 3.10. The endomorphism x_s is called the semisimple part and x_n is called the nilpotent part of x.

Lemma 3.11. If $x \in L \subseteq \operatorname{End} V$ and $x = x_s + x_n$ is the Jordan decomposition, then $\operatorname{ad}(x_s) = \operatorname{ad}(x)_s$ and $\operatorname{ad}(x_n) = \operatorname{ad}(x)_n$.

Proof. By (2.17), $\operatorname{ad}(x_n)$ is nilpotent. Since x_s and x_n commute with x, $\operatorname{ad}(x_s)$ and $\operatorname{ad}(x_n)$ commute with $\operatorname{ad}(x)$. Since $\operatorname{ad}(x) = \operatorname{ad}(x_s) + \operatorname{ad}(x_n)$, it remains to show that $\operatorname{ad}(x_s)$ is semisimple.

Since x_s is semisimple, we find a basis $\{v_i\}$ of V consisting of eigenvectors of x_s, i.e., $x_s(v_i) = \lambda v_i$.

Define $\theta_{ij} \in \operatorname{End} V$ via $v_i \mapsto v_j$, and $v_\ell \mapsto 0$ for $\ell \neq i$. The θ_{ij} form a basis of $\operatorname{End} V$ corresponding to elementary matrices.

Note that $x_s\theta_{ij}(v_i) = \lambda_j v_j$ and $x_s\theta_{ij}(v_\ell) = 0$ for $\ell \neq i$. On the other hand, $\theta_{ij}x_s(v_i) = \lambda_i v_j$ and $\theta_{ij}x_s(v_\ell) = 0$ if $\ell \neq i$.

Thus, $\operatorname{ad}(x_s)(\theta_{ij} = (\lambda_j - \lambda_i)\theta_{ij}$, so the θ_{ij} form a basis of eigenvectors of $\operatorname{ad}(x_s)\colon \operatorname{End} V \to \operatorname{End} V$.

Hence $\operatorname{ad}(x_s)\colon \operatorname{End} V \to \operatorname{End} V$ is diagonalisable, hence its restriction to L is diagonalisable as well, completing the proof. □

Remark. If L is semisimple, then $Z(L) \subseteq R(L) = 0$, since $Z(L)$ is an abelian ideal, so $L \cong \operatorname{ad}(L) \subseteq \operatorname{End} L$ and so we can say that $x \in L$ is semisimple/nilpotent according to whether $\operatorname{ad}(x)$ is semisimple or nilpotent.

Lemma 3.12. Let A and B be subspaces of $\operatorname{End} V$ with $A \subseteq B$. Define $T \coloneqq \{t \in \operatorname{End} V \mid [t, B] \subseteq A\}$.

Let $w \in T$ and suppose that for all $t \in T$ we have $\operatorname{tr}(wt) = 0$. Then w is nilpotent.

Proof. Compute the Jordan decomposition $w = w_s + w_n$. Our goal is to show that $w_s = 0$. Take a basis $\{v_i\}$ of eigenvectors of w_s such that $w_s(v_i) = \lambda_i v_i$.

Define θ_{ij} as in the previous proof. Again we have $\mathrm{ad}(w_s)(\theta_{ij}) = (\lambda_j - \lambda_i)\theta_{ij}$

Assume that $w_s \neq 0$, so there is some j such that $\lambda_j \neq 0$. Let E be the $\mathbb{Q}$-span of $\lambda_i, \ldots, \lambda_n$. Choose any non-zero linear form $f\colon E \to \mathbb{Q}$.

Define $y \in \operatorname{End} V$ via $y(v_i) \coloneqq f(\lambda_i)v_i$. So

$$\mathrm{ad}(y)(\theta_{ij}) = (f(\lambda_j) - f(\lambda_i))\theta_{ij} = f(\lambda_j - \lambda_i)\theta_{ij}$$

by linearity of f.

Let $r(t)$ be a polynomial with vanishing constant term such that

$$r(\lambda_j - \lambda_i) = f(\lambda_j - \lambda_i)$$

for all i, j. The polynomial r exists by polynomial interpolation.

Then

$$\begin{aligned} r(\mathrm{ad}(w_s))(\theta_{ij}) &= \sum_{\ell=0}^{\deg q} q_\ell \,\mathrm{ad}(w_s)^\ell(\theta_{ij}) \\ &= \sum_{\ell=0}^{\deg q} q_\ell(\lambda_j - \lambda_i)^\ell \theta_{ij} \\ &= r(\lambda_j - \lambda_i)\theta_{ij} \\ &= f(\lambda_j - \lambda_i)\theta_{ij} \\ &= \mathrm{ad}(y)(\theta_{ij}), \end{aligned}$$

so $\mathrm{ad}(y) = r(\mathrm{ad}(w_s))$.

By 3.9(ii) and 3.11, the semisimple part of $\mathrm{ad}(w)$ is a polynomial in $\mathrm{ad}(w)$ with zero constant term. So $\mathrm{ad}(y)$ is also such a polynomial. However $w \in T$, so $[w, B] \subseteq A$, which means that $\mathrm{ad}(w)(B) \subseteq A$, and we conclude $\mathrm{ad}(y)(B) \subseteq A$. By definition of T, we have $y \in T$. By assumption, $\mathrm{tr}(wt) = 0$ for all $t \in T$. In particular, $0 = \mathrm{tr}(wy) = \sum \lambda_i f(\lambda_i)$. Recall that $f(\lambda_i) \in \mathbb{Q}$. But f is linear, so applying f we get $\sum f(\lambda_i)^2 = 0$. Hence, $f(\lambda_i) = 0$ for all i, but since the λ_i span E, f is identically zero, a contradiction.

Hence, we must have $w_s = 0$. ☐

Propositton 3.13. Let L be a finite-dimensional Lie algebra in characteristic zero.

(i) If L is semisimple, then L is a direct sum of non-abelian simple ideals.
(ii) If $0 \neq J$ is an ideal of $L = \bigoplus L_i$, then J is a direct sum of a subset of the L_i.
(iii) If L is a direct sum of nonabelian simple ideals, then L is semisimple.

Proof. We will prove part (i) by induction on $\dim L$. Let J be an ideal of the semisimple Lie algebra L. By the Cartan-Killing criterion, the Killing form is non-degenerate. Consider the orthogonal space $J^\top$, which is an ideal. We have $\dim J + \dim J^\top = \dim L$.

By Cartan's criterion applied to $\mathrm{ad}(J \cap J^\top)$, we find that $\mathrm{ad}(J \cap J^\top)$ is soluble. Hence $J \cap J^\top$ is an ideal and it is soluble, since the kernel of the adjoint representation is an abelian ideal. We conclude $J \cap J^\perp \subseteq R(L) = 0$. By a dimension argument, we conclude $L = J \oplus J^\top$. Note that any ideal of J or in $J^\top$ is also an ideal of L (because J is a direct summand of L). Thus, J and $J^\perp$ are also semisimple. Since $L \neq J \neq 0$, J and $J^\perp$ are direct sums of non-abelian simple ideals. This completes the proof of part (i).

For part (ii), suppose $J \cap L_i = 0$. Then $[L_i, J] = 0$, since J and L_i are ideals and hence $J \subseteq \bigoplus_{j \neq i} L_j$.

Conversly, if $J \cap L_j \neq 0$, then by simplicity of L_i we have $L_i \subseteq J$. Hence $J = \bigoplus_{L_i \subseteq J} L_i$.

For part (iii), assume at L is a direct sum of non-abelian simples. By (ii), the ideal $R(L)$ is the direct sum of some of them. However, $R(L)$ is soluble and so cannot contain nonabelian simple ideals. Hence $R(L) = 0$, so L is semisimple. □

Remark. Almost everybody define simple Lie algebras to be nonabelian, i.e., they exclude the case of the one-dimensional Lie algebra.

According to the definition used here, we have that L is simple if and only if ad is irreducible.

Example. Assume that $\operatorname{char} k \neq 0$. We will show that $\mathfrak{sl}_2$ is simple. We have met irreducible representations of $\mathfrak{sl}_2$. In particular $\mathfrak{sl}_2 \to \operatorname{End}(V_n)$, where V_n are the homogenous polynomials in two variables of degree n. We noted that when $n = 2$, then this is just the adjoint representation. Thus the adjoint representation is irreducible, hence $\mathfrak{sl}_2$ is simple.

CHAPTER 4

Cartan Subalgebras

Remark. In this chapter, L is a finite-dimensional Lie algebra over $k = \mathbb{C}$.

Definition 4.1. For $0 \neq y \in L$, $\lambda \in \mathbb{C}$ define

$$L_{\lambda,y} \coloneqq \{x \in L \mid \exists r > 0\colon (\mathrm{ad}(y) - \lambda\iota)^r x = 0\}.$$

This is called the generalised λ-eigenspace for $\mathrm{ad}(y)$.

Remark. Note that for all y we have $y \in L_{0,y}$ since $[y, y] = 0$. We write $L_{\lambda,y} = 0$ if λ is not an eigenvalue of $\mathrm{ad}(y)$.

Lemma 4.2.
- (i) We have $[L_{\lambda,y}, L_{\mu,y}] \subseteq L_{\lambda+\mu,y}$. In particular $L_{0,y}$ is a subalgebra of L.
- (ii) We have $L = \bigoplus L_{\lambda,y}$, summing over the eigenvalues of $\mathrm{ad}(y)$.
- (iii) If $L_{0,y}$ is contained in a subalgebra A of L, then $\mathrm{Id}_L(A) = A$. In particular, $L_{0,y} = \mathrm{Id}(L_{0,y})$.

Proof. For part (i), note that using the fact that adjoints are derivations we have

$$(\mathrm{ad}(y) - (\lambda + \mu)\iota)([x, z]) = [(\mathrm{ad}(y) - \lambda\iota)x, z] + [x, (\mathrm{ad}(y) - \mu\iota)z]$$

and so

$$(\mathrm{ad}(y) - (\lambda + \mu)\iota)^n([x, z]) = \sum_{i+j=n} \binom{n}{i} [(\mathrm{ad}(y) - \lambda\iota)^i(x), (\mathrm{ad}(y) - \mu\iota)^j(z)]$$

Hence if $x \in L_{\lambda,y}$, $z \in L_{\mu,y}$, then $[x, z] \in L_{\lambda+\mu,y}$.

Part (ii) is just standard linear algebra about generalized eigenspaces.

For part (iii), we have already noticed that $[y, y] = 0$ and so $\mathrm{ad}(y)$ has 0 as an eigenvalue. Hence the characteristic polynomial of $\mathrm{ad}(y)$ can be written as $t^m f$ with $m \geq 1$ and $t \nmid f$. By coprimality we find polynomials q, r such that $1 = qt^m + rf$.

Let $b \in \mathrm{Id}_L(A)$. Then

$$b = q(\mathrm{ad}(y))(\mathrm{ad}(y))^m(b) + r(\mathrm{ad}(y))f(\mathrm{ad}(y))(b). \tag{$\star$}$$

But $m \geq 1$ and $y \in A$, so the first term of the RHS of ($\star$) is in A. Also $(\mathrm{ad}(y))^m f(\mathrm{ad}(y))(b) = 0$ by Cayley-Hamilton. So $f(\mathrm{ad}(y))(b) \in L_{0,y} \subseteq A$. Hence the secod term of the RHS of ($\star$) is in A, hence $b \in A$. Thus $\mathrm{Id}(A) = A$. □

Definition 4.3. A Cartan subablgebra (CSA) of L is a nilpotent subalgebra equal to its own idealiser in L.

Theorem 4.4. *H is a minimal subalgebra of the form $L_{0,y}$ with respect to inclusion if and only if H is a Cartan subalgebra of L.*

Proof. Suppose $H = L_{0,z}$ is minimal. We must show that it is nilpotent and equal to its own idealiser. Then $\mathrm{Id}(H) = H$ by 4.2(iii). Take $K = H$ in 4.9 to deduce that $H = L_{0,z} \subseteq L_{0,y}$ for all $y \in H$. Thus $\mathrm{ad}(y)|_H \colon H \to H$ is nilpotent for $y \in H$ since 0 is the only eigenvalue. Hence $\mathrm{ad}(H)$ is nilpotent by (the corollary of) Engel (TODO: why?) and so H is nilpotent (since the quotient by the center (the kernel of ad) is nilpotent). Thus H is a Cartan subalgebra.

Conversely, sat H is a Cartan subalgebra. Then $H \subseteq L_{0,y}$ for all $y \in H$ since H is nilpotent. Suppose we have strict inequality for all y.

Choose $L_{0,z}$ as small as possible with $z \in H$. By 4.9 with $K = H$ we have $L_{0,z} \subseteq L_{0,y}$ for all $y \in H$. But $H \subseteq L_{0,z}$, hence $\mathrm{ad}(H)(L_{0,z}) \subseteq L_{0,z}$. For $y \in H$ we have $L_{0,z} \subseteq L_{0,y}$, hence $\mathrm{ad}(y)$ acts nilpotently on $L_{0,z}$. Hence all elements of $\mathrm{ad}(H)$ act nilpotently on $L_{0,z}/H$. By Engel, there is a common eigenvector $x + H$ with $x \in L_{0,z} \setminus H$ such that $[H, x] \subseteq H$. Therefore, $x \in \mathrm{Id}(H) \setminus H$, but H is a Cartan subalgebra, so we have a contradiction.

Therefore, we must have some z such that $H = L_{0,z}$ for some $z \in H$. Note that H is nilpotent and so satisfies the idealiser condition. But 4.2iii says that $\mathrm{Id}_L(L_{0,y}) = L_{0,y}$ for any y. Hence, no $L_{0,y}$ is a proper subalgebra of H and so we know that $H = L_{0,z}$ is minimal among the $L_{0,y}$ for $y \in H$. □

Definition 4.5. (i) The rank of a Lie algebra L is the minimal dimension of $L_{0,y}$ for $y \in L$.
(ii) y is called regular if the dimension of $L_{0,y}$ is equal to the rank of L.

Corollary 4.6. If y is regular, then $L_{0,y}$ is a CSA.

Proof. Immediate from (4.4). □

Remark. On the face of it, we could have minimal $L_{0,y}$ of different dimensions, but that is not the case.

Theorem 4.7. Any two CSAs are conjugate under the group of automorphisms of L generated by $e^{\mathrm{ad}(y)} = 1 + \mathrm{ad}(j) + \frac{\mathrm{ad}(y)^2}{2!} + \cdots$ for y such that $\mathrm{ad}(y)$ is nilpotent.

Not proved in this course. □

Remark. There is geometry concerning the set of regular elements.

Theorem 4.8. The set of regular elements of L is a connected, Zariski dense and open subset of L.

Not proved in this course. □

Lemma 4.9. Let K be a subalgebra of L and $z \in K$ such that $L_{0,z}$ is minimal in the set $\{L_{0,y} \mid y \in K\}$.

Supppose $K \subseteq L_{0,z}$. Then $L_{0,z} \subseteq L_{0,y}$ for all $y \in K$.

Proof. We start with an observation. Let $\theta, \phi \in \operatorname{End} V$ iand $c \in k = \mathbb{C}$. Suppose $\theta + c\phi$ has characteristic polynomial

$$f(t, c) = t^n + f_1(c)t^{n-1} + \cdots + f_n(c).$$

Then f_i is a polynomial in c of degree at most i. (TODO: why?)

For $y \in K$ consider the set $S := \{\mathrm{ad}(z + cy) \mid c \in \mathbb{C}\}$. Write $H := L_{0,z}$. Each $z + cy \in K \subseteq H$ by hypothesis. Elements of S induce endomrphisms of H and L/H since $\mathrm{ad}(z, cy)(H) \subseteq H$. Write $f(t, c)$ for the characteristic polynomial of $\mathrm{ad}(z + cy)$ on H and $g(t, c)$ for the characteristic polynomial of $\mathrm{ad}(z, cy)$ on L/H. If $\dim L = n$, $\dim L = m$, then

$$\begin{aligned} f(t, c) &= t^m + f_1(c)t^{m-1} + \cdots + f_m(c), \\ g(t, c) &= t^{n-m} + g_1(c)t^{n-m-1} + \cdots + g_{n-m}(c), \end{aligned}$$

where f_i and g_i are polynomials of degree at most i by the initial observation.

But $\mathrm{ad}(z)$ has no zero eigenvalue on L/H since H is the generalized eigenspace. Hence $g_{n-m}(0) \neq 0$, so g_{n-m} is not the zero polynomial. Hence we can find $c_1, \ldots, c_{m+1} \in k$ with $g_{n-m}(c_j) \neq 0$ for each j. Hence $\mathrm{ad}(z + c_j y)$ has no zero eigenvalue on L/H and so $L_{0,z+c_jy} \subseteq H$. But H was chosen to be minimal among $L_{0,y}$ and so $L_{0,z+c_jy} = H$. Therefore, 0 is the only eigenvalue of the map

$$\mathrm{ad}(z + c_jy)|_H : H \to H.$$

This means that $f(t, c_j) = t^m$ for $1 \leq j \leq m + 1$. Therefore $f_i(c_j) = 0$ for $1 \leq j \leq m + 1$, but since $\deg f_i \leq i < m + 1$, this means that f_i is the zero polynomial. Hence $f(t, c) = t^m$ for all $c \in k$, which implies that $H \subseteq L_{0,z+cy}$ for all $c \in \mathbb{C}$.

But $y \in K$ was arbitrary, hence $H \subseteq L_{0,y}$ for any $y \in K$. $\square$

Example. Let $L = \mathfrak{sl}_2$ and recall that we have the basis

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

where $[e, f] = h$, $[h, e] = 2e$, $[h, f] = -2f$.

We have

$$L_{0,h} = \langle h \rangle \qquad L_{2,h} = \langle e \rangle \qquad L_{-2,h} = \langle f \rangle.$$

Notice that $[L_{2,h}, L_{-2,h}] \subseteq L_{0,h}$ (as proved earlier). $L_{0,h} = \langle h \rangle$ is a Cartan subalgebra, as it is clearly minimal. Furthermore, h is a regular element (i.e., it generates the $L_{0,y}$ of smallest dimension), and the rank of $\mathfrak{sl}_2$ is 1.

Theorem 4.10. Let H be a Cartan subalgebra of a semisimple Lie algebra. Then

(a) H is abelian,
(b) the centraliser $Z_L(H) = \{x \in L \mid [x, h] = 0 \ \forall h \in H\}$ is H itself, and so H is a maximal abelian subalgebra,
(c) every element of H is semisimple (i.e., $\mathrm{ad}(x)$ is diagonalizable for every $x \in H$),
(d) the restriction of the Killing form of L to H is non-degenerate (note that this restriction to H is not (!) the Killing form of H, unlike in the case of restricting to an ideal).

Proof. By the classification of Cartan subalgebras, $H = L_{0,y}$ for some $y \in H$.

For part (d), consider the decomposition

$$L = L_{0,y} \oplus \bigoplus_{\substack{\lambda \neq 0 \\ \lambda \text{ eigenvalue of } \mathrm{ad}(y)}} L_{\lambda,y}.$$

Recall $[L_{\lambda,y}, L_{\mu,y}] \subseteq L_{\lambda+\mu,y}$. If $u \in L_{\lambda,y}$, $v \in L_{\mu,y}$ with $\lambda + \mu \neq 0$, then applying $\mathrm{ad}(u)\,\mathrm{ad}(v)$ maps each generalized eigenspace into a different one, so $\mathrm{tr}(\mathrm{ad}(u) \circ \mathrm{ad}(v)) = 0$ (think of matrices). Thus when $\lambda + \mu \neq 0$, the spaces $L_{\lambda,y}$ and $L_{\mu,y}$ are orthogonal with respect to the Killing form. Hence,

$$L = L_{0,y} \oplus (L_{\lambda,y} + L_{-\lambda,y}) \oplus \ldots$$

is an orthogonal direct sum with regards to the Killing form. But by the Cartan-Killing criterion, we know that that Killing form is non-degenerate on L. This means that its restriction to each summand is non-degenerate. In particular, the restriction to $L_{0,y}$ is non-degenerate.

For part (a), we notice that H is nilpotent and so $\mathrm{ad}_L(H)$ is nilpotent and hence soluble, and so we can use Cartan's solubility criterion to find that

$$\mathrm{tr}(\mathrm{ad}(x_1) \circ \mathrm{ad}(x_2) = 0$$

for $x_1 \in H$, $x_2 \in H^{(1)}$. Thus, $H^{(1)}$ is orthogonal to H with respect to the Killing form. By (d), the restriction of the Killing form to H is non-degenerate, so we must have $H^{(1)} = 0$, which just means that H is abelian.

