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 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.

Elementary properties of Lie algebras

REMARK. Assume that 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 $JA^T J^{-1} + A = 0$ 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 $\rho_V \colon \mathfrak{L} \to \mathrm{End}\, V$, 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 $\rho_U \colon \mathfrak{L} \to \mathrm{End}\, U$ where $\rho_U(x)(u) := \rho_V(x)(u)$ 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 $\mathfrak{L}$ 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

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

We have (TODO: think about this more)

$$
\mathrm{ad}_{\mathbb{R}^3} \colon \mathbb{R}^3 \to \mathrm{End}\, \mathfrak{L} \cong M_3(\mathbb{R})
$$

$$
e_i \mapsto A_i \in \mathfrak{so}_3(\mathbb{R}) \subseteq \mathfrak{gl}_3
$$

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

EXAMPLE. We define a morphism

$$\rho \colon \mathfrak{sl}_2 \to \mathrm{Der}(k[X,Y]) \subseteq \mathrm{End}(k[X,Y])$$

$$e \mapsto X\frac{\partial}{\partial Y}$$

$$f \mapsto Y\frac{\partial}{\partial X}$$

$$h \mapsto X\frac{\partial}{\partial X} - Y\frac{\partial}{\partial Y}$$

An easy but somewhat lengthy calculation shows that this is a morhpism (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

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

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. □

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: $\mathfrak{L}^{(0)} \coloneqq \mathfrak{L}$, $\mathfrak{L}^{(n+1)} \coloneqq [\mathfrak{L}^{(n)}, \mathfrak{L}^{(n)}]$, 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 $[x + J, y + J] \coloneqq [x, y] + J$.

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. Exercise (TODO).   □

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: $\mathfrak{L}_{(1)} \coloneqq \mathfrak{L}$, $\mathfrak{L}_{(n+1)} \coloneqq [\mathfrak{L}_{(n)}, \mathfrak{L}]$. 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 $\alpha \coloneqq [z, [x, y]] \in \mathfrak{L}_{(i+j+1)}$ by definition. Furthermore, $[z, x]$ in $\mathfrak{L}_{(i+1)}$ by definition, so $\beta \coloneqq [y, [z, x]] \in \mathfrak{L}_{(i+j+1)}$ 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. $\square$

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. $\square$

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 $u_0 := u$ and inductively set $u_{i+1} := x(u_i)$. Define $U_i := \langle u_0, \ldots, u_i \rangle$. 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. $\square$

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 $V_1 := \langle v_1 \rangle$. Define $\overline{\mathfrak{L}} := \{ \overline{x} \mid x \in \mathfrak{L} \} \subseteq \operatorname{End}(V/V_1)$ 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 . $\square$

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

PROOF. We may assume that $\mathfrak{L} = \mathrm{End}\,V$. Let $\theta \colon \mathrm{End}\,V \to \mathrm{End}\,V$ 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 $V_i := \langle v_1, \dots, v_i \rangle$. 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 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

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

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) Trace forms of representations of invariant symmetric bilinear forms.
 (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. $\qquad\square$

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 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, $\mathrm{tr}(xy) = 0$ 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 $T = \{t \in \mathrm{End}\, V \mid [t, L] \subseteq L^{(1)}\}$. Note that $L^{(1)} \subseteq L \subseteq T$. $L^{(1)}$ is spanned by $[x, z]$, $x, z \in L$. Let $t \in T$. Then

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

where $[z, t] \in L^{(1)}$ by definition of T, hence $\mathrm{tr}([x, z] \circ t) = 0$. Thus, $\mathrm{tr}(wt) = 0$ 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. $\qquad\square$

THEOREM 3.5 (Cartan-Killing criterion for semisimplicity). Let 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

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

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

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

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

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 $\mathrm{ad}_{\mathfrak{L}} \colon \mathfrak{L} \to \mathrm{End}\,\mathfrak{L}$ and the image $\mathrm{ad}(J) \subseteq \mathrm{End}\,\mathfrak{L}$. By definition of J, we have $\mathrm{tr}(\mathrm{ad}(x) \circ \mathrm{ad}(y)) = 0$ for all $x \in J, y \in \mathfrak{L}$.

