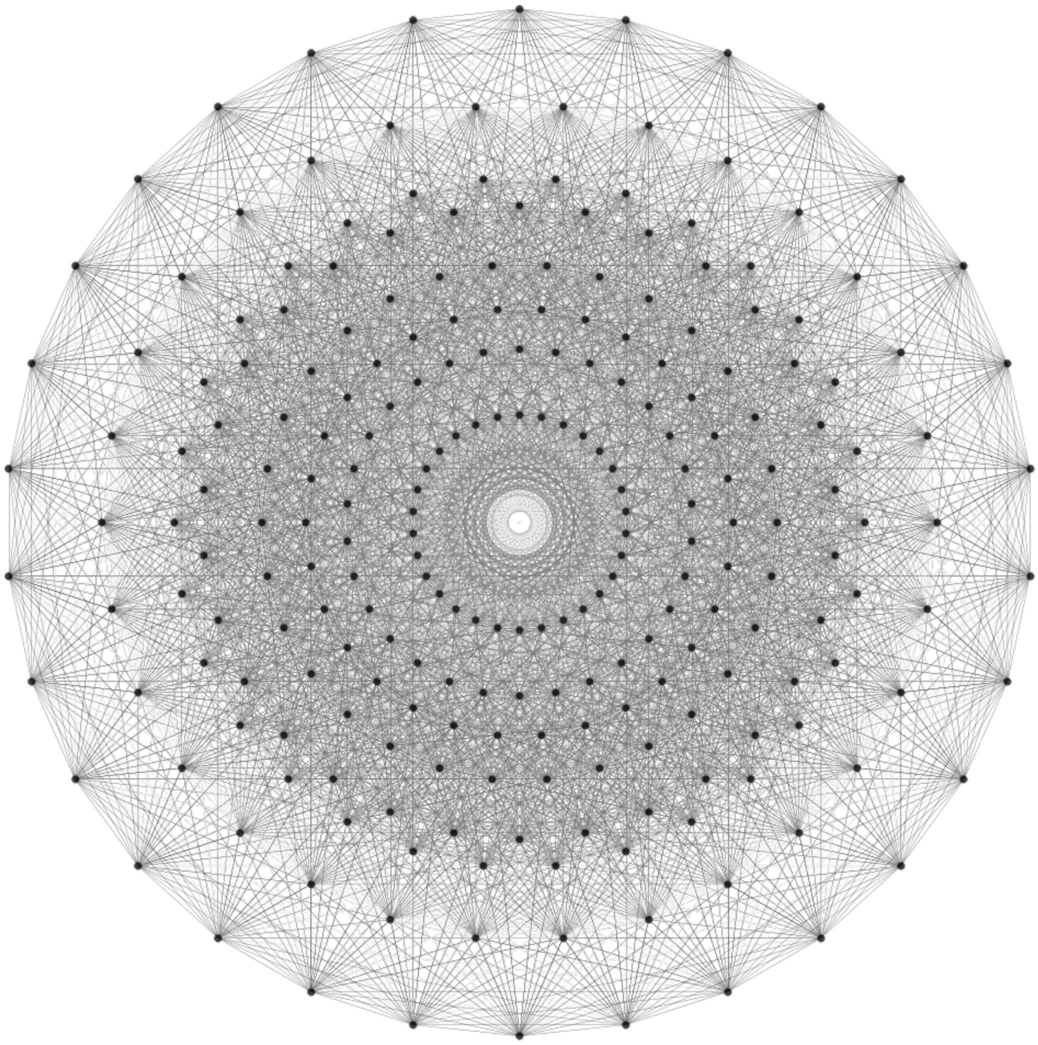
Lecture Notes to a course on Lie Algebras