For part (b), obvserve that $H \subseteq Z_L(H) \subseteq \mathrm{Id}_L(H)$, where the first inclusion follows from (a) and the second inclusion is true by definition. But H is a CSA, so $\mathrm{Id}_L(H) = H$, hence we have equality everywhere, so in particular $Z_L(H) = H$. If $H \subseteq A$ is abelian, then certainly $A \subseteq Z_L(H) = H$, and H is indeed maximal abelian.

Finally, for part (c), recall that L is semisimple, so ad_L is injective. If $x \in H$, then we have a Jordan decomposition $\mathrm{ad}\, x = \mathrm{ad}(x)_s + \mathrm{ad}(x)_n$. Now it can be shown (see Humphreys, Lemma 4.2.B), that $\mathrm{ad}(x)_s$ and $\mathrm{ad}(x)_n$ are derivations. Since L is semisimple, by Theorem 3.7 we have that $\mathrm{ad}(x)_s, \mathrm{ad}(x)_n \in \mathrm{ad}(L)$, hence we find $x_s, x_n \in L$ such that $\mathrm{ad}(x)_s = \mathrm{ad}(x_s)$, $\mathrm{ad}(x)_n = \mathrm{ad}(x_n)$ and, by injectivity, $x = x_s + x_n$.

If $h \in H$, then since $[x, h] = 0$ by abelianness and using the Jacobi idendity, we find that $\mathrm{ad}(x) \circ \mathrm{ad}(h) = \mathrm{ad}(h) \circ \mathrm{ad}(z)$. By 3.9(ii), we know that $\mathrm{ad}(h)$ commutes with $\mathrm{ad}(x_s)$ and $\mathrm{ad}(x_n)$.

In particular, h commutes with x_n by injectivity of ad, so $x_n \in Z_L(H) = H$. Since $\mathrm{ad}(x_n)$ is nilpotent, since $\mathrm{ad}(h)$ and $\mathrm{ad}(x_n)$ commute, $\mathrm{ad}(h) \circ \mathrm{ad}(x_n)$ is also nilpotent. In particular, $\mathrm{tr}(\mathrm{ad}(h) \circ \mathrm{ad}(x_n)) = 0$, thus $\langle h, n \rangle_{\mathrm{ad}} = 0$ for all $h \in H$.

We have $x_n \in H$ and by (d) the restriction of $\langle\ ,\ \rangle_{\mathrm{ad}}$ to H is non-degenerate, so $x_n = 0$. Hence $x = x_s$ is semisimple. □

Corollary. Every regular element y of a semisimple Lie algebra is semisimple.

Proof. If y is regular, then $L_{0,y}$ is a Cartan subalgebra, which implies that $y \in L_{0,y}$ is semisimple by the previous theorem. □

Remark. Suppose L is a semisimple complex Lie algebra. Then by (4.10) a CSA H of L is abelian and all elements are semisimple. Then L breaks up as a direct sum of common eigenspaces of the elements of $\mathrm{ad}(H)$.

An easy induction on the dimension of H establishes this: take $h_1, \ldots, h_r$ to be a basis of H. By induction L splits as a direct sum of common eigenspaces for $\mathrm{ad}(h_1), \ldots, \mathrm{ad}(h_{r-1})$. These break up as direct sums of eigenspaces of $\mathrm{ad}(h_r)$.

On each of these common eigenspaces, $\mathrm{ad}(h)(x) = \alpha(h)x$ for some linear form $\alpha\colon H \to \mathcal{C}$. Define

$$L_\alpha \coloneqq \{x \in L \mid \mathrm{ad}(h)(x) = \alpha(h)x\}.$$

Notice that $L_0 = H$ since $L_0 = Z_L(H) = H$ using 4.10. Thus we have the following definition.

Definition 4.11. If L is a semisimple complex finite-dimensional Lie algebra, then the decomposition

$$L = L_0 \oplus \left(\bigoplus_{\alpha \neq 0} L_\alpha \right)$$

is called the weight space or Cartan decomposition with regard to the Cartan subalgebra H. Notice that $L_0 = H$. The α such that $L_\alpha \neq 0$ are called the weights. The space L_α is called the weight space if α is a weight. The non-zero weights are called the roots with respect to H.

We denote the set of roots of L by Φ. We define $m_\alpha \coloneqq \dim L_\alpha$.

Remark. In what follows we are relying on this decomposition. Over fields which are not algebraically closed and of characteristic 0, L might not split in this way. For example, there exist real semisimple Lie algebras which are not split semisimple (i.e., semisimple and has a Cartan decomposition).

Lemma 4.13. (a) Let $x, y \in H$. Then we have the formula

$$\langle x, y \rangle_{\mathrm{ad}} = \sum_{\alpha \in \Phi} m_\alpha \alpha(x)\alpha(y).$$

(b) If α, β are weights, then $[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta}$, with $L_{\alpha+\beta} = 0$ if $\alpha + \beta$ is not a weight. If $\alpha + \beta \neq 0$, then $\langle L_\alpha, L_\beta \rangle_{\mathrm{ad}} = 0$.
(c) If $\alpha \in \Phi$, then $-\alpha \in \Phi$.
(d) The restriction of $\langle\ ,\ \rangle_{\mathrm{ad}}$ to H is non-degenerate.
(e) If α is a weight, then $L_\alpha \cap L_{-\alpha}^{\perp} = 0$.
(f) If $0 \neq h \in H$, then $\alpha(h) \neq 0$ for some $\alpha \in \Phi$, hence Φ spans the dual space H^*.

Proof. For (a) choose a basis for each weight space and take the union to obtain a basis of L. Then $\mathrm{ad}(x)$ and $\mathrm{ad}(y)$ are represented by diagonal matrices, and $\mathrm{tr}(\mathrm{ad}(x) \circ \mathrm{ad}(y))$ is precisely the right hand side of the formula.

The first part of (b) is an immediate consequence of the fact that $\mathrm{ad}(h)$ is a derivation. By an argument similar to that used in 4.10(d) we have that $\mathrm{tr}(\mathrm{ad}(x) \circ \mathrm{ad}(y)) = 0$ if $x \in L_\alpha$ and $y \in L_\beta$ and $\alpha + \beta \neq 0$ (TODO). So $\langle L_\alpha, L_\beta \rangle_{\mathrm{ad}} = 0$.

For (c), take $\alpha \in \Phi$ and suppose $-\alpha \notin \Phi$. Then using (b) we have that for all weights β, $\langle L_\alpha, L_\beta \rangle_{\mathrm{ad}} = 0$.

Hence, $\langle L_\alpha, L \rangle_{\mathrm{ad}} = 0$. But by Cartan-Killing $\langle\ ,\ \rangle_{\mathrm{ad}}$ is non-degenerate on L, so $L_\alpha = 0$, which is a contradiction by definition of a root.

Part (d) is the same as 4.10(d).

Take $x \in L_\alpha \cap L_{-\alpha}^{\perp}$. Then $\langle x, L_\beta \rangle_{\mathrm{ad}} = 0$ for all weights β, since if $\beta \neq -\alpha$, this is true by (b), and if $\beta = -\alpha$, then it is true by choice of x. Hence, $\langle x, L \rangle_{\mathrm{ad}} = 0$, so $x = 0$ by non-degeneracy.

Suppose $h \in H$ is such that $\alpha(h) = 0$ for all $\alpha \in \Phi$. Let $x \in H$. Then $\langle h, x \rangle_{\mathrm{ad}} = \sum_{\alpha \in \Phi} m_\alpha \alpha(h) \alpha(x) = 0$. By (d), we have $h = 0$.

Thus, if $h \neq 0$, there is some $\alpha \in \Phi$ such that $\alpha(h) \neq 0$.

The spanning statement can be proved by inductively finding a basis:: if $n = \dim H$, given $k < n$ linearly independent elements in Φ, by the dimension formula the intersection of the kernels of these linear forms has dimension at least $n - k > 0$, so we find some $h \in H$ on which all elements vanish. By what we just proved, we find a new element of Φ that does not vanish on h and therefore is linearly independent from the previous elements. □

Definition 4.14. The α-string through β for $\alpha, \beta \in \Phi$ is the longest arithmetic progression

$$\beta - q\alpha, \dots, \beta, \dots, \beta + p\alpha,$$

such that all terms are weights.

Lemma 4.15. Let $\alpha, \beta \in \Phi$ and p, q as above. Then

(a) We have

$$\beta(x) = -\frac{\sum_{r=-q}^{p} r m_{\beta+r\alpha}}{\sum_{r=-q}^{p} m_{\beta+r\alpha}} \alpha(x)$$

for all $x \in [L_\alpha, L_{-\alpha}]$.
(b) If $0 \neq x \in [L_\alpha, L_{-\alpha}]$, then $\alpha(x) \neq 0$.
(c) We have $[L_\alpha, L_{-\alpha}] \neq 0$.

.

Proof. For (a), define $M = \sum_{r=-q}^{p} L_{\beta+r\alpha}$. Observe that $[L_{\pm\alpha}, M] \subseteq M$ by maximality of p and q. Let U be the Lie subalgebra generated by L_α and $L_{-\alpha}$. Then $\mathrm{ad}(U)(M) \subseteq M$.

Now take $x \in [L_\alpha, L_{-\alpha}]$. We have that $x \in M^{(1)}$, so $\mathrm{ad}(x)|_M \colon M \to M$ (exists by the above) has zero trace, since it is an element of the derived subalgebra of ad(M). But then

$$0 = \operatorname{tr}\operatorname{ad}(x)|_M = \sum_{r=-q}^{p} m_{\beta+r\alpha}(\beta + r\alpha)(x).$$

Rearranging gives part (a). Note that $\sum_{-q}^{p} m_{\beta+r\alpha}$ is nonzero since multiplicities are positive and β is a root, hence its multiplicity is nonzero.

For part (b), let $0 \neq x \in [L_\alpha, L_{-\alpha}]$ and suppose that $\alpha(x) = 0$. Then we deduce from (a) that $\beta(x) = 0$ for all roots β. This contradicts 4.13(f). Hence $\alpha(x) \neq 0$.

Finally, for (c) let $v \in L_{-\alpha}$. By definition, we have $[h, v] = -\alpha(h)v$ for $h \in H$.

Choose $u \in L_\alpha$ and $v \in L_\alpha$ such that $\langle u, v\rangle_{\text{ad}} \neq 0$. This is possible by 4.13(e). Furthermore, choose $h \in H$ such that $\alpha(h) \neq 0$. Define $x \coloneqq [u, v] \in [L_\alpha, L_{-\alpha}]$. Then $\langle x, h\rangle_{\text{ad}} = \langle u, [v, h]\rangle_{\text{ad}} = \alpha(h)\langle u, v\rangle_{\text{ad}} \neq 0$.

In particular, $x \neq 0$ as required. □

Lemma 4.16. (a) For all $\alpha \in \Phi$, we have $m_\alpha = 1$.
If $n\alpha \in \Phi$ for $n \in \mathbb{Z}$, then $n = \pm 1$.
(b) For $x \in [L_\alpha, L_{-\alpha}]$, we have $\beta(x) = \frac{q-p}{2}\alpha(x)$.

Proof. For (a), take u, v, x as in the previous proof, and let A be the Lie subalgebra generated by u and v and N the vector space span of v, H and $\sum_{r>0} L_{r\alpha}$.

We can calculate $[u, N] \subseteq H \oplus \sum L_{r\alpha} \subseteq N$, noting that $[u, v] \in [L_\alpha, L_{-\alpha}] \subseteq L_0 = H$.

Similarly, $[v, N] \subseteq [v, H] + \sum_{r>0}[v, L_{r\alpha}] \subseteq N$, again using the addition formula for the second term. TODO: why is $[v, H] \subseteq N$?

So $[A, N] \subseteq N$. Then $x = [u, v] \in A^{(1)}$. Consider $\operatorname{ad}(x)|_N \colon N \to N$. We have $0 = \operatorname{tr}\operatorname{ad}(x)|_N$ as x is in the derived subalgebra (TODO: but it's the derived subalgebra of A and not N. Why does that not matter?). Hence

$$0 = -\alpha(x) + \sum m_{r\alpha} r\alpha(x) = \left(-1 + \sum r m_\alpha\right)\alpha(x).$$

But $\alpha(x) \neq 0$ by part (b) of the previous lemma. Hence $\sum r m_\alpha = 1$ for all $\alpha \in \Phi$. Thus for $\alpha \in \Phi$ we have $m_\alpha = 1$ and if $n\alpha$ is a root, then $n = \pm 1$ (the negative case comes from considering $-\alpha$ in the argument above.

Part (b) is obtained by combining (a) with 4.15(a). □

Remark. We have $A \cong \mathfrak{sl}_2$ and we saw that we had a representation $\operatorname{ad}|_N \colon A \to \operatorname{End} N$. Some authors use the representation theory of $\mathfrak{sl}_2$ to establish the lemma.

Lemma 4.17. If $\alpha \in \Phi$ and $c\alpha \in \Phi$ with $c \in \mathbb{C}$, then $c = \pm 1$.

Proof. Set $\beta = c\alpha$. Consider the α-string through β

$$\beta - q\alpha, \ldots, \beta, \ldots, \beta + p\alpha.$$

As before, choose $x \in [L_\alpha, L_{-\alpha}]$ such that $\alpha(x) \neq 0$. By the previous lemma, $\beta(x) = \frac{q-p}{2}\alpha(x)$. Hence, $c = \frac{q-p}{2}$. But if $q - p$ is even, then we're done by the previous lemma.

If $q-p$ is odd, then $r = (p-q+1)/2 \in \mathbb{Z}$ and satisfies $-q \leq r \leq p$ and therefore $\beta + r\alpha$ is a root and we have $\beta + r\alpha = (q - p + p - q + 1)/2\alpha = 1/2\alpha$. But then Φ contains $1/2\alpha$ as well as $2(1/2\alpha)$, which is not possible by 4.16. □

Lemma 4.18. (i) For $\alpha \in \Phi$ we can choose $h_\alpha \in H$, $e_\alpha \in L_\alpha$, $e_{-\alpha} \in L_{-\alpha}$ such that
(a) $\forall x \in H \colon \langle h_\alpha, x\rangle_{\text{ad}} = \alpha(x)$,
(b) $h_{\alpha\pm\beta} = h_\alpha \pm h_\beta$, $h_{-\alpha} = -h_\alpha$ and the h_α span H,
(c) $h_\alpha = [e_\alpha, e_{-\alpha}]$, $\langle e_\alpha, e_{-\alpha}\rangle_{\text{ad}} = 1$.

(ii) If $\dim L = n$ and $\dim H = r$, then the number of roots is $2s = n - r$ and $r \leq s$.

PROOF. For (i), define $h^* \in H^*$ via $h^*(x) := \langle h, x\rangle_{\mathrm{ad}}$. There is a linear map $h \mapsto h^*$. This map is injective by non-degeneracy of the restriction, hence an isomorphism by finite-dimensionality. We define h_α to be the preimage of α. Property (a) is then satisfied by contruction, and (b) is satisfied by linearity of $h \mapsto h^*$ and bececause the H^* are spanning by 4.13(f). By 4.13e we find $e_{\pm\alpha} \in L_{\pm\alpha}$ such that $\langle e_\alpha, e_{-\alpha}\rangle \neq 0$. We can scale them in a way such that $\langle e_\alpha, e_{-\alpha}\rangle = 1$. For $x \in H$ we have

$$\langle [e_\alpha, e_{-\alpha}], x\rangle_{\mathrm{ad}} = \langle e_\alpha, [e_{-\alpha}, x]\rangle_{\mathrm{ad}} = \alpha(x)\langle e_\alpha, e_{-\alpha}\rangle_{\mathrm{ad}} = \alpha(x) = \langle h_\alpha, x\rangle_{\mathrm{ad}},$$

using that the Killing form is invariant and the fact that $e_{-\alpha}$ is an eigenvector for $\mathrm{ad}(x)$ with eigenvalue $-\alpha(x)$.

Again by nondegeneracy, we have $h_\alpha = [e_\alpha, e_{-\alpha}]$ as required.

For (ii), each weight space which is not H has dimension 1 by 4.16(a). Hence, the number of roots is $2s = n - r$ (since roots come in pairs α and $-\alpha$). Since the h_α span S, we find $r \leq s$. □

DEFINITION 4.19. For $\alpha, \beta \in H^*$ define

$$(\alpha, \beta) := \langle h_\alpha, h_\beta\rangle_{\mathrm{ad}}$$

where h_α and h_β are the unique elements of H satisfying

$$\langle h_\alpha, x\rangle_{\mathrm{ad}} = \alpha(x), \quad \langle h_\beta, x\rangle_{\mathrm{ad}} = \beta(x).$$

LEMMA 4.20. (a) We have

$$\frac{2\langle h_\beta, h_\alpha\rangle_{\mathrm{ad}}}{\langle h_\alpha, h_\alpha\rangle_{\mathrm{ad}}} \in \mathbb{Z},$$

(b) Furthermore,

$$4\sum_{\beta\in\Phi} \frac{\langle h_\beta, h_\alpha\rangle_{\mathrm{ad}}^2}{\langle h_\alpha, h_\alpha\rangle_{\mathrm{ad}}^2} = \frac{4}{\langle h_\alpha, h_\alpha\rangle_{\mathrm{ad}}} \in \mathbb{Z}.$$

(c) We have $\langle h_\alpha, h_\beta\rangle_{\mathrm{ad}} \in \mathbb{Q}$ for all $\alpha, \beta \in \Phi$.

(d) For all $\alpha, \beta \in \Phi$ we have

$$\beta - 2\frac{\langle h_\beta, h_\alpha\rangle_{\mathrm{ad}}}{\langle h_\alpha, h_\alpha\rangle_{\mathrm{ad}}}\alpha \in \Phi.$$

These results can be reformulated using $(\ , \ )$.

PROOF. We have $\langle h_\alpha, h_\alpha\rangle_{\mathrm{ad}} = \alpha(h_\alpha) \neq 0$ by 4.15(b).

$$2\frac{\langle h_\beta, h_\alpha\rangle_{\mathrm{ad}}}{\langle h_\alpha, h_\alpha\rangle_{\mathrm{ad}}} = 2\frac{\beta(h_\alpha)}{\alpha(h_\alpha)} = \frac{2(q-p)}{2} \in \mathbb{Z},$$

where the α-string through β is

$$\beta - q\alpha, \ldots, \beta, \ldots, \beta + p\alpha.$$

This shows (a).

For (b), let $x, y \in H$, we have

$$\langle x, y\rangle_{\mathrm{ad}} = \sum_{\beta\in\Phi} \beta(x)\beta(y)$$

by 4.13(a) and 4.16(a). Hence,

$$\langle h_\alpha, h_\alpha\rangle_{\mathrm{ad}} = \sum_{\beta\in\Phi} \beta(h_\alpha)^2 = \sum_{\beta\in\Phi} \langle h_\beta, h_\alpha\rangle_{\mathrm{ad}}^2.$$

Pulling out the denominator from the left hand side of the claim and substituting yields (b).

Part (c) follows immediately from (a) and (b).

For (d), notice that

$$\beta - 2\frac{\langle h_\beta, h_\alpha\rangle}{\langle h_\alpha, h_\alpha\rangle}\alpha = \beta + (p-q)\alpha$$

is in the α-string through β, so we are done.

□

Definition. Define $\tilde{H}$ to be the $\mathbb{Q}$-span of $\{h_\alpha\}_{\alpha\in\Phi} \subseteq H$.

Since the h_α span H as a $\mathbb{C}$-vector space there is a subset $\{h_1, \dots, h_r\}$ forming a $\mathbb{C}$-basis.

Lemma 4.21. The Killing form restricted to $\tilde{H}$ is an inner product and $h_1, \dots, h_r$ is a $\mathbb{Q}$-basis of $\tilde{H}$.

Proof. We know that $\langle\ ,\ \rangle_{\mathrm{ad}}$ is symmetric and bilinear and rationally valued on $\tilde{H}$ by 4.20(c).