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

On the other hand, $\ker \mathrm{ad}_{\mathfrak{L}} = Z(\mathfrak{L})$ 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. $\qquad\square$

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 $\mathrm{Der}\,\mathfrak{L} \subseteq \mathrm{End}\,\mathfrak{L}$, and $\mathrm{ad}(\mathfrak{L})$ is a Lie ideal of $\mathrm{Der}\,\mathfrak{L}$.

---

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

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

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

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

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

so we conclude that

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

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 $\mathfrak{L} \to \mathrm{ad}_{\mathfrak{L}}$ is an abelian ideal, it must be zero (since it is trivially soluble), hence $\mathfrak{L} \cong \mathrm{ad}(L)$.

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

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

Thus if $D \in J$, then for all $x \in \mathfrak{L}$ we have $\mathrm{ad}(D(x)) = 0$ 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 $\mathrm{Der}(\mathfrak{L}) = \mathrm{ad}(\mathfrak{L})$ (by linear algebra) and so we are done. $\qquad \square$

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

DEFINITION 3.8. An endomorphism $x \in \mathrm{End}\, V$ 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 $V_i := \ker(x - \lambda_i \iota)^{m_i}$ 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 $x_s := p(x)$, $x_n := q(x)$.

For each i, we have

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

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. $\square$

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. $\square$

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 $T := \{t \in \operatorname{End} V \mid [t, B] \subseteq A\}$.

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 $y \in \mathrm{End}\, V$ via $y(v_i) := f(\lambda_i)v_i$. 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 $w_s = 0$. $\qquad\qquad\square$

PROPOSITION 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. $\square$

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 char $k \neq 0$. We will show that $\mathfrak{sl}_2$ is simple. We have met irreducible representations of $\mathfrak{sl}_2$. In particular $\mathfrak{sl}_2 \to \mathrm{End}(V_n)$, 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

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

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 polynomial q, r such that $1 = qt^m + rf$.

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

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

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 $\mathrm{Id}(A) = A$. $\square$

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 $y \in H$. $\qquad\square$

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). $\qquad\square$

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. $\qquad\square$

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. $\qquad\square$

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 $\theta, \phi \in \mathrm{End}\, V$ 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

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

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 $L_{0,z+c_j y} \subseteq H$. But H was chosen to be minimal among $L_{0,y}$ and so $L_{0,z+c_j y} = H$. Therefore, 0 is the only eigenvalue of the map

$$\operatorname{ad}(z + c_j y)|_H \colon H \to H.$$

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

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

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., $\operatorname{ad}(x)$ 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

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

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 $\operatorname{ad}(u) \operatorname{ad}(v)$ maps each generalized eigenspace into a different one, so $\operatorname{tr}(\operatorname{ad}(u) \circ \operatorname{ad}(v)) = 0$ (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 $\operatorname{ad}_L(H)$ is nilpotent and hence soluble, and so we can use Cartan's solubility criterion to find that

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

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. $\qquad\square$

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. $\qquad\square$

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

REMARK. In what follows are 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 $\langle L_\alpha, L_\beta \rangle_{\text{ad}} = 0$.

(c) If $\alpha \in \Phi$, then $-\alpha \in \Phi$.

(d) The restriction of $\langle \ , \ \rangle_{\text{ad}}$ 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 $\text{ad}(x)$ and $\text{ad}(y)$ are represented by diagonal matrices, and $\text{tr}(\text{ad}(x) \circ \text{ad}(y))$ is precisely the right hand side of the formula.

The proof of the first part of (b) is similar to the proof of 4.2(i) (TODO). By an argument similar to that used in 4.10(d) we have that $\text{tr}(\text{ad}(x) \circ \text{ad}(y)) = 0$ if $x \in L_\alpha$ and $y \in L_\beta$ and $\alpha + \beta \neq 0$ (TODO). So $\langle L_\alpha, L_\beta \rangle_{\text{ad}} = 0$.