Winter 2018, Technion IIT

Lectures by Amos Nevo

Typed by Elad Tzorani



Last updated December 31, 2018

Contents

Preface **iii**
Technicalities iii
Course Literature iii

I Lie Algebras 1

1 Preliminaries **2**
1.1 Basic definitions 2
1.2 Structure constants 4
1.3 Linear representations 5
1.4 Sub-algebras and ideals 6

2 Structure of Lie algebras **9**
2.1 Nilpotent Lie algebras 9
2.1.1 Flags 12
2.2 Solvable Lie algebras 13

3 Jordan-Chevalley decomposition **17**
3.1 The Chinese remainder theorem 17
3.2 Decomposition of vector spaces 17

4 Cartan's criterion for semi-simplicity **21**
4.1 Preliminary results 21
4.2 Cartan's criterion 22

5 Killing form **23**
5.1 Some properties of the Killing form 23

6 Derivations of simple algebras **27**

7 Toral subalgebras **29**

Preface

Technicalities

These aren't formal notes related to the course and henceforward there is *absolutely no guarantee* that the recorded material is in correspondence with the course expectations, or that these notes lack any mistakes.
In fact, there probably are mistakes in the notes! I would highly appreciate if any comments or corrections were sent to me via email at tzorani.elad@gmail.com.
Elad Tzorani.

Course Literature

The recommended course literature is as follows.

Humphreys, James E.: Introduction to Lie algebras and representation theory.

Jacobson, Nathan: Lie algebras. New York, 1962.

Part I

Lie Algebras

Chapter 1

Preliminaries

The course will be entirely algebraic, with possibly few examples from analysis.

This will allow us to discuss issues regarding the algebraic properties of Lie algebras. We might be interested in infinite-dimensional Lie algebras, but in this course we discuss only finite-dimensional algebras. In this course one of our main goals is a classification theorem for simple Lie algebras. We assume knowledge in linear algebras and specifically bilinear forms.

1.1 Basic definitions

Let $\mathbb{F}$ be a field, and V a finite-dimensional vector-space over $\mathbb{F}$.

Definition 1.1.1. V is a ***generalised*** $\mathbb{F}$***-algebra*** if it comes with a map $m\colon V \times V \to V$ which is bilinear.

$$\begin{aligned} m(v_1 + v_2, w) &= m(v_1, w) + m(v_2, w) \\ m(v, w_1 + w_2) &= m(v, w_1) + m(v, w_2) \\ m(av, bw) &= abm(v, w) \end{aligned}$$

Example. Let V be an associative algebra. Here m is an associative operation which is left and right distributive on addition in v. Equivalently: If we denote $m(v, w) = v \odot w$ then

$$\begin{aligned} (v \odot w) \odot u &= v \odot (w \odot u) \\ v \odot (u + w) &= v \odot u + v \odot w \\ (u + w) \odot u &= u \odot v + w \odot v \end{aligned}$$

Remark 1.1.2. Here associativity means the following.

$$m(v, m(w, u)) = m(m(v, w), u)$$

Examples.

1. Every field k is an $\mathbb{F}$–algebra over any subfield $\mathbb{F}$.
2. $M_n(k)$ is an $\mathbb{F}$–algebra.
3. P_n, polynomials over k of degree smaller or equal to n, is an $\mathbb{F}$-algebra.

Definition 1.1.3. A Lie algebra L over $\mathbb{F}$ is an $\mathbb{F}$-algebra, so $\exists m\colon L \times L \to L$, which generally need not be associative, but instead satisfies the following ***Jacobi identity***,

$$m(X, m(Y, Z)) + m(Z, m(X, Y)) + m(Y, m(Z, X)) = 0$$

and additionally, antisymmetry of the multiplication

$$m(X, Y) = -m(Y, X).$$

If $\mathrm{char}\mathbb{F} = 2$ we require $m(X, X) = 0$.

Notation 1.1.4. The "multiplication" in L is called ***bracket***, and denoted $m(X, Y) = [X, Y]$ (X bracket Y).