Let $x \in \tilde{H}$. Then

$$\langle x, x\rangle_{\mathrm{ad}} = \sum_{\alpha\in\Phi} \alpha(x)^2 = \sum_{\alpha\in\Phi} \langle h_\alpha, x\rangle^2$$

by 4.13(a). Each $\langle h_\alpha, x\rangle$ is rational and so $\langle x, x\rangle_{\mathrm{ad}} \geq 0$ with equality only if $\langle h_\alpha, x\rangle = \alpha(x) = 0$ for every $\alpha \in \Phi$. By 4.13(f), this can only happen if $\alpha = 0$.

It remains to show that each h_α is a rational linear combination of $h_1, \dots, h_r$. But if

$$h_\alpha = \sum \lambda_i h_i$$

with $\lambda_i \in \mathbb{Q}$, then

$$\langle h_\alpha, h_j\rangle_{\mathrm{ad}} = \sum \lambda_i \langle h_i, h_j\rangle \in \mathbb{Q}$$

by 4.20(c). Hence the matrix with entries $\langle h_i, h_j\rangle$ is nonsingular by non-degeneracy and has entries in $\mathbb{Q}$. Hence it is invertible over $\mathbb{Q}$, so each $\lambda_i \in \mathbb{Q}$ as required. □

Remark. Now we can translate these results to make similar statements concerning the $\mathbb{Q}$-span of Φ using the symmetric bilinar form $(\ ,\ )$ on H^*.

Let $\tilde{H}^*$ be the rational dual of $\tilde{H}$. Then $\tilde{H}^*$ is the $\mathbb{Q}$-span of Φ by 4.20(c).

The bilinear form $(\ ,\ )$ restricts to $\tilde{H}^*$ and defines an inner product on $\tilde{H}^*$, and subset Φ' of Φ that is a $\mathbb{C}$-basis of H^* and a $\mathbb{Q}$-basis of $\tilde{H}^*$.

CHAPTER 5

Root systems

Definition 5.1. A subset Φ of a real Euclidean vector space E is called a finite root system if

(a) Φ is finite, spans E and does not contain 0,
(b) for each $\alpha \in \Phi$ there is a reflection s_α preserving the inner product such that $s_\alpha(\alpha) = -\alpha$, the set of fixed points of s_α is a hyperplane of E and s_α leaves Φ invariant,
(c) for each $\alpha, \beta \in \Phi$, $s_\alpha(\beta) - \beta$ is an integral multiple of α,
(d) for $\alpha, \beta \in \Phi$, $2\frac{(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}$, and
(e) $s_\alpha(\beta) = \beta - 2\frac{(\beta,\alpha)}{(\alpha,\alpha)}\alpha$ for all $\beta \in E$.

Remark. From 4.21 and the following remark we could take $E = \mathbb{R}$-span of the roots Φ of a semisimple complex Lie algebra.

We have an inner product on E and 4.20 as converted into the language of (α, β) says that Φ forms a finite root system.

Example. The Lie algebra $\mathfrak{sl}_2$ has the Cartan subalgebra

$$H = \{\begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix} \mid \lambda \in \mathbb{C}\}.$$

We have

$$\begin{aligned} \mathfrak{sl}_2 &= L_0 \oplus L_\alpha \oplus L_{-\alpha}, \\ L_\alpha &= \{\begin{pmatrix} 0 & \lambda \\ 0 & 0 \end{pmatrix} \mid \lambda \in \mathbb{C}\}, \\ L_{-\alpha} &= \{\begin{pmatrix} 0 & 0 \\ \lambda & 0 \end{pmatrix} \mid \lambda \in \mathbb{C}\}. \end{aligned}$$

As usual, denote

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

We have that $\mathrm{ad}(h)$ has eigenvalues $-2, 0, 2$. Thus $\alpha \in H^*$ is the linear form $\mathrm{diag}(\lambda, -\lambda) \mapsto 2\lambda$. Hence $\alpha(h) = 2$.

Definition 5.2. The rank of a root system is $\dim_{\mathbb{R}} E$.

If the root system is induced by a semisimple complex Lie algebra L, the rank of the root system coincides with the rank of L.

Definition 5.3. We say that a root system Φ is reduced if for each $\alpha \in \Phi$, α and $-\alpha$ are the only multiples of α inside Φ.

Remark. (1) The root system arising from a complex semisimple Lie algebra is reduced by 4.17.
(2) Non-reduced root systems are still interesting: they arise over fields which are not algebraically closed.

DEFINITION 5.4. The Weyl group $W(\Phi)$ of a root system Φ is the group generated by the reflections s_α ($\alpha \in \Phi$). It is a subgroup of the orthogonal group of E.

Note that since Φ is finite and spans E and each s_α leaves Φ invariant, $W(\Phi)$ must be finite. Hence, $W(\Phi)$ is a finite reflection group.

DEFINITION 5.5. An isomorphism of root systems $(E, \Phi) \to (E', \Phi')$ is a linear isomorphism $\phi \colon E \to E'$ such that $\phi(\Phi) = \Phi'$.

Note that ϕ is not required to be an isometry.

REMARK. Up to isomorphism, the root system from $\mathfrak{sl}_2$ is the only reduced rank 1 root system.

DEFINITION 5.6. (a) The direct sum of two root systems (E, Φ), (E', Φ') is $(E \oplus E', \Phi \cup \Phi')$.
(b) A root system is called irreducible if it cannot be written as a direct sum.

DEFINITION 5.7. If $\alpha \in \Phi$, define the co-root (or inverse root) by $\alpha^\vee := \frac{2}{(\alpha,\alpha)}\alpha$. Thus $(\alpha, \alpha^\vee) = 2$.

EXERCISE. If (E, Φ) is a root system, then $(E, \Phi^\vee)$ is a root system. This is called the root system dual to Φ.

DEFINITION 5.8. For $\alpha, \beta \in \Phi$, write $n(\beta, \alpha) = 2\frac{(\beta,\alpha)}{(\alpha,\alpha)}$ and recall that $n(\beta, \alpha) \in \mathbb{Z}$ by definition of a root system.

REMARK. Observe that $n(\ ,\ )$ is in general not symmetric. Let $|\alpha| = (\alpha, \alpha)^{1/2}$. Then $(\alpha, \beta) = |\alpha||\beta| \cos \phi$, where ϕ is the angle between α and β.

We have
$$n(\beta, \alpha) = 2\frac{|\beta|}{|\alpha|} \cos \phi.$$

LEMMA 5.9. We have $n(\beta, \alpha)n(\alpha, \beta) = 4 \cos^2 \phi$.

PROOF. This is obvious from what we have just seen.

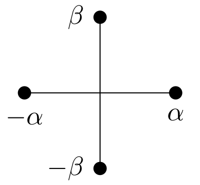
But observe that $n(\beta, \alpha)$ and $n(\alpha, \beta)$ are integers, so $4 \cos^2 \phi$ can only have value 0, 1, 2, 3 or 4 and 4 is only possible when α and β are collinear. Otherwise, there are 7 possibilities.

$n(\alpha,\beta)$	$n(\beta,\alpha)$	ϕ	$\lvert\beta\rvert$
0	0	$\pi/2$	
1	1	$\pi/3$	$\lvert\alpha\rvert$
-1	-1	$2\pi/3$	$\lvert\alpha\rvert$
1	2	$\pi/4$	$\sqrt{2}\lvert\alpha\rvert$
-1	-2	$3\pi/4$	$\sqrt{2}\lvert\alpha\rvert$
1	3	$\pi/6$	$\sqrt{3}\lvert\alpha\rvert$
-1	-3	$5\pi/6$	$\sqrt{3}\lvert\alpha\rvert$

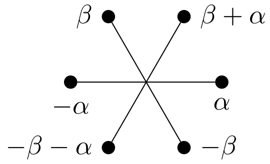
TABLE 1. Exhaustive list of possible combinations of $n(\alpha, \beta)$ and $n(\beta, \alpha)$.

$\square$

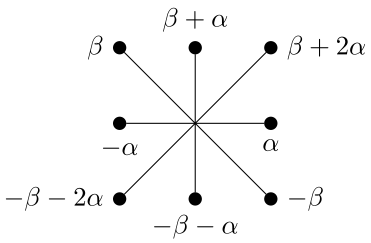
EXAMPLE. Table 2 shows reduced rank 2 root systems. These are the only reduced rank 2 root systems. Of these, A_1, B_2 and G_2 are irreducible and $A_1 \times A_1$ is reducible.

<table>
<tr><th>Picture</th><th>Description</th></tr>
<tr><td>
</td><td>

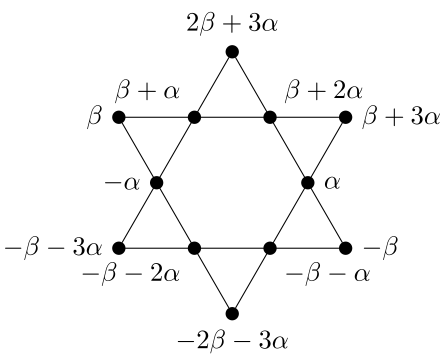
- type $A_1 \times A_1$
- arising from $\mathfrak{sl}_2 \times \mathfrak{sl}_2$
- not irreducible
- Weyl group $C_2 \times C_2$

</td></tr>
<tr><td>
</td><td>

- type A_2
- arising from $\mathfrak{sl}_3$ (TODO: why? CSA are diagonal matrices)
- $\alpha^\vee = \alpha$, $\beta^\vee = \beta$
- $(\alpha, \beta) = -1$
- $\dim \mathfrak{sl}_3 = 6 + 2$, where 6 is the number of roots and 2 is the rank (this comes from the Cartan decomposition!)
- Weyl group $D_6 \cong S_3$

</td></tr>
<tr><td>
</td><td>

- type B_2
- $(\alpha, \alpha) = 1$, $(\beta, \beta) = 2$
- arises from $\mathfrak{sp}_4$ and $\mathfrak{so}_5$
- $\dim L = 10 = 8 + 2$

</td></tr>
<tr><td>
</td><td>

- type G_2
- arising from the Lie algebra of derivations of the octionions
- Weyl group is dihedral D_{12} of order 12

</td></tr>
</table>

Table 2. Reduced rank 2 root systems

Definition 5.10. We call a root system simply laced if all roots have the same length.

The only irreducible simply laced root systems of rank 2 is A_2.

Definition 5.11. A subset Δ of a root system Φ is called a base of Φ if

(i) Δ is a basis of the Euclidean space E,
(ii) each $\gamma \in \Phi$ can be written as a linear combination $\gamma = \sum_{\alpha \in \Delta} k_\alpha \alpha$, where the k_α are either all positive integers or negative integers.

Elements of Δ are called the simple roots and the γ with all $k_\alpha \geq 0$ are called positive roots with regard to Δ, and the other roots are called the negative roots.

We will show that every root system has a base in due course.

Example. In our rank 2 examples, the set $\Delta = \{\alpha, \beta\}$ always forms a base.

Definition 5.12. The Cartan matrix of the root system Φ with respect to the base Δ is the matrix with entries $n(\alpha, \beta)$ with $\alpha, \beta \in \Delta$.

Example. The Cartan matrix for G_2 is $\begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}$. Note that $n(\alpha, \alpha) = 2$ for all $\alpha \in \Phi$ (by definition!), hence we always have 2s on the diagonal of the Cartan matrix.

We will prove that all other entries are nonpositive.

Definition. A Coxeter graph is a finite graph such that each pair of distinct vertices are joined by 0, 1, 2 or 3 edges. Given a root system Φ and a base Δ, the Coxeter graph (or unlabelled Dynkin diagram) of Φ with respect to Δ has as vertices the element of Δ. A vertex α is joined to $\beta \neq \alpha$ by $n(\alpha, \beta)n(\beta, \alpha)$ edges.

Example 5.13. The following table shows the Coxeter Graphs for reduced rank 2 root systems.

A_1	$\bullet$
$A_1 \times A_1$	$\bullet \qquad \bullet$
A_2	$\bullet$ —— $\bullet$
B_2	$\bullet =\!= \bullet$
G_2	$\bullet \equiv \bullet$

Theorem 5.14. Every connected non-empty Coxeter graph associated to an irreducible reduced root system is isomorphic to one of the following

- A_r $(r \geq 1)$ is a path with r vertices and $r - 1$ single edges.
- B_r $(r \geq 2)$ is a path with r vertices and $r - 1$ single edges, but the final edge is a double edge.
- D_r $(r \geq 4)$ is a path with $r - 1$ vertices plus another vertex connected to the second-to-last vertex of the path.
- E_6 is a path with 5 vertices with another vertex connected to the middle vertex.
- E_7 is a path with 6 vertices with another vertex connected to one of the middle vertices.
- E_8 is a path with 7 vertices with another vertex connected to the middle vertices.
- F_4 is a path with 4 vertices and the middle edge is a double edge.
- G_2 consists of two vertices connected by a triple edge.

The ones coming from simply laced irreducible root systems are exactly the ones without double edges, i.e., they are A_r, D_r, E_6, E_7, E_8.

Remark. The Coxeter graphs give insufficient information to recover the Cartan matrix of the root system—we want something additional to indicate the relative lengths of the roots in the multiple edge cases.

We orient the double edges of the Coxeter graphs by having them point towards the shorter root. In this way, the family B_r splits into B_r, where the double edge is directed toward the final vertex, and C_r, wehere the double edge is directed toward the second-to-last vertex.

For F_4 and G_2, we also get directed versions, but they are symmetric.

We do not need arrows in the simply laced cases.

REMARK. For $\gamma \in E$ define $\Phi^+(\gamma) = \{\alpha \in \Phi \colon (\gamma, \alpha) > 0\}$.

$E \setminus \bigcup P_\alpha$ is nonempty, where P_α is the hyperplane of reflections of s_α, where $\alpha \in \Phi$.

DEFINITION 5.15. (a) We call γ regular if $\gamma \in E \setminus \bigcup P_\alpha$.
If γ is regular, we obtain a decomposition $\Phi = \Phi^+(\gamma) \cup (-\Phi^+(\gamma))$. The only thing that can go wrong is that $(\gamma, \alpha) = 0$ for some $\alpha \in \Phi$, but then $s_\alpha(\gamma) = \gamma$, so $\gamma \in P_\alpha$, a contradiction.
(b) We say that $\alpha \in \Phi^+(\gamma)$ is indecomposable if it is not possible to express it as a sum $\alpha_1 + \alpha_2$ with $\alpha_1, \alpha_2 \in \Phi^+(\gamma)$.

LEMMA 5.16. Let $\gamma \in E$ be regular. Then the set $\Delta(\gamma)$ of all indecomposable roots in $\Phi^+(\gamma)$ is a base of Φ. Every base has this form.

PROOF. We will first show that every $\alpha \in \Phi^+(\gamma)$ is a non-negative integral combination of elements of $\Delta(\gamma)$.

Otherwise, choose α with $(\gamma, \alpha) > 0$ which does not satisfy the claim with (γ, α) minimal. Then α is decomposable, say $\alpha = \alpha_1 + \alpha_2$ with $\alpha_i \in \Phi^+(\gamma)$. But $(\gamma, \alpha) = (\gamma, \alpha_1) + (\gamma, \alpha_2)$. By minimality, α_1 and α_2 are good, hence α is also good, a contradiction.

Hence, $\Delta(\gamma)$ spans E and satisfies (ii) of the definition of a base. To show linear independence, it suffices to show that $(\alpha, \beta) \leq 0$ for α, β distinct elements of $\Delta(\gamma)$ (TODO: why?).

Indeed, otherwise we would find α, β such that $(\alpha, \beta) > 0$. By definition of n, this implies that $n(\alpha, \beta) > 0$ and $n(\beta, \alpha) > 0$. Consulting Table 1, we conclude $n(\alpha, \beta) = 1$ or $n(\beta, \alpha) = 1$. If $n(\alpha, \beta) = 1$, then $\alpha - \beta = -(\beta - n(\beta, \alpha)\alpha) = -s_\alpha(\beta)$ is a root. Similarly, if $n(\beta, \alpha) = 1$ then $\alpha - \beta$ is a root. So $\alpha - \beta$ is a root, and such $\alpha - \beta \in \Phi^+(\gamma)$ or $\beta - \alpha \in \Phi^+(\gamma)$.

In the first case $\alpha = (\alpha - \beta) + \beta$, so α is decomposable. Similarly, in the second case, α is also decomposable. This is a contradiction, hence $(\alpha, \beta) \leq 0$ as claimed.

Finally, suppose that Δ is a base. Choose γ such that $(\gamma, \alpha) > 0$ for all $\alpha \in \Delta$ (TODO: I guess that is some geometric voodoo?). Then γ is certainly regular and we will show that $\Delta = \Delta(\gamma)$.

Certainly, we have $\Phi^+ \subseteq \Phi^+(\gamma)$ (using linearity of the scalar product). Hence $-\Phi^+ \subseteq -\Phi^+(\gamma)$. But since Φ splits into $\Phi^+ \cup -\Phi^-$ and $\Phi^+(\gamma) \cup -\Phi^+(\gamma)$, we conclude that the converse inclusion is also true, hence $\Phi^+ = \Phi^+(\gamma)$. But Δ is a base, so each of Φ^+ is a positive integral combination of Δ and so elements of Δ are indecomposable. This implies $\Delta \subseteq \Delta(\gamma)$. But $|\Delta| = \dim E = |\Delta(\gamma)|$, so $\Delta = \Delta(\gamma)$. □

LEMMA 5.17. Let Δ be a base in Φ, where Φ is reduced.

(a) If $\alpha, \beta \in \Delta$, then $\alpha - \beta \notin \Phi$ and $(\alpha, \beta) \leq 0$. Hence, the non-diagonal entries of the Cartan matrix are less than or equal to 0.
(b) If $\alpha \in \Phi^+$ and $\alpha \notin \Delta$, then $\alpha - \beta \in \Phi^+$ for some $\beta \in \Delta$.
(c) Each $\alpha \in \Phi^+$ is of the form $\beta_1 + \cdots + \beta_s$, where each partial sum $\beta_1 + \cdots + \beta_i \in \Phi^+$ and where each β_i is a simple root (not necessarily distinct).
(d) If $\alpha \in \Delta$, then s_α permutes $\Phi^+ \setminus \{\alpha\}$. Set $\rho = \frac{1}{2} \sum_{\beta \in \Phi^+} \beta$, then $s_\alpha(\rho) = \rho - \alpha$.

PROOF. For (a), if $\alpha - \beta \in \Phi$, then part (ii) of the definition of base would be violated. We already saw $(\alpha, \beta) \leq 0$ in the previous lemma.

If $(\alpha, \beta) \leq 0$ for all $\beta \in \Delta$, then $\Delta \cup \{\alpha\}$ would be linearly independent. Hence, we have $(\alpha, \beta) > 0$ for some $\beta \in \Delta$, and using the same argument as in the previous lemma, $\alpha - \beta$ is a root.

For (b), if $\alpha = \sum_{\gamma\in\Delta} k_\gamma \gamma$, then k_γ for at least two $\gamma \in \Delta$, and so for at least one $\gamma \neq \beta$. So $\alpha - \beta \in \Phi^+$.

(c) follows from (b) via an induction on the sum of the coefficients.

For (d), if $\beta = \sum_{\gamma\in\Delta} k_\gamma\gamma \in \Phi^+ \setminus \{\alpha\}$, then there is some $k_\gamma > 0$ with $\gamma \neq \alpha$. But the coefficient of γ in $s_\alpha(\beta) = \beta - 2\frac{(\beta,\alpha)}{(\alpha,\alpha)}\alpha$ is $k_\gamma > 0$. So $s_\alpha(\beta) \in \Phi^+$, so $s_\alpha(\beta) \in \Phi^+ \setminus \alpha$.

For $\rho = \frac{1}{2}\sum_{\beta\in\Phi^+}\beta$ set $\rho' = \rho - \alpha/2$. We have $s_\alpha(\rho') = \rho'$ since s_α permutes. Thus $s_\alpha(\rho) = \rho - \alpha$. □

LEMMA 5.18. Let Δ be a base of a root system Φ.

(a) If $\sigma \in \mathrm{GL}(E)$ is orthogonal and satisfies $\sigma(\Phi) = \Phi$, then $\sigma s_\alpha \sigma^{-1} = s_{\sigma(\alpha)}$.
(b) Let $\alpha_1, \ldots, \alpha_t \in \Delta$ not necessarily distinct.
If $s_t \cdots s_2(\alpha_1)$ is negative, then for some $1 \leq a \leq t$, then $s_t \cdots s_1 = s_t \cdots s_{a+1} s_{a-1} \cdots s_2$, where $s_i := s_{\alpha_i}$.
(c) If $\sigma = s_t \cdots s_1$ is an expression for $\sigma \in W(\Phi)$ in terms of simple reflections $\alpha_1, \ldots, \alpha_t$ with t minimal, then $\sigma(\alpha_1)$ is negative.