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

Hence, $\langle L_\alpha, L \rangle_{\text{ad}} = 0$. But by Cartan-Killing $\langle \ , \ \rangle_{\text{ad}}$ 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 $\langle x, L_\beta \rangle_{\text{ad}} = 0$ 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, $\langle x, L \rangle_{\text{ad}} = 0$, 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 $\langle h, x \rangle_{\text{ad}} = \sum_{\alpha \in \Phi} m_\alpha \alpha(h) \alpha(x) = 0$. By (d), we have $h = 0$.

Thus, if $h \neq 0$, there is some $\alpha \in \Phi$ such that $\alpha(h) \neq 0$. TODO: why does this imply that Φ spans H^*? $\qquad\square$

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

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

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 $\text{ad}(U)(M) \subseteq M$.

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

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

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

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

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 $\mathrm{ad}(x)|_N : N \to N$. We have $0 = \mathrm{tr}\,\mathrm{ad}(x)|_N$ 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

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

But $\alpha(x) \neq 0$ by part (b) of the previous lemma. Hence $\sum rm_\alpha = 1$ 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). $\square$

REMARK. We have $A \cong \mathfrak{sl}_2$ and we saw that we had a representation $\mathrm{ad}\,|_N : A \to \mathrm{End}\,N$. 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 $r = (pq + 1)/2 \in \mathbb{Z}$ 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. $\square$

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) $\forall x \in H : \langle h_\alpha, x \rangle_{\mathrm{ad}} = \alpha(x)$,
        (b) $h_{\alpha \pm \beta} = h_\alpha \pm h_\beta$, $h_{-\alpha} = -h_\alpha$ and the h_α span H,
        (c) $h_\alpha = [e_\alpha, e_{-\alpha}]$, $\langle e_\alpha, e_{-\alpha} \rangle_{\mathrm{ad}} = 1$.
    (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 satisfies 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 sway 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 $r \leq s$. $\qquad\square$

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

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

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 trough β, so we are done.

$\square$

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 $\{h_1, \ldots, h_r\}$ forming a $\mathbb{C}$-basis.

LEMMA 4.21. The Killing form restricted to $\tilde{H}$ is an inner product and $h_1, \ldots, h_r$ 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 $.h_\alpha$ is a rational linear combination of $h_1, \ldots, h_r$. 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. $\square$

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 $\tilde{H}^*$.

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

$$\mathfrak{sl}_2 = L_0 \oplus L_\alpha \oplus L_{-\alpha},$$

$$L_\alpha = \{\begin{pmatrix} 0 & \lambda \\ 0 & 0 \end{pmatrix} \mid \lambda \in \mathbb{C}\},$$

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

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 $(\alpha, \beta) = |\alpha||\beta| \cos\varphi$, 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	$|\beta|$
0	0	$\pi/2$	
1	1	$\pi/3$	$|\alpha|$
-1	-1	$2\pi/3$	$|\alpha|$
1	2	$\pi/4$	$\sqrt{2}|\alpha|$
-1	-2	$3\pi/4$	$\sqrt{2}|\alpha|$
1	3	$\pi/6$	$\sqrt{3}|\alpha|$
-1	-3	$5\pi/6$	$\sqrt{3}|\alpha|$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

EXAMPLE. Table 5 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.

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

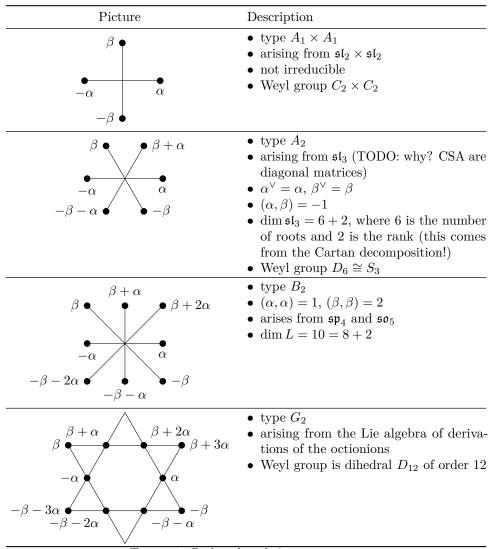The only irreducible simply laced root systems of rank 2 is A_2.