Remark 1.1.5. In these terms we write the Jacobi identity as follows.

$$[X,[Y,Z]]+[Z,[X,Y]]+[Y,[Z,X]]=0$$

Definition 1.1.6. A ***Lie algebra*** L is a vector space over $\mathbb{F}$ with a bilinear map $[,]: L\times L\to L$, which is anti-symmetric and satisfies the Jacobi identity.

Definition 1.1.7. Given a Lie algebra L, a vector subspace $L_0\subseteq L$ is called a ***Lie sub-algebra*** if it is closed under brackets. I.e.

$$X,Y\in L_0 \implies [X,Y]\in L_0.$$

Examples. 1. *Abelian Lie algebras:* The bracket is the zero form.

$$\forall X,Y\in L\colon\ [X,Y]=0$$

Example. $\mathbb{F}$ is itself a Lie algebra as well as any $\mathbb{F}$-vector space V under the bracket

$$\forall u,v\in V\colon\ [u,v]=0.$$

Example. Let A be any associative $\mathbb{F}$-algebra, and define on A *another* bilinear operation, namely

$$[a,b]=ab-ba.$$

This is called ***the commutator of*** a ***and*** b. Then $[,]: A\times A\to A$.

Exercise. This bracket satisfies the Jacobi identity, and is anti-symmetric.

Given a solution to this exercise, $(A,[,])$ is a Lie algebra.
In particular, $M_n(k)$ is a Lie algebra under the bracket $[A,B]=AB-BA$. This algebra is *very important* and is denoted $\mathfrak{gl}_n(k)$.

Exercise. Consider the subspace

$$\{A\in\mathfrak{gl}_n(k)\mid \mathrm{tr}A=0\}\subseteq\mathfrak{gl}_n(k).$$

Is the subspace a Lie algebra? Yes! Since for any $A,B\in\mathfrak{gl}_n(k)$ we have that $\mathrm{tr}(AB)=\mathrm{tr}(BA)$, we get that $\mathrm{tr}[A,B]=0$. The sub-Lie-algebra of zero-trace matrices is denoted $\mathfrak{sl}_n(k)$.

Exercise (Lie algebras associated with bilinear forms). Let V be a vector space over $\mathbb{F}$, and $B\colon V\times V\to\mathbb{F}$ be a bilinear form. Assume B is anti-symmetric. Define

$$L_B=\{X\in\mathrm{End}(V)\mid B(Xv,w)=-B(v,Xw)\}.$$

Check that L_B is a vector subspace of $\mathrm{End}(V)$. Consider the bracket operation on $\mathrm{End}(V)$, defined by $[T,S]=TS-ST$. Is L_B closed under brackets?

Solution. We compute as follows.

$$\begin{aligned}B([X,Y]v,w)&=B((XY-YX)v,w)\\&=B(XYv,w)-B(YXv,w)\\&=-B(Yv,Xw)+B(Xv,Yw)\\&=B(v,YXw)-B(v,XYw)\\&=B(v,(YX-XY)w)\\&=-B(v,[X,Y]w)\end{aligned}$$

In conclusion, L_B is a sub-Lie-algebra of $\mathrm{End}(V)$, the Lie algebra associated with the form B.

Exercise. Let S be a symmetric bilinear form, and let

$$L_S=\{X\in\mathrm{End}(V)\mid S(Xv,w)=-S(v,Xw)\}.$$

Then again, L_S is a Lie sub-algebra.

Examples (Sub-algebras of $\mathfrak{gl}_n(\mathbb{F})$). 1.

$$\mathfrak{t}(n,\mathbb{F}) = \left\{ \begin{pmatrix} a_{11} & & a_{i,j} \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix} \middle| \, a_{ij} \in \mathbb{F} \right\}$$

is closed under the bracket operation, for if $A, B \in \mathfrak{t}(n,\mathbb{F})$ then $AB \in \mathfrak{t}(n,\mathbb{F})$ and so $AB - BA \in \mathfrak{t}(n,\mathbb{F})$.