PROOF. For (a), let $\alpha \in \Phi$, $\beta \in E$. Then

$$(\sigma s_\alpha \sigma^{-1})\sigma(\beta) = \sigma s_\alpha(\beta) = \sigma(\beta) - n(\beta,\alpha)\sigma(\alpha).$$

Hence, $\sigma s_\alpha \sigma^{-1}$ fixes pointwise $\sigma(P_\alpha)$ and sends $\sigma(\alpha) \mapsto -\sigma(\alpha)$. But then we must have $\sigma^{-1} s_\alpha \sigma = s_{\sigma(\alpha)}$ by orthogonality.

For (b), take a minimal such that $s_a \cdots s_2(\alpha_1)$ is negative. Then $1 < a \leq t$. Then $\beta := s_{a-1} \cdots s_2(\alpha_1)$ is positive by minimality. By the first part of 5.17(d), we have $\beta = \alpha_a$. Define $\sigma := s_{a-1} \cdots s_2$. We have $s_a = s_\beta = s_{\sigma(\alpha_1)} = \sigma s_{\alpha_1}\sigma^{-1} = \sigma s_1 \sigma^{-1}$ using (a). The claim now follows after rearranging (recall that reflections are self-inverse).

For (c), notice that if $\sigma(\alpha_1) = s_t \cdots s_1(\alpha_1) = -s_t \cdots s_2(\alpha_1)$ is positive, then $s_t \cdots s_2(\alpha_1)$ is negative, so by the previous part the expression would not be minimal. □

LEMMA 5.19. Let $W = W(\Phi)$ denote the Weyl group of a reduced root system Φ.

(a) If γ is a regular element of E, then we find $\sigma \in W(\Phi)$ such that $(\sigma(\gamma), \alpha) > 0$ for all $\alpha \in \Delta$. Furthermore, W permutes the bases transitively.
(b) If $\alpha \in \Phi$, then $\sigma(\alpha) \in \Delta$ for some $\sigma \in W$.
(c) $W = \langle s_\alpha \mid \alpha \in \Delta\rangle$.
(d) If $\sigma(\Delta) = \Delta$ for $\sigma \in W$, then $\sigma = \mathrm{id}$, i.e., W permutes the bases regularly.

PROOF. Define $W' = \langle s_\alpha \mid \alpha \in \Delta\rangle \subseteq W$. We will first prove (a) and (b) for W' in place of W.

For (a), define $\rho := \frac{1}{2}\sum_{\alpha\in\Phi^+}\alpha$ and let γ be regular. Choose $\sigma \in W'$ with $(\sigma(\gamma), \rho)$ maximal. Then for $\alpha \in \Delta$ we have $s_\alpha, \sigma \in W'$. Hence

$$(\sigma(\gamma), \rho) \leq (s_\alpha\sigma(\gamma), \rho) = (\sigma(\gamma), s_\alpha(\rho)) = (\sigma(\gamma), \rho) - (\sigma(\gamma), \alpha),$$

using the fact that reflections preserve the inner product and 5.17(d). Hence $(\sigma(\gamma), \alpha) \geq 0$. Equality would imply $(\gamma, \sigma^{-1}(\alpha)) = 0$, hence $\gamma \in P_{\sigma^{-1}(\alpha)}$, which is a contradiction since γ is regular.

Also, $\sigma^{-1}(\Delta)$ is a base with $(\gamma, \alpha') > 0$ for all $\alpha' \in \sigma^{-1}(\Delta)$. By an argument as in the proof of 5.16, we find that $\sigma^{-1}(\Delta) = \Delta(\gamma)$. Since any base is of the form $\Delta(\gamma)$ by 5.16 transitivity follows.

For (b), it will suffice to show that each root α is in a base and then use (a). Choose $\gamma_1 \in P_\alpha \setminus \bigcup_{\beta\neq\pm\alpha} P_\alpha$. Define $\varepsilon := \frac{1}{2}\min\{|(\alpha,\beta)| \mid \beta \neq \pm\alpha\}$. Pick γ_2 in such a way that $|(\gamma_2, \beta)| < \varepsilon$ for each $\beta \neq \pm\alpha$. Define $\gamma := \gamma_1 + \gamma_2$. Then $0 < (\gamma, \alpha) < \varepsilon$ and $|(\gamma, \beta)| > \varepsilon$ for each $\beta \neq \pm\alpha$.

Hence, α is an indecomposable element of $\Phi^+(\gamma)$, and so $\alpha \in \Delta(\gamma)$.

For (c), it will be enough to show that $\alpha \in \Phi \implies s_\alpha \in W'$. By (b), we find some $\sigma \in W'$ with $\sigma(\alpha) \in \Delta$. In particular, $s_{\sigma(\alpha)} \in W'$. But $s_{\sigma(\alpha)} = \sigma^{-1} s_\alpha \sigma$ by 5.18(a). Rearranging gives $s_\alpha = \sigma s_{\sigma(\alpha)} \sigma^{-1} \in W'$ as required.

Suppose we find $\sigma \in W$ such that $\sigma(\Delta) = \Delta$ but $\sigma \neq \mathrm{id}$. Write σ as a product of simple reflections in the shortest possible way. By 5.18(c), this means that there is some $\alpha \in \Delta$ whose image under σ is negative, i.e., not an element of Δ. This is a contradiction. □

THEOREM 5.20. We have

$$W(\Phi) = \{s_\alpha \mid \alpha \in \Delta, s_\alpha^2 = 1, (s_\alpha s_\beta)^{m(\alpha,\beta)} = 1, m(\alpha, \beta) \in \{2, 3, 4, 6\}\},$$

where $m(\alpha, \beta)$ depends on the angle between α and β: $\frac{\pi}{2}$, $\frac{2\pi}{3}$, $\frac{3\pi}{4}$ or $\frac{5\pi}{6}$.

NOT PROVED IN THIS COURSE. □

1. Construction of root systems from Dynkin diagrams/Cartan matrices

REMARK. Our strategy will be the following: let $e_1, \dots, e_n$ be an orthonormal basis in Euclidean n-space. Denote by I the set of integral combinations of elements of the form $\frac{1}{2}e_i$. J will be a subgroup of I and x, y fixed real numbers strictly greater than zero with $\frac{x}{y} \in \{1, 2, 3\}$. Define $\Phi = \{\alpha \in J \mid |\alpha|^2 \in \{x, y\}\}$, $\mathbb{E} = \langle \Phi \rangle$. We need that each reflection s_α preserves lengths and $s_\alpha(\Phi) = \Phi$ and so ensure $n(\beta, \alpha) \in \mathbb{Z}$.

Note that if $J \subseteq \sum \mathbb{Z} e_i$ and $x, y \in \{1, 2\}$, then this is satisfied.

We will first consider A_r for $r \geq 1$. Take $n = r + 1$ and

$$J = (\sum \mathbb{Z} e_i) \cap \langle \sum_{i=1}^{r+1} e_i \rangle^\perp.$$

Define

$$\Phi := \{\alpha \in J \mid |\alpha|^2 = 2\} = \{e_i - e_j \mid i \neq j\}.$$

The elements $\alpha_i = e_i - e_{i+1}$ with $i \leq r$ are linearly independent and if $i < j$ then $e_i - e_j = \sum_{k=i}^{j-1} \alpha_k$. Hence the α_i form a base for Φ. We have $(\alpha_i, \alpha_j) = 0$ unless $j \in \{i, i+1\}$, and $(\alpha_i, \alpha_i) = 2$, and $(\alpha_i, \alpha_{i+1}) = -1$. Hence, the Dynkin diagram of Φ is A_r as required.

Each permutation of $(1, \dots, r + 1)$ induces an automorphism of Φ. Hence $W(\Phi) \cong S_{r+1}$, since s_{α_i} switches $i, i+1$ and the transpositions $(i, i + 1)$ generate S_{r+1}. This is the root system for $\mathfrak{sl}_{r+1}$.

Next, we will consider B_r for $r \geq 2$. Set $n = r$ and define $J = \sum \mathbb{Z} e_i$. Then $\Phi = \{\alpha \in J \mid |\alpha|^2 \in \{1, 2\}\} = \{\pm e_i, \pm e_i \pm e_j \mid i \neq j\}$. Take $\alpha_i := e_i - e_{i+1}$ with $i < r$ and $\alpha_r = e_r$. These are linearly independent and we have $e_i = \sum_{k=i}^{r} \alpha_k$, $e_i + e_j$ is the sum of two such expressions, and $e_i - e_j = \sum_{k=i}^{j-1} \alpha_k$. So $\alpha_1, \dots, \alpha_r$ form a base of Φ. This root system corresponds to the Dynkin diagram B_r.

For the action of $W(\Phi)$, observe that all permutations and sign changes of $e_1, \dots, e_r$ have an effect, hence $W(\Phi)$ is isomorphic to a split extension of C_2^r by S_r, i.e., we have a normal subgroup isomorphic to C_2^r and a subgroup isomorphic to S_r, such that S_r acts on C_2^r by conjugation. This is known as the permutation wreath product. This arises from the Lie algebra $\mathfrak{so}_{2r+1}$.

Next, we will consider C_r for $r \geq 3$. Set $n = r$ and define $J = \sum \mathbb{Z} e_i$. Then $\Phi = \{\alpha \in J \mid |\alpha|^2 \in \{2, 4\}\} = \{\pm 2e_i, \pm e_i \pm e_j \mid i \neq j\}$. This is the dual root system for B_r. We have a base $e_1 - e_2, e_2 - e_3, \dots, e_{r-1} - e_r, 2e_r$. The Weyl group is identical to the one of B_r. This arises from $\mathfrak{sp}_{2r}$.

Type	Number of elements	Weyl group	$\dim L$
A_r	$\frac{1}{2}r(r+1)$	S_{r+1}	$r(r+2)$
B_r, C_r	r^2	$C_2^r \rtimes S_r$	$r(2r+1)$
D_r	$r^2 - r$	index 2 subgroup of above	$r(2r-1)$
E_6	36	$72 \cdot 6!$	78
E_7	63	$8 \cdot 9!$	133
E_8	120	$2^6 \cdot 3 \cdot 10!$	248
F_4	24	1152	52
G_2	6	$D_1 2$	14

TABLE 3. The irreducible root systems

Next, we will consider D_r ($r \geq 4$). Set $n = r$ and $J = \sum \mathbb{Z}e_i$ and

$$\Phi = \{\alpha \in J \mid |\alpha|^2 = 2\} = \{\pm e_i \pm e_j \mid i \neq j\}.$$

Set $\alpha_i = e_i - e_{i+1}$ for $i < r$ and $\alpha_r = e_{r-1} - e_r$. These form a base and the simple reflectionsd cause permutation and an even number of sign changes of ${}_1, \ldots, e_r$. Hence, $W(\Phi)$ is a split extension and $C_2^{r_1}$ by S_r, which is of index 2 inside the wreath product $C_2^r \rtimes S_r$. This arises from $\mathfrak{so}_{2r}$.

For E_8, we set $n = 8$ and take $f \coloneqq \frac{1}{2}\sum_{i=1}^{8} e_i$. Define

$$J \coloneqq \{cf + \sum_{c_i e_i} \mid c, c_i \in \mathbb{Z}, c + \sum c_i \in \mathbb{Z}\}.$$

Then

$$\Phi = \{\alpha \in J \mid |\alpha|^2 = 2\} = \{\pm e_i \pm e_j \mid i \neq j\} \cup \{\frac{1}{2}\sum_{i=1}^{8}(-1)^{k_i} e_i \mid \sum k_i \in 2\mathbb{Z}\}.$$

Set $\alpha_1 = \frac{1}{2}(e_1 + e_8 - \sum_{i=3}^{7} e_i)$, $\alpha_2 = e_1 + e_2$, $\alpha_i = e_{i-1} - e_{i-2}$ for $i \geq 3$.

Next, we consider E_7 and E_6. Take Φ from E_8 and take the intersections

$$\Phi \cap \langle y \rangle^\perp \qquad \Phi \cap \langle h, y \rangle^\perp$$

for suitable h and y. We obtain $\alpha_1, \ldots, \alpha_k$ for $k = 7, 6$ with Dynkin diagrams E_7, E_6.

For F_4, take $n = 4$ and set $h = \frac{1}{2}(e_1 + e_2 + e_3 + e_4)$. Define $J = \sum \mathbb{Z}e_i + \mathbb{Z}h$. Then

$$\Phi = \{\alpha \in J \mid |\alpha|^2 \in \{1, 2\}\} = \{\pm e_i, \pm e_i \pm e_j, \frac{1}{2}(\pm e_1 \pm e_2 \pm e_3 \pm e_4) \mid i \neq j\}$$

and we have a base $e_2 - e_3$, $e_3 - e_4$, e_4, $\frac{1}{2}(e_1 - e_2 - e_3 - e_4)$.

Finally, for G_2, take $n = 3$, $J = \sum \mathbb{Z}e_i \cap \langle e_1 + e_2 + e_3 \rangle^\perp$. Then

$$\Phi = \{\alpha \in J \mid |\alpha|^2 \in \{2, 6\}\} = \{\pm(e_i - e_j), \pm(2e_i - e_j - e_k) \mid i, j, k \text{ distinct}\}$$

and we have a base $\alpha_1 = e_1 - e_2$, $\alpha_2 = -2e_1 + e_2 + e_3$.

REMARK. One may have orthogonal automorphisms of Φ that are not in $W(\Phi)$ (which we showed was generated by the simple reflections).

For example, one can take an automorphism of the Dynkin diagram (that indeed induces an automorphism of the root system), e.g., we can flip A_r. This automorphism is not part of the Weyl group since we showed in 5.19(d) that if an element of $W(\Phi)$ leaves the set of simple roots invariant, then it has to be the identity. Since the vertices of the diagram are exactly the simple roots, the flip satisfies this condition.

REMARK.

THEOREM 5.21. There is a semisimple complex Lie algebra giving rise to each of these irreducible Lie algebras.

NOT PROVED IN THIS COURSE. $\square$

REMARK. We conclude with a few remarks on the proof of the classification of connected Coxeter graphs arising from irreducible reduced root systems.

Given such a Coxeter graph, one can define a symmetric bilinear form on the $\mathbb{R}$-span of the vertices v_i represented with respect to the maxtrix $v_1, \ldots, v_r$ by a matrix (q_{ij}) with $q_{ij} = 2$ for $i = j$ and otherwise $q_{ij} = -\sqrt{t_i j}$, where t_{ij} is the number of edges joining v_i to v_j.

If the graph is coming from the simple roots Δ if a root system Φ in Euclidean space E, then $q_{ij} = 2\frac{(\alpha_i,\alpha_j)}{(\alpha_i,\alpha_i)}$, where $(\ ,\ )$ is the inner product in E.

Note that the matrix is the same as 2 times the matrix representing the inner product $(\ ,\ )$ with respect to the basis $\{\frac{\alpha_i}{|\alpha_i|} \mid \alpha_i \in \Delta\}$, the normalised simple roots.

The matrix is therefore positive definite. Our task is to classify positive definite Coxeter graphs, i.e., Coxeter graphs for which the bilinear form as defined above is positive definite. Some remarks:

(1) Positive semidefinite Coxeter graphs are also of interest for infinite Lie algebras.
(2) Positive definite Coxeter graphs also arise in the representation theory of quivers (directed graphs).

LEMMA 5.22. A connected positive Coxeter graph with r vertices has exactly $r - 1$ pairs of vertices joined by at least one edge.

Ignoring multiplicity of edges, this means that we have a tree.

PROOF. Let e denote the number of pairs of vertices with a least one edge between them. Let $v := \sum_{i=1}^{r} v_i$ (as an element of the $\mathbb{R}$-span). Then

$$0 < q(v,v) = 2r + 2\sum_{i<j} q_{ij},$$

but for $i < j$ we have $q_{ij} \leq 0$, and so

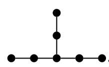
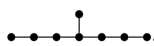
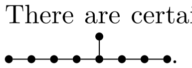
$$r > -\sum_{i<j} q_{ij} = \sum \sqrt{t_{ij}} \geq e.$$

Hence, $e \leq r - 1$, so by connectedness, we have $e = r - 1$.

$\square$

REMARK. The strategy is now as follows. The details are left as an exercise.

(1) The only connected positive definitive Coxeter graphs with 3 vertices are paths of length 2, either with two single edges or a single edge and a double edge. The other cases fail to satisfy the inequality in the previous lemma.
(2) If there is a triple edge, there cannot be any other multiple edges.
(3) Removing some of the vertices and all of the edges attached to them from a positive definite Coxeter graph yields a positive definite Coxeter graph.
(4) Contracting an edge of a positive definite Coxeter graph yields a positive definite Coxeter graph.

 Similarly, if a positive definite Coxeter graph contains the configuration , then collapsing the two edges to yields a positive definite Coxeter graph.
(5) There are certain invalid graphs: , , , and .

(6) Show that every positive definite Coxeter graph is one of A_r, B_r, D_r, E_6, E_7, E_8, F_4, G_2.

CHAPTER 6

Finite dimensional associative algebras

Example. (1) Let D be a finite-dimensional division k-algebra. Then we have an algebra $M_n(D)$ of $n \times n$-matrices over D. An example of a division algebra are the quaternions $\mathbb{H}$.

Let $A \in R := M_n(D)$. The right ideal generated by A is $AR = \{B \in R \mid (\star)\}$, where $(\star)$ means that the columns of B lie in the right span of columns of A (i.e., using right scalar multiplication for the linear combination).

A general right ideal is of the form $\{B \mid (\star)\}$, where $(\star)$ means that the columns of B lie in some fixed right D-subspace of column vectors.

Similarly, the left ideal RA consists of matrices B such that the rows of B lie in the left span of A, and a general left ideal consists of matrices B whose rows lie inside some left D-subspace of row vectors.

Note that the only 2-sided ideals are 0 and $M_n(D)$. Hence $M_n(D)$ is a simple algebra.

(2) Let k be a field and G be a finite group. Then kG is the k-vector space with basis G. A general element of kG looks like $\sum_{g\in G} \lambda_g g$, where $\lambda_g \in k$. We can declare a product via

$$\left(\sum_{g\in G} \lambda_g g\right)\left(\sum_{h\in G} \mu_h h\right) = \sum_{\ell\in G} v_\ell \ell,$$

where $v_k = \sum_{gh=\ell} \lambda_g \mu_h$.

Definition 6.1. The Jacobson radical $J(R)$ of R is the intersection of the maximal right ideals.

This is actually a two-sided ideal (as we will show in a minute).

Definition 6.2. An R-module M is called simple if its only submodules are 0 and M.

Note that I is a maximal right ideal if and only if R/I is a simple right R-module.

Let M be a simple right R-module and $m \in M$. Then $\operatorname{Ann}_R(m) = \{r \in R \mid mr = 0\}$ is a right ideal, but not necessarily a two-sided ideal.

However, $\operatorname{Ann}_R(M) = \bigcap_{m\in M} \operatorname{Ann}(m)$, the annihilator of the module, is a two-sided ideal: if $r \in \operatorname{Ann}_R(M)$ and $x \in R$, then $m(xr) = (mx)r = 0$, since $r \in \operatorname{Ann}_R(mx)$. Hence $xr \in \operatorname{Ann}_R(M)$, so $\operatorname{Ann}_R(M)$ is a left ideal.

If M is simple, then the $\operatorname{Ann}_R(m)$ for $m \neq 0$ are maximal right ideals, because $mR = M$ by simplicity of M, so $R/\operatorname{Ann}_R(m) \cong M$ (via the map $r \mapsto mr$) is simple, so $\operatorname{Ann}_R(m)$ is maximal. So we can see that $J(R) = \bigcap_{M \text{ simple}} \operatorname{Ann}_R(M)$ is a two-sided ideal (all ideals of the form $\operatorname{Ann}_R(M)$ are maximal, and if I is a maximal right ideal, then $I = \operatorname{Ann}_R(R/I)$).

Lemma 6.3 (Nakayama's lemma). The following are equivalent for a right ideal I.

(1) $I \subseteq J(R)$,

(2) if M is a finitely generated R-module and $N \subseteq M$ satisfying $N + MI = M$, then $N = M$,
(3) $G \coloneqq \{1 + x \mid x \in I\}$ is a subgroup of $R^\times$.