Picture	Description
	• type $A_1 \times A_1$ • arising from $\mathfrak{sl}_2 \times \mathfrak{sl}_2$ • not irreducible • Weyl group $C_2 \times C_2$
	• 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 $D_6 \cong S_3$
	• type B_2 • $(\alpha, \alpha) = 1$, $(\beta, \beta) = 2$ • arises from $\mathfrak{sp}_4$ and $\mathfrak{so}_5$ • $\dim L = 10 = 8 + 2$
	• type G_2 • arising from the Lie algebra of derivations of the octionions • Weyl group is dihedral D_{12} of order 12

TABLE 1. Reduced rank 2 root systems

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.

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

$$[v, x] = [\lambda_1 x + \lambda_2 y, x] = [\lambda_1 x, x] + [\lambda_2 y, x] = -\lambda_2 v,$$
$$[v, y] = [\lambda_1 x + \lambda_2 y, y] = [\lambda_1 x, y] + [\lambda_2 y, y] = \lambda_1 v.$$

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 $[(1, 0), (0, 1)] = (1, 0)$. □

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

$$\mathrm{ad}_L([x, y], z) = [[x, y], z]$$
$$= -[z, [x, y]],$$
$$[\mathrm{ad}_L(x), \mathrm{ad}_L(y)](z) = (\mathrm{ad}_\mathfrak{L}(x) \circ \mathrm{ad}_\mathfrak{L}(y) - \mathrm{ad}_\mathfrak{L}(y) \circ \mathrm{ad}_\mathfrak{L}(x))(z)$$
$$= \mathrm{ad}_\mathfrak{L}(x)(\mathrm{ad}_\mathfrak{L}(y)(z)) - \mathrm{ad}_\mathfrak{L}(y)(\mathrm{ad}_\mathfrak{L}(x)(z))$$
$$= [x, [y, z]] + [y, [z, x]]. \qquad \square$$

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. $\square$

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 $n + 1$. $\square$

EXERCISE. A finite-dimensional Lie algebra L is nilpotent if and only if $\mathrm{ad}(x)$ is nilpotent for all x in L.

SOLUTION. If L is nilpotent, then $\mathrm{ad}(x)$ is obviously nilpotent for all $x \in L$.

Conversely, if $\mathrm{ad}(x)$ is nilpotent for all $x \in L$, then $\mathrm{ad}(L) \subseteq \mathrm{End}\, L$ satisfies the condition of Engel, hence by (2.14) $\mathrm{ad}(L)$ is isomorphic to a subalgebra of $\mathfrak{n}_n$ for some n. In particular $\mathrm{ad}(L)$ is nilpotent. But we have $L/Z(L) \cong \mathrm{ad}(L)$, so by the previous lemma L is nilpotent. $\square$

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 $\mathrm{ad}(x)$ is nilpotent for every $x \in L$. Let $S \subseteq L$ be a proper Lie subalgebra. Define

$$\rho \colon S \to \mathrm{End}\, L/S$$
$$x \mapsto (y + S \mapsto [x, y] + S),$$

this is well-defined and a representation, because S is a subalgebra. We have that $\rho(S) \subseteq \mathrm{End}\, L/S$ 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 $y \in \mathrm{Id}(S) \setminus S$, 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 $\mathrm{Id}^0(S) := S$, $\mathrm{Id}^{n+1}(S) := \mathrm{Id}(\mathrm{Id}^n(S))$. We claim that if $y \in \mathrm{Id}^n(\langle x \rangle)$, then $\mathrm{ad}(x)^{n+1}(y) = 0$.