Proof. Suppose $I \subseteq J(R)$, M is a finitely generated right R-module and $N \subseteq M$ is a submodule such that $N + MI = M$. If $N \neq M$, then $N \subseteq N'$ for a maximal right R-module N' (this follows from Zorn's lemma, as upper bounds exist: if the union of a chain of proper submodules was not a proper submodules, then all of the finitely many generators of M would be contained in some member of the chain). Let $m \in M$. By assumption we find $n \in N$, $m'_k \in M$ and $i_k \in I$ such that $m = n + \sum_k m'_k i_k$. Since M/N' is simple, $\operatorname{Ann}(m'_k + N')$ is maximal, so $I \subseteq \operatorname{Ann}(m'_k + N')$. In particular, $(m'_k + N')i_k = 0$, so $m'_k i_k \in N'$. Hence $m \in N + N' \subseteq N'$, so $M = N'$, a contradiction. So $N = M$ as required.

Next, assume (2) holds and let $x \in I$. Let $M \coloneqq R$, $N \coloneqq (1+x)R$. Let $r \in R$. Then $r = (1+x)r + 1 \cdot (-xr) \in (1+x)R + RI$. Hence, by (2) we have $(1+x)R = R$, so we find $y' \in R$ such that $(1+x)y' = 1$. Define $y \coloneqq y' - 1$, then $1 + y = y'$, so $(1+x)(1+y) = 1$. In particular, $x + y + xy = 0$, so $y = -x - xy \in I$.

Repeating the argument for y, we find $z \in R$ such that $(1+y)z = 1$. Then $(1+y)(1+x) = (1+y)(1+x)(1+y)z = (1+y)z = 1$, so $1 + x$ is a unit with two-sided inverse $1 + y$. Hence G is a subset of $R^\times$ and closed under taking inverses. Since G is obviously closed under multiplication since I is an ideal, (3) follows.

Finally, assume that G is a subgroup of $R^\times$ and J is a maximal right ideal. Suppose $I \nsubseteq J$. Then we find $i \in I$ such that $i \notin J$. Since J is maximal, this implies $J + iR = R$, i.e., we find $j \in J$, $r \in R$ such that $j + ir = 1$. But then $j = 1 + (-ir)$, but $-ir \in I$, so j is a unit, which is a contradiction. Thus, $I \subseteq J$ for all maximal right ideals J, so $I \subseteq J(R)$. $\square$

Remark. $J(R)$ is the largest two-sided ideal J in R such that $\{1 + x \mid x \in J\}$ is the subgroup of the unit group and so $J(R)$ is indepent on whether it is defined via right or left multiplication.

Recall that we called R semisimple if $J(R) = 0$.

Example. (1) $M_n(D)$, where D is a finite dimensional division algebra, is simple, so it is semisimple.
(2) Let $G \coloneqq \mathbb{Z}/p\mathbb{Z}$. And let $k \coloneqq \mathbb{F}_p$. Then $\mathbb{F}_pG \cong \mathbb{F}_p[X]/(X^p - 1)$. By Frobenius, $X^p - 1 = (X-1)^p$, hence $J(\mathbb{F}_pG) \cong J(\mathbb{F}_p[X]/(X-1)^p) = (X-1)/(X-1)^p \neq 0$ (TODO: verify this calculation). Hence, $\mathbb{F}_pG$ is not semisimple.
Remark: If $p \mid |G|$, then kG is not semisimple if $\operatorname{char} k = p$. The proof is nontrivial.

Lemma 6.4. Let R be a semisimple finite-dimensional k-algebra. Then R is the direct sum of finitely many simple right R-modules.

This can be throught of as a non-commutative Chinese remainder theorem.

Proof. By semisimplicity, the intersection of all maximal right ideals is trivial. Consider $R \supsetneq I_1 \supsetneq I_1 \cap I_2 \supsetneq \dots$, where I_i are maximal right ideals. This must terminate since we are dealing with finite-dimensional k-vector spaces, i.e., there is some n such that $0 = I_1 \cap \cdots \cap I_n$. Choose n minimal. Consider the homomorphism of right R-modules $\theta\colon R \to \bigoplus R/I_i$ defined via $r \mapsto (r + I_1, \dots, r + I_n)$. Note that $\bigcap_{j \neq i} I_j \neq 0$ by minimality and the restriction of the quotient map $R \to R/I_i$ to $\bigcap_{j \neq i} I_j$ is injective (since the kernel is just the intersection of all I_i). Hence, the image is non-zero in R/I_i, but the quotient is simple, so the map is actually an isomorphism $\bigcap_{j \neq i} I_j \cong R/I_i$. In particular, $I_2 \cap \cdots \cap I_n$ gets mapped to $(R/I_1, 0, \dots, 0)$ under θ, and so we see that θ is surjective, so it is an isomorphism as the kernel of θ is again the intersection of the I_i. □

LEMMA 6.5. Let R be a semisimple finite-dimensional k-algebra and M be a finite-dimensional right R-module. Then M is a direct sum of finitely many simple R-modules.

PROOF. Exercise (TODO). □

DEFINITION 6.6. A module which is a direct sum of finitely many simple modules is said to be completely reducible.

DEFINITION 6.7. The socle $\mathrm{soc}(M)$ of a nonzero finite-dimensional R-module is the sum of all its simple submodules.

LEMMA 6.8. We have $\mathrm{soc}(M) = \{m \in M \mid mJ(R) = 0\}$.

PROOF. Each simple submodule is annihilated by $J(R)$ as we have seen. Hence, $J(R)$ annihilates $\mathrm{soc}(M)$.

Conversely, if $mJ(R) = 0$, then mR may be regarded as a $R/J(R)$-module. By 6.5, it is a sum of simple modules. Hence $mR \subseteq \mathrm{soc}(M)$. □

REMARK. We have that $\mathrm{soc}(M)$ is a sum of simple modules and is completely reducible.

DEFINITION 6.9. The socle series of a module M is

$$0 = \mathrm{soc}_0(M) \subsetneq \mathrm{soc}_1(M) \subsetneq \cdots,$$

where $\mathrm{soc}_i(M)/\mathrm{soc}_{i-1}(M) = \mathrm{soc}(M/\mathrm{soc}_{i-1}(M))$ (i.e., to define $\mathrm{soc}_i(M)$, pull back $\mathrm{soc}(M/\mathrm{soc}_{i-1}(M))$, which is a submodule of $M/\mathrm{soc}_{i-1}(M)$, along the quotient map), as long as $\mathrm{soc}_{i-1}(M) \neq M$.

REMARK. (1) The series is strict until $\mathrm{soc}_i(M) = M$.

(2) By the previous lemma, $\mathrm{soc}_i(M) = \{m \in M \mid mJ(R)^i = 0\}$. Indeed, the case $i = 0$ is trivial, and if we know that the claim is true for i and if $\pi\colon M \to M/\mathrm{soc}_i(M)$ is the quotient map, then by definition and the inductive claim we have

$$\begin{aligned}
\mathrm{soc}_{i+1}(M) &= \pi^{-1}(\mathrm{soc}(M/\mathrm{soc}_i(M))) \\
&= \{m \in M \mid (m + \mathrm{soc}_i(M))J(R) = 0\} \\
&= \{m \in M \mid mJ(R) \in \mathrm{soc}_i(M)\} \\
&= \{m \in M \mid mJ(R)^{i+1}\}
\end{aligned}$$

as required.

PROPOSITTON 6.10. Let R be a finitely dimensional algebra. Then $J(R)$ is nilpotent.

PROOF. Let $J = J(R)$. Consider $J \supsetneq J \supseteq J^2 \supseteq \cdots$. This must terminate by finite-dimensionality, so $J^n = J^{n+1}$ for some n. Hence, the sockle series must terminate by the previous remark. Hence, $R = \mathrm{soc}_n(R)$. But then J^n annihilates 1, so $J^n = 0$ as required. □

1. The Artin-Wedderburn theorem

LEMMA 6.11 (Schur's lemma). Let S be a simple right R-module. Then $\mathrm{End}_R(S)$ is a division ring (by simplicity of S). If S_1 and S_2 are non-isomorphic simple right R-modules, then $\mathrm{Hom}_R(S_1, S_2) = 0$.

Note that S is a left $\mathrm{End}_R(S)$-module.

PROOF. Let $\phi\colon S \to S$ be an R-linear map. Then either $\phi(S) = 0$ and hence $\phi = 0$ or $0 \neq \phi(S) = S$ since S is simple. Furthermore, $\ker \phi$ is a submodule of S, so it is either 0 or S. Hence, if $\phi \neq 0$, it is an isomorphism (so it has a two-sided inverse). Thus, $\mathrm{End}_R(S)$ is a division ring.

If S_1 and S_2 are non-isomorphic simple right R-modules and $0 \neq \phi\colon S_1 \to S_2$ is R-linear, then $\mathrm{im}\,\phi = S_2$ and $\ker \phi = 0$, hence ϕ is an isomorphism, which is a contradiction. □

LEMMA 6.12. Denote by R_R the right R-module R. Then $\mathrm{End}_R(R_R) \cong R$ via multiplication on the left by $r \in R$.

PROOF. A morphism $\phi \in \mathrm{End}_R(R_R)$ is determined by $\phi(1)$. The map $\mathrm{End}(R_R) \to R$, $\phi \mapsto \phi(1)$ is an isomorphism, noting that multiplication on the left by $r \in R$ is an endomorphism. □

THEOREM 6.13 (Artin-Wedderburn theorem). Let R be a semisimple finite-dimensional associative algebra. Then $R = \bigoplus_{i=1}^r R_i$, where $R_i = M_{n_i}(D_i)$, where D_i is a finite-dimensional division algebra. Moreover, the R_i are uniquely determined.

R has exactly r isomorphism classes of right simple modules S_i and $D_i = \mathrm{End}_R(S_i)$ and $n_i = \dim_{D_i}(S_i)$.

Furthermore, if k is algebraically closed, then $D_i \cong k$ for each i and thus $\mathbb{C}G$ for a finite group G is $\bigoplus_{i=1}^r M_{n_i}(\mathbb{C})$, where $\mathbb{C}G$ has r isomorphism classes of simple modules of degree n_i.

PROOF. Since R is semisimple, by 6.5 R_R is a finite direct sum of simple right R-modules.

Group together those that are isomorphic.

$$R_R \cong (S_{11} \oplus \cdots \oplus S_{1n_1}) \oplus (S_{21} \oplus \cdots \oplus S_{2n_2}) \cdots$$

so that $S_{ij} \cong S_{kl}$ if and only if $i = k$ and define $S_i := S_{i1}$.

Define $R_i := S_{i1} \oplus \cdots \oplus S_{in_i}$. Then $R_R = \bigoplus_{i=1}^r R_i$.

Let S be a simple R-submodule of R_R. Consider the projections $\pi_{ik}\colon R \to S_{ik}$ restricted to S. By Schur's lemma, $\pi_{ik}|_S$ is either an isomorphism or the zero map. Note that at least one of these restrictions must be nonzero, since S is non-zero. We deduce that $\pi_{ik}|_S$ is non-zero for exactly one i (and possibly several k) and thus we deduce that $S \subseteq R_i$. Thus R_i can be expressed as the sum of the simple submodules of R_R isomorphic to S_i and is hence uniquely determined.

Consider $\mathrm{End}_R(R_i) = \mathrm{End}_R(S_{i1} \oplus \cdots \oplus S_{in_i}) \cong M_{n_i}(D_i)$, where $M_{n_i}(D_i)$, where $D_i = \mathrm{End}_R(S_i)$ by Schur, which is a division algebra (also by Schur). Indeed, $\phi \in \mathrm{End}_R(S_{i1} \oplus \cdots \oplus S_{in_i})$ is represented by a matrix $(\phi_{m\ell})$, where $\phi_{m\ell} \in \mathrm{Hom}(S_{im}, S_{i\ell})$. Hence $R \cong \mathrm{End}_R(R_R)$ (using 6.12) is a matrix algebra consisting of block diagonal matrices with blocks $M_{n_i}(D_i)$, where the other blocks of the form $\mathrm{Hom}_R(S_{ij}, S_{kl})$ with $i \neq k$ are zero by Schur.

Now recall the example about right ideals of $M_n(D)$. The minimal right ideals consist of matrices B with columns of the form $\begin{pmatrix} d_1 & \cdots & d_n \end{pmatrix}^\top \lambda$, where $\lambda \in D$ and the column vector is fixed.

The simple right submodules of $M_{n_i}(D_i)$ are all of dimension n_i as a D_i-vector space and so $\dim_{D_i}(S_i) = n_i$.

Finally, it remains to show that if D is a division algebra over an algebraically closed field k, then $D = k$. Indeed, let $x \in D$. Consider $k(x) \subseteq D$. This is a finite extension of k (since D is finite-dimensional), so it is in particular algebraic. Hence, we find a monic polynomial $f \in k[X]$ such that $f(x) = 0$. Choose f of minimal degree. Since k is algebraically closed, it has a root $\lambda \in k$ and we may write $f = g(X - \lambda)$. By minimality, $g(x) \neq 0$, so $x = \lambda \in k$, so $D = k$. □

REMARK. If k is a finite field and D is a finite-dimensional division k-algebra then D is a finite field. This is Wedderburn's little theorem (1905). We will not prove it in this course (though it admits short elementary proofs).

COROLLARY 6.14. If G is a finite group, then $Z(\mathbb{C}G)$ is an r-dimensional $\mathbb{C}$-vector space, where r is the number of isomorphism classes of simple modules, which coincides with the number of conjugacy classes of G.

PROOF. Any class sum $\sum_{g' \in \mathrm{ccl}(g)} g'$ is contained in $Z(\mathbb{C}G)$. Moreover, any element of $Z(\mathbb{C}G)$ is a linear combination of class sums.

Linear independence of the class sums is clear. Hence, the class sums form a basis of $Z(\mathbb{C}G)$, so $\dim Z(\mathbb{C}G)$ is just the number of conjugacy classes.

But $Z(M_n(\mathbb{C})) = \{\lambda I \mid \lambda \in \mathbb{C}\}$ and so $\dim Z(M_n(\mathbb{C})) = 1$. By Artin-Wedderburn, $\mathbb{C}G \cong \bigoplus_{i=1}^r M_{n_i}(\mathbb{C})$, so $Z(\mathbb{C}G) = \bigoplus_{i=1}^r Z(M_{n_i}(\mathbb{C}))$ and $\dim Z(\mathbb{C}G) = r$ is just the number of isomorphism classes of simples modules. □

EXAMPLE. Consider $G = S_3$ and let k be an algebraically closed field. Denote $g := \begin{pmatrix} 1 & 2 \end{pmatrix}$, $h := \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$. In characteristic 0 there are three simple kG-modules up to isomorphism, hence there are three conjugacy classes. We have the trivial one-dimensional module $U_1 = k$, on which g, h act like 1. We also have the one-dimensional module U_2 on which g acts like -1 and h acts like 1. Finally, we have the two-dimensional module U_2. We write elements of U_2 as row vectors $\begin{pmatrix} \lambda & \mu \end{pmatrix}$ and G acts on the right. The element g acts as $\begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}$ and h acts as $\begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}$. Geometrically, g is a reflection and h is a rotation.

If $\operatorname{char} k = 2$ or $\operatorname{char} k = 3$, then we can work modulo 2 or 3. For example, in characteristic 2, we have $\overline{U_1} = \overline{U_2}$, but $\overline{U_3}$ remains simple. Hence, we have at least two simple modules. By Artin-Wedderburn, $kG/J(kG) \cong M_1(k) \oplus M_2(k)$, which has dimension 5, and we can't have anything else because then $J(kG) = 0$, which is not the case.

Hence $J(kG)$ is one-dimensional, and the group sum $1 + h + h^2 + g + gh + gh^2$ is contained in the centre and spans $J(kG)$.

Furthermore, $\mathrm{soc}(kG) = \{r \in kG \mid rJ(kG) = 0\}$ is the span of $\gamma - 1$, where $\gamma \in G$. This is precisely the kernel of the map $kG \to k$ sending $\gamma \mapsto 1$.

The characteristic 3 case will apear on the example sheet. Some hints: $\overline{U_1}$ and $\overline{U_2}$ are simple and not isomorphic, but $\overline{U_3}$ is not simple, as g and h have a common eigenvector.

We find that $kS_3/J(kS_3) \cong M_1(k) \oplus M_1(k)$ and $J(kS_3)$ is the kernel of the map $kS_3 \to kC_2$ which sends even permutations to 1 and odd permutations to the generator α of C_2.

2. Indecomposable Modules

REMARK. In the introduction, we called a non-zero R-module M indecomposable if it is not expressible as $M = M_1 \oplus M_2$ with M_1, M_2 non-zero.

DEFINITION 6.15. (a) M has the unique decomposition property if
  (i) M is a finite direct sum of indecomposable modules, and
  (ii) whenever $M = \bigoplus_{i=1}^m M_i = \bigoplus_{i=1}^n M_i'$, with M_i, M_i' indecomposable, then $m = n$ and we can reorder them in such a way that $M_i \cong M_i'$.
(b) A ring R has a the unique decomposition property if every finitely generated R-module does.

DEFINITION 6.16. A ring R is called local if it has a unique maximal right ideal. This ideal is then necessarily equal to $J(R)$ and it is also the unique maximal left ideal, because its complement consists of units.

REMARK. If R is local, then $R/J(R)$ is a division ring and every element not in $J(R)$ is invertible.

THEOREM 6.17 (Krull-Schmidt). Suppose M is a finite direct sum of indecomposable right R-modules M_i with each $\operatorname{End}_R(M_i)$ local. Then M has the unique decomposition property.

PROOF. Let $M = \bigoplus_{i=1}^m M_i = \bigoplus_{i=1}^n M_i'$. We will proceed by induction on m. If $m = 1$, then M is indecomposable, so the claim follows.

Next, assume that $m > 1$. Let

$$\alpha_i \colon M_i' \to M \to M_1 \qquad \beta_i \colon M_1 \to M \to M_i'$$

be the obvious maps. Then $\sum(\alpha_i \circ \beta_i)$ is the identity on M_1. Now since $\operatorname{End}_R(M_1)$ is local by assumption, some of the $\alpha_i \circ \beta_i$ must be invertible. Otherwise, they would all be in the maximal ideal, so their sum would be in the ideal and thus not equal to the identity.

This means that we find a left inverse $f \colon M_1 \to M_1$ of $\alpha_i \circ \beta_i$. But then the short exact sequence

$$0 \longrightarrow \operatorname{im}\beta_i \overset{\iota}{\hookrightarrow} M_i' \longrightarrow \operatorname{coker}\iota \longrightarrow 0$$

splits on the left via the map $\beta_i \circ f \circ \alpha_i \colon M_i' \to \operatorname{im}\beta_i$, hence $M_i' \cong \operatorname{im}\beta_i \oplus \operatorname{coker}\iota$ by the splitting lemma, so we conclude $M_i' = \operatorname{im}\beta_i$ by indecomposability. Hence β_i is surjective, so β_i is an isomorphism, so $\alpha_i = \alpha_i \circ \beta_i \circ \beta_i^{-1}$ is also an isomorphism.

Renumber the M_i' such that $\alpha_i \circ \beta_i$ is invertible and $M_1 \cong M_1'$. Our next will be to find an R-linear automorphism of M sending $M_1 \to M_1'$.

Consider the map $\mu = 1 - \theta$, where θ is the composite

$$M \twoheadrightarrow M_1 \xrightarrow{\alpha_1^{-1}} M_1' \hookrightarrow M \twoheadrightarrow \bigoplus_{i=2}^m M_i \hookrightarrow M.$$

Observe that $\mu(M_1') = M_1$. Indeed, on M_1', θ acts like $M_1' \to M \to \bigoplus_{i=2}^m M_i$, so on M_1' μ acts like, $M_1' \to M \to M_1$, and we know that this is surjective. Moreover, $\mu(\bigoplus_{i=2}^m M_i) = \bigoplus_{i=2}^m M_i$, since θ vanishes on this submodule (this is obvious: it is precisely the kernel of the first map). This shows that μ is surjective.

Next, if $\mu(x) = 0$, then $x = \theta(x)$, so $x \in \bigoplus_{i=2}^m M_i$ (since that is the image of the final map). But then $x = \theta(x) = 0$ as seen above. Hence μ is an automorphism of M satisfying $\mu(M_1') = M_1$. Hence

$$\bigoplus_{i=2}^n M_i' \cong M/M_1' \cong M/M_1 \cong \bigoplus_{i=2}^m M_i,$$

using μ in the second step. Hence, we are done using the inductive step. □

LEMMA 6.18 (Fitting). Suppose and R-module M is finite-dimensional over k and $f \in \operatorname{End}_R(M)$. Then for sufficiently large n we have $M = \ker f^n \oplus \operatorname{im} f^n$.

PROOF. Choose n large enough so that $f^n \colon f^n(M) \to f^{2n}(M)$ is an isomorphism. This is possible by finite-dimensionality.

Let $m \in M$ and write $f^n(m) = f^{2n}(m_1)$. Then $m = f^n(m_1) + (m - f^n(m_1))$, but the first summand is in the image of f^n and the second summand is in the kernel of f^n. If $f^n(x) \in \ker f^n$, then $f^n(x) = 0$ since $f^n|_{\operatorname{im} f^n}$ is an isomorphism. Hence $\operatorname{im} f^n \cap \ker f^n = 0$, completing the proof. □

LEMMA 6.19. Suppose M is an indecomposable R-module that is finite-dimensional over k. Then $\operatorname{End}_R(M)$ is local.

PROOF. Let $E = \mathrm{End}_R(M)$. Choose a maximal right ideal I of E and take $x \notin I$. Our goal is to show that x is invertible, which will imply that I is the unique maximal right ideal. Then $E = xE + I$ by maximality of I. In particular, $1 = x\lambda + \mu$ for some $\lambda \in E$ and $\mu \in I$. By Fitting's lemma, $M = \ker \mu^n \oplus \operatorname{im} \mu^n$ for some n. These summands are R-modules, so by idecomposability of M we have $\ker \mu^n = M$ or $\operatorname{im} \mu^n = M$. The latter implies that μ is invertible with inverse μ^{-1}. But if $\mu \in \mathrm{End}_R(M)$, then we must have $\mu^{-1} \in \mathrm{End}_R(M)$ and so $\mu E = E$, which is a contradiction to $\mu \in I$.

Hence, we must have $\ker \mu^n = M$. Thus, $x\lambda = 1-\mu$ has inverse $1+\mu+\ldots+\mu^{n-1}$ and so x is invertible as claimed. □

COROLLARY 6.20. Let R be a finite-dimensional k-algebra. Then R has the unique decomposition property.

PROOF. This follows by combining 6.17 and 6.19. □

REMARK. An R-R-bimodule M is an abelian group which is both a left R-module and a right R-module with the obvious associativity property: $(rm)s = r(ms)$.

A right R-module can be thought of a left R^{op}-module, where R^{op} is the opposite ring.

Thus an R-R-bimodule may be viewed as a left $R^{\mathsf{op}} \otimes_k R$-module. Similarly it is a right $R^{\mathsf{op}} \otimes_k R$-module.

For example, $(kG)^{\mathsf{op}} \cong kG$ via the map $g \mapsto g^{-1}$.

If R is finite-dimensional, then $R^{\mathsf{op}} \otimes R$ is finite-dimensional, so the unique decomposition property holds by Krull-Schmidt.

Hence we have a decomposition $R \cong \bigoplus B_j$ into indecomposable ideals (i.e., sub-bimodules, i.e., $R^{\mathsf{op}} \otimes R$-submodules) that are unique up to reordering.

DEFINITION 6.21. The blocks of an algebra are the indecomposable ideals above.

REMARK. Let $R \cong \bigoplus B_j$ be a block decomposition. Write $1 = e_1 + \cdots + e_n$, where $e_j \in B_j$. Then $e_i e_j \in B_i \cap B_j = 0$ if $i \neq j$. Hence $1 = 1^2 = e_1^2 + \cdots + e_n^2$ with $e_j^2 + B_j$. Since we have a direct sum, comparison of components yields $e_j^2 = e_j$, i.e., e_j is an idempotent. Now if $r = r_1 + \cdots + r_n \in R$ with $r_i \in B_j$, then

$$1r = (e_1 + \cdots + e_n)(r_1 + \cdots + r_n) = \sum_j e_j r_j,$$

again since $e_i r_j = 0$ for $i \neq j$. Similarly, $r1 = \sum r_j e_j$. By comparison of components, we conclude $re_j = e_j r$ for all r. In order words, the e_j are central idempotents. Each e_j is the multiplicative identity of the block $B_j = e_j R$ (the last equality is obvious: if $b \in B_j$, then $b = 1b = e_j b \in e_j R$).

Conversely, if $1 = e_1 + \cdots + e_n$ with $e_i e_j = 0$ for $i \neq j$, $e_j^2 = e_j$ and e_j central, then $R = \bigoplus_j e_j R$, where the $e_j R$ are ideals since e_j is central.

EXAMPLE. Consider the group algebra kS_3, where $\operatorname{char} k = 2$. Let h denote a 3-cycle and g denote a transposition. The element $e_1 := 1 + h + h^2$ is a linear combination of class sums and $e_2 := h+h^2$ is a class sum, hence they are both central. It is easy to verify that they are idempotent. Since we are in characteristic two, their sum is 1 and we can calculate that their product vanishes. Write $B_1 := e_1 kS_3$, $B_2 := e_2 kS_3$. Then these are indecomposable ideals and $kS_3 = B_1 \oplus B_2$. We may verify that $B_1 \cong kC_2$ and B_2 is isomorphic to the matrix algebra $M_2(k)$.

As an exercise (TODO), consider the case that $\operatorname{char} k = 3$ and show that there is one block. As a hint, try to calculate the central idempotents.

REMARK. Recall that $R/J(R)$ is semisimple. If $R = \bigoplus B_j$, then $R/J(R) = \bigoplus B_j/B_j J(R)$, where the $B_j/B_j J(R)$ are semisimple R-modules (TODO: why?

See Drozd, Kirichenko: Finite Dimensional Algebras). Recall that $B_j = e_j R$, so $B_j/B_j J(R) = e_j R/e_j J(R)$.

Now by Artin-Wedderburn, each $B_j/B_j J(R)$ is a direct sum of matrix algebras—it may be a sum of various of the matrix algebras associated with simple R-modules S_i (TODO: why R-modules?). The point here is that while B_i is indecomposable, $B_i/B_j J(R)$ will in general not be indecomposable.

The following question arises: when does the matrix algebra associated with S_i appear in the same $B_j/B_j J(R)$ as the matrix algebra associated with S_ℓ?

One way to tackle this question is using the Ext quiver. Its vertices are labelled by the isomorphism classes of simple modules S_i. The matrix algebras associated with S_i and S_j appear in the same $B_j/B_j J(R)$ if and only if S_i and S_ℓ are in the same component of the Ext quiver.' In other words, blocks correspond to components.

CHAPTER 7

Quivers

DEFINITION 7.1. A quiver Q is a directed multigraph. It has vertices (usually labelled i, j, $\dots$) and arrows. There is no restriction on the number of arrows between any ordered pair of vertices (in particular, we allow loops).

As a matter of notation, if $i \xrightarrow{x} j$ is an arrow, we say that i is the source of x and j is the target of x.

DEFINITION 7.2. A representation M of a quiver Q is a direct sum of k-vector spaces M_i for each vertex i of Q, i.e., $M = \bigoplus_i M_i$, together with a linear map $\theta_x \colon M_i \to M_j$ for every arrow $x \colon i \to j$.

EXAMPLE. Consider the quiver with vertices $1, 2$ and two arrows $x, y \colon 1 \to 2$. Let $M_1 = k$, $M_2 = 0$, then we must define θ_x and θ_y to be zero. We have $M = M_1 \bigoplus M_2$.

DEFINITION 7.3. A morphism of representations $M \to M'$ of a common quiver Q is a collection of linear maps $M_i \to M_i'$ that commute with the linear maps θ_x, θ_x'.

DEFINITION 7.4. A path of length $\ell \geq 1$ is a concatenation of ℓ compatible arrows, i.e., the target of one arrows is the source of the next.

A path of length 0 is a vertex.

DEFINITION 7.5. The path algebra kQ is the k-vector space with basis labelled by paths of any length (including 0). The multiplication is given on basis elements by concatenation of paths. If two paths cannot be concatenated, then the product is defined to be zero.

In the previous example, we have two paths of length 0, e_1 and e_2 and two paths of length 1, x and y, and no paths of any other length. The products come out to $e_1 x = x$, $e_y = y$, $e_2 x = 0$, $e_2 y = 0$, $x e_1 = 0$, $x e_2 = x$, $y e_1 = 0$, $y e_2 = y$, $xy = 0$, $yx = 0$, $e_1^2 = e_1$, $e_2^2 = e_2$.

Observe that paths of length 0 are idempotent elements of the path algebra.

LEMMA 7.6. (a) kQ is a finite-dimensional k-algebra if and only if Q is finite and has no directed cycles.
(b) kQ is finitely generated as a k-algebra if and only if Q is finite.

PROOF. The first claim is immediate since kQ is finite-dimensional if and only if there are finitely many paths.

The second claim follows since kQ is generated by the paths of length 0 and 1. $\square$

REMARK. Suppose $M = \bigoplus M_i$ is a representation of a quiver Q. If $x \colon i \to j$ is an arrow, then x acts on M_i (on the right!) via θ_x and on all other M_ℓ as zero.

Also, e_i acts on M via the projection onto the component M_i.

This definition makes $\bigoplus M_i$ into a right module over the path algebra kQ. In fact, we have a correspondence between kQ-modules and quiver representations.

Note that there are simple kQ-modules S_i for each vertex i corresponding to the representation which is k at the vertex i, zero everywhere else and has all maps trivial.

Example. Consider again the quiver from before. Then the module e_1kQ corresponds to the representation which has $M_1 = k$, $M_2 = k \oplus k$, θ_x is the first inclusion, and θ_y is the second inclusion.

Indeed, $e_1kG = \{e_1p \mid p \in kG\}$ is the k-span of the set of paths starting at 1. In our case, we e_1kG is generated as a k-vector space by e_i, x and y. Now the corresponding representation has $M = e_1kG$ and $M_i = Me_i$, thus $M_1 = \langle e_1 \rangle$ and $M_2 = \langle x, y \rangle$ (these are k-spans). For an arrow z, θ_z is given by $m \mapsto mz$. In our case, this means θ_x maps $e_1 \mapsto e_1x = x$ and θ_y maps $e_1 \mapsto e_1y = y$, completing the proof.

Example. Consider now that quiver Q which has a single vertex 1 and a loop $x\colon 1 \to 1$. Consider the representation given by $M_1 = k$ and $\theta_x\colon \lambda \mapsto \lambda\mu$ for some fixed $\mu \in k$.

The corresponding kQ-module is indecomposable, and for distinct μ_1, μ_2 these representations are not isomorphic. Indeed, if $\sigma\colon M_1^{\mu_1} \to M_1^{\mu_2}$ is a morphism of representations, then the commutativity condition in the definition of representation says that $\mu_1\sigma(1) = \mu_2\sigma(2)$. Hence, if $\mu_1 \neq \mu_2$, then $\sigma(1) = 0$, hence σ is not an isomorphism.

In particular, if k is infinite, then there are infinitely many indecomposables.

This generalises to the case where Q has a directed cycle: take a directed cycle, put a copy of k at each vertex of the cycle, and make each arrow of the path act like the identity, except for one, which acts as multiplication by some non-zero element. Make everything else zero. Then we get a representation, and it is indecomposable, since it is generated as a kQ-module by any non-zero element of any of the copies of k, and so it is generated by any non-zero element (projecting to a component using e_i if necessary). Again, different choices of the multiplicative constant lead to non-isomorphic representations.

Exercise (TODO): what if k is finite?

Next, let Q be the quiver considered at the beginning of the chapter. For every $\mu \in k$, we get a representation by setting $M_1 = M_2 = k$, $\theta_x = \mathrm{id}$ and $\theta_y\colon \lambda \mapsto \lambda\mu$. These representations are indecomposable for the same reasons as seen before, and they are pairwise non-isomorpic, since if $\mu_1, \mu_2 \in k$ and we have an isomorphism $\sigma\colon M^{\mu_1} \to M^{\mu_2}$, then the compatibility condition for θ_x says that $\sigma_1 = \sigma_2$ (as maps $k \to k$) and thus if we set $\lambda \coloneqq \sigma_1(1) = \sigma_2(1)$, then the compatibility condition for θ_y gives that $\lambda\mu_1 = \lambda\mu_2$. Since σ is an isomorphism, $\lambda \neq 0$, hence $\mu_1 = \mu_2$.

Example. Let Q be a finite quiver with no directed cycles (hence kQ is finite dimensional).

Let J be the k-span of paths of length $\ell \geq 1$. Then J^r is the k-span of paths of length $\ell \geq r$. Since Q is finite and has no directed cycles, we find $J^n = 0$ for some n. Thus $J \subseteq J(kQ)$ (for example using Nakayama's lemma and the standard telescoping trick).

Now $kQ/J \cong k \oplus \cdots \oplus k$, where we get one copy of k for every vertex.

Recall that S_i was the simple kQ-module corresponding to the representation that has k at vertex i, 0 everywhere else and all maps are zero.

Now if an element of $J(kQ)$ has a non-zero component for a path of length 0, then by multiplying with the appropriate e_i and scaling we find $e_i \in J(kQ)$ for some i. But then $e_iS_i \neq 0$, which is a contradiction since $J(kQ)$ is the intersection of the annihilators of the simple modules, so in particular e_i should annihilate S_i. Hence we conclude $J(kQ) \subseteq J$, i.e., $J = J(kQ)$.

Now using Artin-Wedderburn and the explicit description of kQ/J from above, we conclude that the S_i are actually all simple kQ-modules.

Definition 7.7. A finite-dimensional algebra R is called basic if $R/J(R)$ is a direct sum of copies of k.

Remark. This occurs when all simple modules are one-dimensional.

Example.
(1) kQ is basic when Q is finite and has no directed cycles.
(2) kS_3 is basic when $\operatorname{char} k = 3$. There were two simple modules which were both one-dimensional.

Remark. For any finite-dimensional k-algebra R, where k is algebraically closed, there is a basic algebra R_1 with R Morita equivalent (we will not define this) to R_1 (and so the category of finite-dimensional R-modules and R_1-modules are equivalent).

Definition 7.8. The Ext quiver of a finite-dimensional associative algebra R has vertices labelled by isomorphism classes of simple R-modules and the number of arrows $x\colon S_i \to S_j$ is the maximal number n of copies of S_j such that there is an indecomposable R-module X such that there is a short exact sequence

$$0 \longrightarrow \bigoplus_{k=1}^n S_j \longrightarrow X \longrightarrow S_i \longrightarrow 0.$$

We will show later that this number is always finite.

Remark. Since $J(R)$ annihilates S_i and S_j, we know that $J(R)^2$ annihilates X. The Ext quiver is just giving us information about $R/J(R)^2$.

There are other quivers, e.g., the Auslander-Reiten quiver, that give more information about R-modules.

Example.
(1) Let $R = k[X]/(X^p)$. Then there is only one isomorphism class of simple modules, the trivial module $S = k$ with X acting like 0[1].

    There is the indecomposable $X_1 := k[X]/(X^2)$ and a short exact sequence

$$0 \longrightarrow (X)/(X^2) \longrightarrow X_1 \longrightarrow k[X]/(X) \longrightarrow 0,$$

    where the left and right terms are both isomorphic to S. Hence, the Ext quiver has at least one arrow $S \to S$.

    By the structure theorem of modules over the PID $k[X]$, the finite-dimensional $k[X]/(X^p)$-modules are of the form $\bigoplus_i k[X]/(X^{r_i})$ for some $r_i \le p$. Hence, for X_i to be indecomposable, it must be of the form $k[X]/(X^r)$ for some r, but there is no injective map $k^n \to k[X]/(X^r)$ for any r and any $n \ge 2$, since the fact that X acts like 0 on the left hand side forces the entire map to have image in $(X^{r-1})/(X^r)$.

    Hence, the Ext-quiver consists of one vertex S with one arrow x.

    Thus, $kQ \cong k[X]$ as a k-algebra, and the kQ-modules are just representations M with a linear map $\theta_x\colon M \to M$.
(2) Let $R = kS_3$, where $\operatorname{char} k = 2$. There are two isomorphism classes of simple R-modules: the trivial S_1 and the two-dimensional S_3.

    We saw that $kS_3 = kC_2 \oplus M_2(k)$ was the block decomposition. In characteristic 2, we have $kC_2 = k[X]/(X-1)^2$. We have an exact sequence

$$0 \longrightarrow \frac{(X-1)}{(X-1)^2} \longrightarrow \frac{k[X]}{(X-1)^2} \longrightarrow \frac{k[X]}{(X-1)} \longrightarrow 0,$$

[1]Let S be a simple R-module. Suppose there is some $s \in S$ such that $Xs \neq 0$. Then $S = RXs$ by simplicity, so we find $r \in R$ such that $s = rXs = r^pX^ps = 0$, a contradiction. Hence X acts like 0 on S and may define a map $k \to S$ via $\lambda \mapsto \lambda x$, where x is any non-zero element of S. This map is R-linear. The kernel is an R-submodule, so it is a k-submodule, and it is not k, so it is zero. Since $S = Rx$ by simplicity, the map is surjective. Hence $S \cong k$ as an R-module as required.

and $kC_2 \cong k[X]/(X-1)^2$ is indecomposable. Hence there is a loop $x\colon S_1 \to S_1$ in the Ext quiver. In fact, the Ext quiver is given by the loop x and an isolated vertex for S_3 (TODO: work this out). Observe that we had two blocks and have two components in the Ext quiver.

As an additional exercise, work out the case where $\operatorname{char} k = 3$. In this case there are two isomorphism classes of simple modules, both of dimension 1, the trivial module S_1 and the signature S_2. One should show that there is an indecomposables X_1 that fits in a sequence

$$0 \longrightarrow k \longrightarrow X_1 \longrightarrow k \longrightarrow 0,$$

where the left k is the signature and the right k is trivial. Furthermore, one should show that there is X_2 that fits into

$$0 \longrightarrow k \longrightarrow X_2 \longrightarrow k \longrightarrow 0,$$

where this time the left k is trivial and the right k is the signature. Hence, the Ext quiver contains two vertices and at least one arrow in either direction. As an exercise (TODO) show that this is already the full Ext quiver.

There is one block and one component.

In both characteristics, there is a directed cycle, so the path algebra is not finite-dimensional in either case.

THEOREM 7.9 (Gabriel). Let R a basic finite-dimensional algebra with k algebraically closed. Then $R \cong kQ/I$, where Q is the Ext quiver and I is a suitable ideal of R such that $I \subseteq J(R)^2$.

NOT PROVED IN THIS COURSE. Recall from the definition of the Ext quiver that the Ext quiver is just giving info about $R/J(R)^2$. Hence, the Ext quiver of R is the Ext quiver of $R/J(R)^2$. □

THEOREM 7.10. Let S_1 and S_2 be two simple R-modules. Then S_1 and S_2 arise from the same block if and only if the vertices of the Ext quiver corresponding to S_1 and S_2 are in the same component of the Ext quiver (TODO: undirected component or strongly connected component)?

Hence, the blocks of R correspond to the components of the Ext quiver.

NOT PROVED IN THIS COURSE. □

DEFINITION 7.11. An algebra R has finite representation type if there are only finitely many indecomposable R-modules up to isomorphism.

EXAMPLE. We have previously seen that the path algebra kQ, where Q ia a quiver consisting of one vertex and a loop, does not have finite representation type.

THEOREM 7.12 (Gabriel 1972). Let Q be a connected quiver without directed cycles. If k is algebraically closed, then kQ has finite representation type if and only if the underlying graph of the quiver Q is A_r $(r \geq 1)$, D_r $(r \geq 4)$, E_6, E_7 or E_8. These are exactly the simply laced Coxeter graphs.

Here, the underlying graph of a Quivers is obtained by forgetting the direction of arrows.

Note that this is independent of the direction of the arrows in the quiver.

PROOF. First observe that if Q has finite representation type, then any subquiver obtained by removing a subset of the vertices is also of finite representation type (because any representation of the subquiver can be promoted to a representation on Q by making everything else zero. Of course, this representation is still indecomposable).

We have already seen that if we have a directed cycle, then Q does not have finite representation type. We have also seen that if any two vertices are connected by more than one edge, then the result does not have finite representation type. We deduce that any subquiver obtained as above leaving two vertices has at most one edge.