We will prove the claim by induction. If $y \in \mathrm{Id}^0(\langle x \rangle) = \langle x \rangle$, then $\mathrm{ad}(x)(y) = [x, y] = 0$. If the claim holds for $n \in \mathbb{N}_0$, let $y \in \mathrm{Id}^{n+1}(\langle x \rangle)$. By definition of the idealiser, we have that for any $z \in \mathrm{Id}^n(\langle x \rangle)$, $[y, z] \in \mathrm{Id}^n(\langle x \rangle)$. In particular, $x \in \mathrm{Id}^n(\langle x \rangle)$, so we find $[x, y] \in \mathrm{Id}^n(\langle x \rangle)$. Hence $\mathrm{ad}(x)^{n+2}(x)(y) = \mathrm{ad}^{n+1}(x)([x, y]) = 0$ by the inductive hypothesis, completing the proof.

Consider the sequence

$$\mathrm{Id}^0(\langle x \rangle) \subseteq \mathrm{Id}^1(\langle x \rangle) \subseteq \ldots.$$

By the idealiser condition and finite-dimensionality, we must have $\mathrm{Id}^n(\langle x \rangle) = L$ for some n. Then $\mathrm{ad}(x)^{n+1}(L) = 0$, so $\mathrm{ad}(x)^{n+1} = 0$, so $\mathrm{ad}(x)$ is nilpotent. By Exercise 9, we conclude that L is nilpotent. $\square$

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 $V = \langle v \rangle$. $\qquad\square$

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. $\qquad\square$

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. $\qquad\square$

Exercise 14.

EXERCISE. $\operatorname{ad}(L)$ is a Lie ideal of the Lie algebra of derivations $\operatorname{Der} L$ of the Lie algebra L.

SOLUTION. First of all, let $x, y, z \in L$. Then we have

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

so $\operatorname{ad}(x)$ is a derivation and we conclude that $\operatorname{ad}(L) \subseteq \operatorname{Der} L$. Since adjoints are obviously closed under addition and scalar multiplication, $\operatorname{ad}(L)$ is a subspace of $\operatorname{Der} L$.

Furthermore, let $D \in \operatorname{Der} L$ and $x, y \in L$. Then we have

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

so we conclude that $[D, \operatorname{ad}(x)] = \operatorname{ad}(D(x))$, hence $\operatorname{ad}(L)$ is a Lie ideal of $\operatorname{Der} L$. $\square$

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 $\operatorname{Der} L / \operatorname{ad}(L)$.

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 $\operatorname{ad}(x)$ sends y to z and other basis elements to 0, $\operatorname{ad}(y)$ sends x to $-z$ and other basis elements to 0 and $\operatorname{ad}(z)$ is the zero derivation. Hence $\operatorname{ad}(L)$ is a two-dimensional subalgebra of $\operatorname{Der} L$.

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 $\operatorname{Der} L$ consists of the endomorphisms that are precisely of the form above. In particular, $\operatorname{Der} L$ 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 $\operatorname{Der} L / \operatorname{ad}(L)$ is a 4-dimensional Lie algebra. We can give representatives $D, E, F, G \in \operatorname{Der} L$ 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 $\operatorname{Der} L / \operatorname{ad}(L) \cong L_2 \oplus L_2$. $\square$

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

$$\operatorname{ad}(\alpha x + \beta y)(x) = [\alpha x + \beta y, x] = -\beta y,$$
$$\operatorname{ad}(\alpha x + \beta y)(y) = [\alpha x + \beta y, y] = \alpha y.$$

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

$$\mu_1 x + \mu_2 y = D(y) = D([x, y]) = [D(x), y] + [x, D(y)]$$
$$= [\lambda_1 x + \lambda_2 y, y] + [x, \mu_1 x + \mu_2 y] = \lambda_1 y + \mu_2 y.$$

Hence $\mu_1 = \lambda_1 = 0$ and $D = \operatorname{ad}(\mu_2 x - \lambda_2 y)$, finishing the proof. $\square$