From this, we deduce that the underlying graph of Q must be a tree. Recall that we defined a symmetric bilinear form on the real span of the vertices via

$$q(v_i, v_j) = \begin{cases} 2, & i = j, \\ -1, & \text{there is an edge between } i \text{ and } j, \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that this symmetric bilinear form is not positive definite. Then we find non-negative integers k_i such that $q(v, v) \leq 0$ with $v = \sum k_i v_i \neq 0$[2].

Thus $2\sum k_i^2 \leq 2\sum_{i,\, j \text{ connected}} k_i k_j$, and hence

$$(\star) \qquad \sum k_i^2 \leq \sum_{i,\, j \text{ connected}} k_i k_j.$$

Let M_i be a vector space of dimension k_i. We will show that there are infinitely many isomorphism classes of representations with this dimension vector $\sum k_i v_i$.

We need to assigne linear maps $\theta_x \colon M_i \to M_j$ for each edge $x \colon i \to j$.

Two such representations are isomorphic if and only if there are automorphisms $\prod_i \mathrm{GL}(M_i)$ such that the diagram

$$\begin{array}{ccc} M_i & \xrightarrow{\theta_x} & M_j \\ \big\downarrow{\scriptstyle f_i} & & \big\downarrow{\scriptstyle f_j} \\ M_i' = M_i & \xrightarrow{\theta_x'} & M_j' = M_j \end{array}$$

commutes.

We will consider oribts of $\prod_i \mathrm{GL}(M_i)$ on $\prod_{x \colon i \to j} \mathrm{Hom}(M_i, M_j)$. Two representations are isomorphic if and only if the homomorphisms representing the arrows yield elements of $\prod_i \mathrm{Hom}(M_i, M_j)$ in the same orbit. But $\prod \mathrm{GL}(M_i')$ is an algebraic variety of dimension $\sum k_i^2$. The dimension of $\prod \mathrm{Hom}(M_i, M_j)$ is $\sum_{x \colon i \to j} k_i k_j$.

Notice that the scalar multiplication in $\prod \mathrm{GL}(M_i)$ act trivially on $\prod \mathrm{Hom}(M_i, M_j)$ and so we have that

$$V := \frac{\prod \mathrm{GL}(M_i)}{k^\times}$$

operates on the homs. We have $\dim V = \left(\sum k_i^2\right) - 1$. By $(\star)$, we find $\dim V < \sum k_i k_j = \dim \prod \mathrm{Hom}(k_i k_j)$, and so the orbits have dimension strictly less than $\dim \prod \mathrm{Hom}(M_i, M_j)$. This implies that we have infinitely many orbits, which means that there are infinitely many isomorphism classes of representations with dimension vector $\sum k_i v_i$.

But if Q has finite represenation type, then there are only finitely many indecomposable representations. By Krull-Schmidt, every representation decomposes as a direct sum of indecomposables in an essentially unique way. Hence, there are only finitely many representations of a given dimension up to isomorphism. In particular, there can only be finitely many isomorphism classes of representations of a given dimension vector, so we have arrived at a contradiction.

[2]To see that we may assume the k_i to be positive integers, approximate by a rational and multiply with a large number to get to integers. Then write $v = v_+ + v_-$, where v_+ is the sum of those $k_i v_i$ where $k_i > 0$ and v_- is defined analogously. Now $q(v, v) = q(v_+, v_+) + 2q(v_+, v_-) + q(v_-, v_-)$. But notice that the middle term is nonnegative, since $q(v_i, v_j)$ is zero or negative, and $k_i k_j$ is negative. Hence, either $q(v_+, v_+)$ or $q(v_-, v_-)$ is negative, so replace v either by v_+ or by $-v_-$.

Hence, the underlying graph is positive definite. □

REMARK. (1) There is an alternative proof analogous to the strategy used in the classification of positive definite Coxeter graphs. We find certain quivers that cannot appear as full subquivers.

(2) Suppose that k is not algebraically closed. In this situation, we should modify the definition of a basic algebra to say that $R/J(R)$ is a direct sum of division algebras and so has simple modules corresponding to each division algebra (which were the endomorphism rings of the simple R-modules).

So when we look at the Ext quiver, it is a good idea to include information in the quiver about the endomorphism algebras of the simple modules.

This additional information is of a similar sort to stipulating that vertices have different lengths with respect to the symmetric bilinear form. In fact, we get a positive definite quiver (with this additional information). We get a positive definite Coxeter graph (not necessarily simply laced).

REMARK. If working with k not algebraically closed (e.g., $\mathbb{R}$), then the other positive definite Coxeter graphs can arise and there is a more general theorem which sats that Q is of finite representation type if and only if the underlying graph of Q is a positive Coxeter graph.

REMARK. The underlying graph of a quiver which consists of two edges $x, y\colon i \to j$ is not the same as the Dynkin diagram consisting of two vertices and a double edge. In fact, we have proved above that this quiver does not have finite representation type.

REMARK. Given a quiver with r vertices, consider the $\mathbb{R}$-span of a basis v_i labelled by the vertices.

DEFINITION 7.13. The dimension vector of a representation M is $\sum(\dim M_i)v_i$, where $M = \bigoplus M_i$.

THEOREM 7.14 (Gabriel). Suppose the underlying graph Γ of a quiver Q is one of the simply laced positive definite Coxeter graphs. Then the isomorphism clasesses of indecomposable representations correspond to the positive roots. This gives a correspondence between dimension vectors of irreducible representations and positive roots, where $\sum k_i v_i$ corresponds to $\sum k_i\alpha_i$, where the α_i are the simple roots.

In particular, kQ has finite representation type.

PROOF. This proof is due to Bernstein-Gel'fand-Panomarev 1972.

Choose an admissible numbering of Q, let $\mathcal{C}^+$ be the corresponding Coxeter functor, and let c be the corresponding Coxeter element.

Suppose V is a finite-dimensional indecomposable representation with dimension vector $v \in \mathbb{R}^\ell$. From the discussion about Coxeter elements, there is some $m \geq 1$ such that $c^m(v)$ is not positive.

By 7.26, $(\mathcal{C}^+)^m(V) = 0$ (TODO: why?). Choose m as small as possible with $(\mathcal{C}^+)^m(V) = 0$. Thus for some j, we have that $\mathcal{S}^+_{j+1}\cdots\mathcal{S}^+_r(\mathcal{C}^+)^{m-1}(V) \neq 0$, but $\mathcal{S}^+_j\cdots\mathcal{S}^+_r(\mathcal{C}^+)^{m-1}(V) = 0$.

Thus by 7.22 we know that $\mathcal{S}^+_{j+1}\mathcal{S}^+_{j+2}\cdots\mathcal{S}^+_r(\mathcal{C}^+)^{m-1}(V)$ is the simple module concentrated at j (since that is the only thing mapped to zero) and $V \cong (\mathcal{C}^-)^{m-1}\mathcal{S}_r\cdots\mathcal{S}^-_{j+1}(S_j)$.

Thus the dimension vector v of V is given by

$$c^{-m+1}s_{\alpha_r}\cdots s_{\alpha_j+1}(v_j),$$

where v_j is the basis element belonging to the vertex j.

Note that this argument shows that any indecomposable with the same dimension vector as V is isomorphic to V.

Conversely, if v is a positive root, then for some $m \geq 1$ the element $c^m(v)$ is not positive. Choose the shortest expression of the form $s_{\alpha_j} \cdots s_{\alpha_r} c^{m-1}(v)$ which is not a positive root.

Then $s_{\alpha_{j+1}} \cdots s_{\alpha_r} c^{m-1}(v) = v_j$ using 5.17(d). Thus, $(\mathcal{C}^-)^{m-1} \mathcal{S}_r^- \cdots S_{j+1}^-(S_j)$ has dimension vector v. □

Remark. The proof is constructive. We have a recipe for producing the indecomposable representations with the given dimension vector.

Remark. This approach can be extended to infinite root systems (with other Dynkin diagrams/Coxeter graphs) and quivers (Kac 1980).

Definition 7.15. A vertex of Q is a sink if it is the target of all arrows meeting the vertex (i.e., it does not have any outgoing arrows).

Similarly, a vertex is a source if it does not have any incoming arrows.

Clearly, any finite quiver without directed cycles has at least one source and at least one sink.

Definition 7.16. Given a quiver Q, define a new quiver $s_i Q$ with the same vertices and edges, except that the direction of the edges touching vertex i are reversed.

For example, if $Q = 1 \to 2 \to 3$, then $s_2 Q = 1 \leftarrow 2 \leftarrow 3$ and $s_1 Q = 1 \leftarrow 2 \to 3$.

Remark. We can number the vertices of our Coxeter graphs such that the numbering gives a topological ordering. In particular, vertex 1 is a source and vertex r is a sink.

Definition 7.17. A quiver whose vertices have been numbered to give a topological ordering is called a standardised quiver.

Lemma 7.18. Let Q be a standardised quiver.

(i) If $1 \leq j < r$, then j is a sink and $j+1$ is a source of the quiver $s_j \cdots s_2 s_1 Q$.
(ii) If $1 < j \leq r$, then j is a source and $j-1$ is a sink of the quiver $s_j s_{j+1} \cdots s_{r-1} s_r Q$.
(iii) $s_r \cdots s_2 s_1 Q = s_1 s_2 \cdots s_r Q = Q$.

Proof. Follows from Q being standardised and that if $j_1, \ldots, j_s$ are distinct vertices, then $i \to j$ in $s_{j_1} s_{j_2} \cdots s_{j_s} Q$ if

- $i \to j$ in Q and either none or both of i, j appear in $j_1, \ldots, j_s$, or
- $i \leftarrow j$ in Q and exactly one of i, j appears in $j_1, \ldots, j_s$.

□

Definition 7.19. A numbering of vertices is called admissible if for each j we have that j is a sink of $s_{j+1} \cdots s_r Q$.

Lemma 7.20. There exists an admissible numbering of the vertices of Q if and only if there are no directed cycles in Q.

Proof. Use 7.18(ii) in the acyclic case. If there is an oriented cycle, then it is clearly impossible. □

Exercise 7.21. If Q and Q' have the same underlying graph, and the graph is a tree, then we find $j_1, \ldots, j_s$ such that $s_{j_1} \cdots s_{j_s} Q = Q'$.

Indeed, root Q at any vertex and work recursively. At a vertex v, if the parent edge is oriented in the wrong way, add s_v to the list and proceed with the childen. This clearly produces the desired transformation.

REMARK. In what follows, let j be a sink of the quiver Q.

DEFINITION 7.22. We define functors

$$\mathcal{S}_j^+ \colon \mathsf{Rep}_Q \to \mathsf{Rep}_{s_j Q}$$
$$\mathcal{S}_j^- \colon \mathsf{Rep}_{s_j Q} \to \mathsf{Rep}_Q.$$

Given a represenation of Q, we define $\mathcal{S}_j^+(V) \coloneqq W$, where $W_i = V_i$ for $i \neq j$, and W_j is the kernel of the map $\phi \coloneqq \bigoplus_{x\colon i\to j} \theta_x \colon \bigoplus_{x\colon i\to j} V_i \to V_j$. Hence, we have an exact sequence

$$(\dagger) \qquad 0 \longrightarrow W_j \longrightarrow \bigoplus_{x\colon i\to j} V_i \xrightarrow{\phi} V_j.$$

Observe that for each $x\colon i \to j$ (which implies $i \neq j$ since Q does not have directed cycles) there are obvious maps $W_j \to V_i = W_i$ given by the inclusion followed by the projection. This makes $W = \bigoplus W_i$ into a representation of $s_i Q$.

As an exercise (TODO), check that we get functorial induced morphisms.

The functor $\mathcal{S}_j^-$ is dual to this: given a representation of $s_j Q$, we let $V_i = W_i$ if $i \neq j$ and define V_j to be the cokernel of the direct sum of maps from j, i.e., we have an exact sequence

$$(\ddagger) \qquad W_j \xrightarrow{\psi} \bigoplus_{x\colon j\to i \text{ in } s_j Q} \longrightarrow V_j \longrightarrow 0.$$

If V is a representation of Q for which ϕ in ($\dagger$) is surjective, then $\mathcal{S}_j^- \mathcal{S}_j^+(V) = V$ (TODO: check this). Hence, $\mathcal{S}_j^+$ and $\mathcal{S}_j^-$ form an equivalence of categories between the subcategory of Rep_Q where ϕ is surjective and the subcategory of $\mathsf{Rep}_{s_j Q}$ where ψ is injective.

Now consider indecomposable representations. If ϕ is not surjective in ($\dagger$), then we can express V as a direct sum of representations $V = V' \oplus V''$, where V'' is a representation with $V_j'' = \operatorname{coker} \phi$ and 0 elsewhere. Also, V' is the same as V but with $\operatorname{im} \varphi$ at vertex j. If V is indecomposable, then either φ is surjective in ($\dagger$) or V is the simple representation with k at vertex j and 0 elsewhere.

LEMMA 7.23. The functors $\mathcal{S}_j^+$ and $\mathcal{S}_j^-$ give a bijection between indecomposable representations of Q not equal to a simple representation concentrated at j and indecomposable representations of $s_j Q$ not equal to a simple representation concentrated at j.

PROOF. This is just what we just saw. TODO: why do the functors map indecomposables to indecomposables? □

COROLLARY 7.24. The path algebra kQ has finite representation type if and only if $ks_j Q$ has finite representation type.

REMARK. Applying this to the exercise (about trees) we see that if two quivers have the same underlying graph which is a tree, then one has finite representation type if and only if the other has finite representation type.

REMARK. Next, we consider dimension vectors. If V is a representation of Q with ϕ surjective in ($\dagger$), then $\dim W_i = \dim V_i$ for $i \neq j$ and

$$\dim W_j = \left(\sum_{x\colon i\to j} \dim V_i \right) - \dim V_j.$$

Hence, the effect of applying $\mathcal{S}_j^+$ is to send the dimension vector of V to s_{α_j} applyied to the dimension vector of V, where s_{α_j} is the simple reflection labelled by α_i. Here we use that we have a simply laced Coxeter graph. TODO: figure out what is happening here.

Similarly, if W is a representation of s_jQ with ψ injective in (‡), then the effect on the dimension vector of W is the same as applying the simple reflection s_{α_j}.

This sort of extension to considering functors between various categories of representations with functors corresponding to reflections in Euclidean space is generally called categorification

Definition 7.25. A Coxeter element c of the Weyl group $W(\Phi)$ is one that is a product of each simple reflection exactly once (but in any order). The Coxeter elements are not unique, but they are conjugate in $W(\Phi)$ (TODO: show this), and so they have the same order. This order is called the Coxeter number h of $W(\Phi)$ (or of Φ).

Example. If Φ has type A_r, then we have $W(\Phi) \cong S_{r+1}$, and the simple reflections correspond to the transpositions $\begin{pmatrix} i & i+1 \end{pmatrix}$ for $1 \leq i \leq r$.

For example, when $r = 2$, i.e., $W(\Phi) = S_3$, then the simple reflections are $\begin{pmatrix} 1 & 2 \end{pmatrix}$, $\begin{pmatrix} 2 & 3 \end{pmatrix}$, and their product is a 3-cycle.

In general, we get an $(r+1)$-cycle.

The Coxeter number h is actually the number of roots divided by the rank r (TODO: why?), and so the dimension of the corresponding Lie algebra, which is the number of roots plus the rank, is $(h+1)r$.

Example.

Type	Coxeter number
A_r	$r+1$
B_r, C_r	2_r
D_r	$2r-2$
E_6	12
E_7	18
E_8	30
F_4	12
G_2	6

Remark. Coxeter elements were studied by Coxeter (1950s). He showed that given a Coxeter element c, there is a unique plane P on which c acts by rotation by $2\pi/h$.

A Coxeter element c has no nonzero fixed points in $\mathbb{R}^r$, and given any $v \in \mathbb{R}^r$ for some $m \geq 0$, the vector $c^m(v)$ is not positive (if $c^m(v)$ is positive for all m, then $\sum_{i=0}^{h-1} c^i(v)$ would be a non-zero fixed point of c, a contradiction).

Definition 7.26. Suppose $1, \ldots, n$ is an admissible numbering of the vertices of Q (recall: for each j, j is a sink of $s_{j+1}s_{j+2}\cdots s_rQ$). Then the Coxeter functor with respect to the numbering is

$$\mathcal{C}^+ := \mathcal{S}_1^+ \circ \cdots S_{r-1}^+ \circ S_r^+.$$

This is a functor $\mathsf{Rep}_Q \to \mathsf{Rep}_Q$, noting that each arrow has been turned around twice, so $s_1 \cdots s_r Q = Q$.

Similarly, we define

$$\mathcal{C}^- := \mathcal{S}_r^- \circ \cdots \circ \mathcal{S}_1^- : \mathsf{Rep}_Q \to \mathsf{Rep}_Q.$$

Lemma 7.27. Given any finite-dimensional indecomposable representation V of Q, one of the following is true:

(i) $\mathcal{C}^-\mathcal{C}^+(V) = V$,
(ii) $\mathcal{C}^+(V) = 0$.

Furthermore, in case (i), if v is the dimension vector of V, then the dimension vector of $\mathcal{C}^+(V)$ is given by $s_{\alpha_1} \cdots s_{\alpha_r}(v)$ for a Coxeter element $c = s_{\alpha_1} \cdots s_{\alpha_r}$.

PROOF. If any of the $\mathcal{S}^+_{j+1} \cdots \mathcal{S}_r(V)$ is the simple module concentrated at j, we are in case (ii).

Otherwise, we are in case (i) by the discussion in 7.22. $\square$

Exercises

Example Sheet 1

Exercise 2.

Exercise. There are exactly two Lie algebras of dimension 2 up to isomorphism.

Solution. Let L be a Lie algebra over k of dimension 2. If L is abelian, then L is isomorphic to k^2 with the trivial Lie bracket.

Otherwise, there are $x, y \in L$ such that $v := [x, y] \neq 0$. Since $v \neq 0$, x and y are linearly independent, so x and y form a basis of L and we have $v = \lambda_1 x + \lambda_2 y$ for some $\lambda_1, \lambda_2 \in k$ which are not both zero. We calculate

$$\begin{aligned} [v, x] &= [\lambda_1 x + \lambda_2 y, x] = [\lambda_1 x, x] + [\lambda_2 y, x] = -\lambda_2 v, \\ [v, y] &= [\lambda_1 x + \lambda_2 y, y] = [\lambda_1 x, y] + [\lambda_2 y, y] = \lambda_1 v. \end{aligned}$$

Now if $\lambda_1 \neq 0$, then setting $w := \lambda_1^{-1} y$, we find that $[v, w] = \lambda_1^{-1}[v, y] = v$. Hence L is isomorphic to k^2 with the bracket given by $[(1,0),(0,1)] = (1,0)$.

If $\lambda_1 = 0$, then we must have $\lambda_2 \neq 0$. Setting $w := -\lambda_2^{-1} x$, we find that $[v, w] = -\lambda_2^{-1}[v, x] = v$. Again, L is isomorphic to k^2 with the bracket given by $[(1,0),(0,1)] = (1,0)$. $\square$

Exercise 6.

Exercise. The Jacobi identity is equivalent to the adjoint representation being a homomorphism.

Solution. Indeed, if $x, y, z \in L$, then by definition of the adjoint representation, we have

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

$\square$

Exercise 7.

Exercise. $L^{(n)}$ lies in $L_{(2^n)}$ for all positive n.

Solution. We will first show that for natural numbers i and j we have $[\mathfrak{L}_{(i)}, \mathfrak{L}_{(j)}] \subseteq \mathfrak{L}_{(i+j)}$.

We do induction on j. The case $j = 1$ is true by definition.

Now assume that for some $j \in \mathbb{N}$ and all $i \in \mathbb{N}$ we have $[\mathfrak{L}_{(i)} + \mathfrak{L}_{(j)}] \subseteq \mathfrak{L}_{(i+j)}$. Let $i \in \mathbb{N}$. We need to show that $[\mathfrak{L}_{(i)}, \mathfrak{L}_{(j+1)}] \subseteq \mathfrak{L}_{(i+j+1)}$. We will check this on generators, so let $x \in \mathfrak{L}_{(i)}$, $y \in \mathfrak{L}_{(j)}$ and $z \in \mathfrak{L}$. We need to show that $[x, [y, z]] \in \mathfrak{L}_{(i+j+1)}$.

Indeed, $[x,y] \in \mathfrak{L}_{(i+j)}$ by our inductive hypothesis, so $\alpha := [z,[x,y]] \in \mathfrak{L}_{(i+j+1)}$ by definition. Furthermore, $[z,x]$ in $\mathfrak{L}_{(i+1)}$ by definition, so $\beta := [y,[z,x]] \in \mathfrak{L}_{(i+j+1)}$ by inductive hypothesis. Therefore $[x,[y,z]] = -\alpha - \beta \in \mathfrak{L}_{(i+j+1)}$ as required, completing the proof of the lemma.

Now we will proceed to prove the claim, again by induction. The base case is again trivial, and for $n \in \mathbb{N}$ we have

$$\mathfrak{L}^{(n+1)} = [\mathfrak{L}^{(n)}, \mathfrak{L}^{(n)}] \subseteq [\mathfrak{L}_{(2^n)}, \mathfrak{L}_{(2^n)}] \subseteq \mathfrak{L}_{(2^{n+1})},$$

using the inductive hypothesis and our lemma. This completes the proof. □

Exercise 9.

Lemma. If L is a Lie algebra and $L/Z(L)$, where $Z(L)$ is the centre of L, is nilpotent, then so is L.

Proof. If $L/Z(L)$ is nilpotency class n, then all expressions of the form $[[\cdots[[x_0,x_1],x_2]\cdots],x_n]$ are contained in $Z(L)$. Hence all expressions of the form $[[\cdots[[x_0,x_1],x_2]\cdots],x_{n+1}]$ vanish in L, i.e., L is nilpotent of nilpotency class at most $n+1$. □

Exercise. A finite-dimensional Lie algebra L is nilpotent if and only if $\operatorname{ad}(x)$ is nilpotent for all x in L.

Solution. If L is nilpotent, then $\operatorname{ad}(x)$ is obviously nilpotent for all $x \in L$.

Conversely, if $\operatorname{ad}(x)$ is nilpotent for all $x \in L$, then $\operatorname{ad}(L) \subseteq \operatorname{End} L$ satisfies the condition of Engel, hence by (2.14) $\operatorname{ad}(L)$ is isomorphic to a subalgebra of $\mathfrak{n}_n$ for some n. In particular $\operatorname{ad}(L)$ is nilpotent. But we have $L/Z(L) \cong \operatorname{ad}(L)$, so by the previous lemma L is nilpotent. □

Exercise 10.

Exercise. A finite-dimensional Lie algebra L is nilpotent if and only if it satisfies the idealiser condition.

Solution. If L is nilpotent, then by Exerice 9 we have that $\operatorname{ad}(x)$ is nilpotent for every $x \in L$. Let $S \subseteq L$ be a proper Lie subalgebra. Define

$$\begin{aligned} \rho\colon S &\to \operatorname{End} L/S \\ x &\mapsto (y+S \mapsto [x,y]+S), \end{aligned}$$

this is well-defined and a representation, because S is a subalgebra. We have that $\rho(S) \subseteq \operatorname{End} L/S$ consists of nilpotent endomorphisms. By Engel's theorem we find $y \in L \setminus S$ such that for all $x \in S$ we have $[x,y] + S = 0 + S$. Hence $y \in \operatorname{Id}(S) \setminus S$, so the idealiser condition is satisfied.

Conversely, assume that the idealiser condition is satisfied and $x \in L$. For a submodule S of L, define $\operatorname{Id}^0(S) := S$, $\operatorname{Id}^{n+1}(S) := \operatorname{Id}(\operatorname{Id}^n(S))$. We claim that if $y \in \operatorname{Id}^n(\langle x\rangle)$, then $\operatorname{ad}(x)^{n+1}(y) = 0$.

We will prove the claim by induction. If $y \in \operatorname{Id}^0(\langle x\rangle) = \langle x\rangle$, then $\operatorname{ad}(x)(y) = [x,y] = 0$. If the claim holds for $n \in \mathbb{N}_0$, let $y \in \operatorname{Id}^{n+1}(\langle x\rangle)$. By definition of the idealiser, we have that for any $z \in \operatorname{Id}^n(\langle x\rangle)$, $[y,z] \in \operatorname{Id}^n(\langle x\rangle)$. In particular, $x \in \operatorname{Id}^n(\langle x\rangle)$, so we find $[x,y] \in \operatorname{Id}^n(\langle x\rangle)$. Hence $\operatorname{ad}(x)^{n+2}(x)(y) = \operatorname{ad}^{n+1}(x)([x,y]) = 0$ by the inductive hypothesis, completing the proof.

Consider the sequence

$$\operatorname{Id}^0(\langle x\rangle) \subseteq \operatorname{Id}^1(\langle x\rangle) \subseteq \dots.$$

By the idealiser condition and finite-dimensionality, we must have $\operatorname{Id}^n(\langle x\rangle) = L$ for n. Then $\operatorname{ad}(x)^{n+1}(L) = 0$, so $\operatorname{ad}(x)^{n+1} = 0$, so $\operatorname{ad}(x)$ is nilpotent. By Exercise 9, we conclude that L is nilpotent. □

Exercise 11.

EXERCISE. All irreducible finite-dimensional representations of complex soluble Lie algebras are one-dimensionsal.

SOLUTION. Let $\rho\colon L \to \operatorname{End} V$ be a finite-dimensional representation. The Lie algebra $\rho(L) \subseteq \operatorname{End} V$ is isomorphic to a quotient of $\mathfrak{L}$, hence soluble by (2.9). By Lie's theorem, we find $0 \neq v \in V$ such that $\langle v \rangle$ is a ρ-invariant subspace, hence $V = \langle v \rangle$. $\square$

Exercise 12.

EXERCISE. (a) If L is the 3-dimensional Heisenberg Lie algebra, then there is a Lie algebra representation $\rho\colon L \to \operatorname{End}(k[X])$ such that x is mapped to $\frac{\mathrm{d}}{\mathrm{d}X}$, y is mapped to multiplication by X and z maps to the identity map.
(b) In characteristic $p > 0$ the ideal (X^p) of $k[X]$ is mapped into itself by the image of ρ, hence ρ induces a representation $\theta\colon L \to \operatorname{End}(k[X]/(X^p))$.
(c) θ is irreducible.

SOLUTION. (a) Easy verification.
(b) The claim is obvious for $\rho(y)$ and $\rho(z)$, and for $fX^p \in (X^p)$ we have

$$\frac{\mathrm{d}}{\mathrm{d}X}(fX^p) = \left(\frac{\mathrm{d}}{\mathrm{d}X}f\right)X^p + f\frac{\mathrm{d}}{\mathrm{d}X}X^p,$$

and the left summand is clearly in (X^p), and since we're in characteristic p, the right summand vanishes, hence the claim follows.
(c) Let $V \subseteq k[X]/(X^p)$ be a nontrivial θ-subspace. Then we find $0 \neq f \in V$. By repeatedly applying $\rho(x)$ to f we find that V contains (an element represented by) a nonzero constant (we use here that k does not have zero divisors), hence $1 + (X^p) \in V$. By repeatedly applying $\rho(y)$ we find that $X^i + (X^p) \in V$ for all $0 \leq i < p$, hence V contains a basis of $k[X]/(X^p)$ and thus $V = k[X]/(X^p)$, so θ is indeed irreducible. $\square$

Exercise 13.

EXERCISE. Let J be a Lie ideal of a Lie algebra L equipped with an invariant symmetric bilinear form $\langle\ ,\ \rangle$. Then $J^\perp$ is a Lie ideal.

Furthermore, the restriction to J of the Killing form on L is the Killing form on J.

SOLUTION. Let $x \in J^\perp$, $y \in L$. We will show that $[x, y] \in J^\perp$. Indeed, let $z \in J$. Then $[y, z] = -[z, y] \in J$ since J is a Lie ideal. But then, using invariance we have $\langle [x, y], z\rangle = \langle x, [y, z]\rangle = 0$ since $x \in J^\perp$. Hence $J^\perp$ is a Lie ideal.

Choose a basis $v_1, \ldots, v_n$ of L such that there is some $m \leq n$ such that $v_1, \ldots, v_m$ is a basis of J. Let x, y in J, and let M be the $m \times m$ matrix corresponding to $\operatorname{ad}_J(x) \circ \operatorname{ad}_J(y)$ under our basis. Since $\operatorname{ad}(y)(L) = [y, L] \subseteq J$ since J is a Lie ideal, the $n \times n$ matrix corresponding to $\operatorname{ad}_L(x) \circ \operatorname{ad}_L(y)$ under our basis has the block form

$$N = \begin{pmatrix} M & 0 \\ \star & 0 \end{pmatrix}.$$

Hence, if $\langle\ ,\ \rangle_J$ and $\langle\ ,\ \rangle_L$ denote the respective Killing forms, we have

$$\langle x, y\rangle_J = \operatorname{tr} M = \operatorname{tr} N = \langle x, y\rangle_L,$$

so the Killing form of J is the restriction of the Killing form of L to J. $\square$

Exercise 14.

Exercise. $\mathrm{ad}(L)$ is a Lie ideal of the Lie algebra of derivations $\mathrm{Der}\, L$ of the Lie algebra L.

Solution. First of all, let $x, y, z \in L$. Then we have

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

so $\mathrm{ad}(x)$ is a derivation and we conclude that $\mathrm{ad}(L) \subseteq \mathrm{Der}\, L$. Since adjoints are obviously closed under addition and scalar multiplication, $\mathrm{ad}(L)$ is a subspace of $\mathrm{Der}\, L$.

Furthermore, let $D \in \mathrm{Der}\, L$ and $x, y \in L$. Then we have

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

so we conclude that $[D, \mathrm{ad}(x)] = \mathrm{ad}(D(x))$, hence $\mathrm{ad}(L)$ is a Lie ideal of $\mathrm{Der}\, L$. $\square$

Exercise 15.

Exercise. Let L be the 3-dimensional Heisenberg Lie algebra. There are non-inner derivations of L and we can determine the Lie algebra $\mathrm{Der}\, L/\mathrm{ad}(L)$.

Solution. Let x, y, z denote a basis of the Heisenberg Lie algebra such that

$$[x,y] = z, \quad [x,z] = 0 \quad [y,z] = 0.$$

It immediately follows that $\mathrm{ad}(x)$ sends y to z and other basis elements to 0, $\mathrm{ad}(y)$ sends x to $-z$ and other basis elements to 0 and $\mathrm{ad}(z)$ is the zero derivation. Hence $\mathrm{ad}(L)$ is a two-dimensional subalgebra of $\mathrm{Der}\, L$.

On the other hand, if $\alpha, \beta, \gamma, a, b, c \in k$ we define

$$D(x) := \alpha x + \beta y + \gamma z, \quad D(y) := ax + bx + cx,$$

and since we want D to be a derivation, we must set

$$D(z) = D([x,y]) = [D(x), y] + [x, D(y)] = (\alpha + b)z.$$

Is is then easily cheked that the conditions on $D([x,z])$ and $D([y,z])$ are vacuous. Hence, we conclude that $\mathrm{Der}\, L$ consists of the endomorphisms that are precisely of the form above. In particular, $\mathrm{Der}\, L$ is a 6-dimensional Lie algebra, so there are derivations that are not inner (for example, the derivation given by $D(x) = z$, $D(y) = 0$, $D(z) = 0$).

Now $\mathrm{Der}\, L/\mathrm{ad}(L)$ is a 4-dimensional Lie algebra. We can give representatives $D, E, F, G \in \mathrm{Der}\, L$ whose images in the quotient form a basis by setting

$$\begin{array}{llll} D(x) = x & E(x) = 0 & F(x) = z & G(x) = 0 \\ D(y) = 0 & E(y) = z & F(y) = 0 & G(y) = y \\ D(z) = z & E(z) = 0 & F(z) = 0 & G(z) = z. \end{array}$$

We find that $[D, E] = E$ and $[F, G] = -F$ and all other Lie brackets of basis elements vanish. Hence, if L_2 is the non-abelian two-dimensional Lie algebra (cf. Exercise 2), then $\mathrm{Der}\, L/\mathrm{ad}(L) \cong L_2 \oplus L_2$. $\square$

Exercise 16.

Exercise. Let L be the non-abelian Lie algebra with basis x, y such that $[x, y] = y$. Then $\operatorname{Der} L = \operatorname{ad}(L)$.

Solution. Let $\alpha, \beta \in k$. We have

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

On the other hand, let $D\colon L \to L$ be any derivation. We have $\lambda_1, \lambda_2, \mu_1, \mu_2 \in k$ such that

$$D(x) = \lambda_1 x + \lambda_2 y, \quad D(y) = \mu_1 x + \mu_2 y.$$

We calculate

$$\begin{aligned}\mu_1 x + \mu_2 y = D(y) = D([x, y]) &= [D(x), y] + [x, D(y)]\\ &= [\lambda_1 x + \lambda_2 y, y] + [x, \mu_1 x + \mu_2 y] = \lambda_1 y + \mu_2 y.\end{aligned}$$

Hence $\mu_1 = \lambda_1 = 0$ and $D = \operatorname{ad}(\mu_2 x - \lambda_2 y)$, finishing the proof. □

Example Sheet 2

Exercise 9.

Exercise. Let $\alpha \neq \beta$ be roots of a semisimple Lie algebra with respect to a Cartan subalgebra H and suppose that neither $\alpha + \beta$ nor $\alpha - \beta$ is a root. Then $\langle h_\alpha, h_\beta \rangle_{\mathrm{ad}} = 0$.

Solution. The β-string through α consists of just α, so we calculate

$$\langle h_\alpha, h_\beta \rangle_{\mathrm{ad}} \overset{4.18(\mathrm{i})(\mathrm{a})}{=} \alpha(h_\beta) \overset{4.16(\mathrm{b})}{=} \frac{0}{2}\beta(h_\beta) = 0.$$ □

Example Sheet 3

Exercise 2.

Exercise. Let L be a semisimple Lie algebra with Cartan subalgebra H. Let L_α be a root space. Then the Lie subalgebra generated by L_α and $L_{-\alpha}$ is isomorphic to $\mathfrak{sl}_2$.

Solution. Let $e_\alpha \in L_\alpha$, $e_{-\alpha} \in L_{-\alpha}$ and $h_\alpha \in H$ be the elements from Lemma 4.18. It follows immediately from 4.18 that all of these elements are nonzero, so since $\dim L_{\pm\alpha} = 1$, these elements form a basis of the subalgebra generated by L_α and $L_{-\alpha}$, which we will call L.

We have $[h_\alpha, e_\alpha] \in L_\alpha$ by 4.13(b), so $[h_\alpha, e_\alpha] = \lambda e_\alpha$ for some $\lambda \in \mathbb{C}$. We may calculate

$$\lambda = \lambda \langle e_\alpha, e_{-\alpha} \rangle_{\mathrm{ad}} = \langle [h_\alpha, e_\alpha], e_{-\alpha} \rangle_{\mathrm{ad}} = \langle h_\alpha, [e_\alpha, e_{-\alpha}] \rangle_{\mathrm{ad}} = \langle h_\alpha, h_\alpha \rangle = \alpha(h_\alpha) \neq 0,$$

using 4.18(c), 4.18(a) and 4.15(b). By a similar argument, $[h_\alpha, e_{-\alpha}] = -\alpha(h_\alpha)$. Now define a linear map $\Phi\colon \mathfrak{sl}_2 \to L$ via

$$e \mapsto \sqrt{\frac{2}{\alpha(h_\alpha)}} e_\alpha, \qquad f \mapsto \sqrt{\frac{2}{\alpha(h_\alpha)}} e_{-\alpha}, \qquad h \mapsto \frac{2}{\alpha(h_\alpha)} h_\alpha.$$

This is clearly an isomorphism of vector spaces, and it is an isomorphism of Lie algebras, since

$$[\Phi(e), \Phi(f)] = \frac{2}{\alpha(h_\alpha)}[e_\alpha, e_{-\alpha}] = \Phi(h) = \Phi([e, f]),$$

$$[\Phi(h), \Phi(e)] = \left(\frac{2}{\alpha(h_\alpha)}\right)^{3/2} [h_\alpha, e_\alpha] = 2\sqrt{\frac{2}{\alpha(h_\alpha)}} e_\alpha = 2\Phi(e) = \Phi([h, e]),$$

$$[\Phi(h), \Phi(f)] = \left(\frac{2}{\alpha(h_\alpha)}\right)^{3/2} [h_\alpha, e_{-\alpha}] = -2\sqrt{\frac{2}{\alpha(h_\alpha)}} e_{-\alpha} = -2\Phi(f) = \Phi([h, f]),$$

completing the proof. □

Exercise 4.

Exercise. Suppose Φ is an irreducible reduced root system and that α, β and $\alpha + \beta$ are roots. If Φ is simply laced, then $(\alpha, \beta) < 0$. This does not hold in general if Φ is not simply laced.

Solution. We have the calculation

$$(\alpha + \beta, \alpha + \beta) = (\alpha, \alpha) + 2(\alpha, \beta) + (\beta + \beta).$$

If Φ is simply laced, then we actually have

$$(\alpha + \beta, \alpha + \beta) = (\alpha, \alpha) = (\beta, \beta) > 0,$$

and the claim follows at once.

As a counterexample when Φ is not simply laced, we consider the root system B_2. Using the names from the lecture, we find that α, $\beta + \alpha$ and $\beta + 2\alpha$ are all roots, but $\langle \alpha, \beta + \alpha \rangle = 0$. □

Exercise 5.

Exercise. Let L be a semisimple Lie alagebra with root system Φ with respect to a Cartan subalgebra H. Suppose Φ is not simply laced and let Φ' be the subset of roots of maximal length. Then $L_0 + \sum \alpha \in \Phi' L_\alpha$ is a Lie subalgebra of L.

Solution. Let $\alpha, \beta \in \Phi'$. Then

$$\frac{(\alpha + \beta, \alpha + \beta)}{(\alpha, \alpha)} = 2 + n(\alpha, \beta) \in \mathbb{Z}.$$

In particular, $\alpha + \beta$ cannot be strictly shorter than α, so if $\alpha + \beta$ is a root, then $\alpha + \beta \in \Phi'$ and the claim follows from 4.13(b). □

Exercise. There is a semisimple Lie algebra L and a Cartan subalgebra H of L such that the corresponding root system Φ is not simply laced and and if Φ' is the subset of roots of minimal length, then $L_0 + \sum_{\alpha \in \Phi'} L_\alpha$ is not a Cartan subalgebra.

Solution. Consider the Lie algebra $L = \mathfrak{sp}_4$. From a previous exercise, we know that L has a Cartan subalgebra generated by $D_1 = \operatorname{diag}(1, 0, -1, 0)$ and $D_2 = \operatorname{diag}(0, 1, 0, -1)$. We have also calculated the roots (warning: the naming is different from the naming in the list of reduced rank 2 root systems)

$$\begin{array}{lll} \alpha & D_1 \mapsto -1 & D_2 \mapsto -1, \\ \beta & D_1 \mapsto -1 & D_2 \mapsto 1, \\ \gamma & D_1 \mapsto 2 & D_2 \mapsto 0, \\ \delta & D_1 \mapsto 0 & D_2 \mapsto 2. \end{array}$$

We immediately see $\langle D_1, D_2\rangle_{\mathrm{ad}} = 0$, $\langle D_1, D_1\rangle_{\mathrm{ad}} = \langle D_2, D_2\rangle_{\mathrm{ad}} = 2$, and so the construction of the h_α for $\alpha \in \Phi$ yields

$$h_\alpha = -\frac{1}{2}(D_1 + D_2), \qquad h_\beta = \frac{1}{2}(D_2 - D_1),$$
$$h_\gamma = D_1, \qquad h_\delta = D_2.$$

This allows us to calculate $(\alpha, \alpha) = \langle h_\alpha, h_\alpha\rangle_{\mathrm{ad}} = 1$, $(\beta, \beta) = 1$, and $(\alpha+\beta, \alpha+\beta) = (-\gamma, -\gamma) = 2$. Thus, $\alpha, \beta \in \Phi'$, but $\alpha + \beta = -\gamma \notin \Phi'$. Furthermore, the Lie bracket of the generators of L_α and L_β is a nonzero element of $L_{-\gamma}$, so $L_0 + \sum_{\alpha\in\Phi'} L_\alpha$ is not a Lie subalgebra of $L = \mathfrak{sp}_4$. $\square$