2.

$$\mathfrak{n}(n,\mathbb{F}) = \left\{ \begin{pmatrix} 0 & & a_{i,j} \\ & \ddots & \\ 0 & & 0 \end{pmatrix} \middle| \, a_{ij} \in \mathbb{F} \right\}$$

is a Lie sub-algebra of $\mathfrak{t}(n,\mathbb{F})$.

3.

$$\mathfrak{d}(n,\mathbb{F}) = \left\{ \begin{pmatrix} a_{1} & & 0 \\ & \ddots & \\ 0 & & a_{n} \end{pmatrix} \middle| \, a_{i} \in \mathbb{F} \right\}$$

an abelian sub-algebra.

1.2 Structure constants

Let L be a Lie algebra and let $X_1, \dots, X_n$ be a basis of L, Then the bracket operation is completely determined by the structure constants with respect to the basis.

$$[X_i, X_j] = \sum_{k=1}^{n} c_k^{i,j} X_k$$

The ***structure constants*** $c_k^{i,j}$ contain full information on the bracket operation of course. These satisfy two properties associated with anti-symmetry and the Jacobi identity of the brackets. The property associated to anti-symmetry is $c_k^{i,j} = -c_k^{j,i}$. The other property (associated to the Jacobi identity) is left as an **Exercise**.

Example.

$$\mathfrak{gl}_n(\mathbb{F}) = \operatorname{span}\{E_{i,j} \mid 1 \le i, j \le n\}$$

In the basis E_{ij} the structure constants are very simple. We have the following.

$$[E_{i,j}, E_{k,l}] = \delta_{j,k} E_{i,l} - \delta_{l,i} E_{k,j}$$

Hence all the structures constants are 1 or -1.

Definition 1.2.1. Let L_1, L_2 be Lie algebras. A ***Lie algebra homomorphism*** between L_1 and L_2 is a linear map $T\colon L_1 \to L_2$ satisfying

$$T[X,Y] = [TX, TY].$$

Definition 1.2.2. Let L be a Lie algebra. A sub-space $I \subseteq L$ is called a ***Lie-ideal*** of L if for all $X \in L$ and $Y \in I$, we have that $[X,Y] \in I$. This is written also by

$$[L, I] = \operatorname{span}\{[X,Y] \mid X \in L, Y \in I\} \subseteq I.$$

Definition 1.2.3. Let L be a Lie algebra and $L_0 \subseteq L$ be a sub-space. The ***Lie normaliser*** of L_0 is

$$N(L_0) = \{X \in L \mid [X, L_0] \subseteq L_0\}.$$

The ***Lie centraliser*** of L_0 is

$$Z(L_0) = \{X \in L \mid [X, L_0] = 0\}.$$

Definition 1.2.4. Let L be a Lie algebra. If $[X, Y] = 0$ one says that X and Y commute. We sometimes refer to the bracket as the commutator.

Example. Two sub-spaces $L_1, L_2 \subseteq L$ of a Lie algebra commute if their commutators are zero. I.e.

$$[L_1, L_2] = 0.$$

Remark 1.2.5. Although we have linearity of the bracket, we do need to take the span in the above example. If we take $X, X' \in L_1$ and $Y, Y' \in L_2$ we can't always express $[X, Y] + [X', Y']$ as a bracket of two elements, although it certainly is in the span.

1.3 Linear representations

Definition 1.3.1. A ***linear representation*** of a Lie algebra L over $\mathbb{F}$ is a Lie-algebra homomorphism $T\colon L \to \mathrm{End}(V) \cong \mathfrak{gl}_n(\mathbb{F})$ where V is an n-dimensional vector space over $\mathbb{F}$.

Remark 1.3.2. The bracket operation on $\mathrm{End}(V)$ is the usual one, namely $[A, B] = AB - BA$.

Let us define another large collection of Lie algebras. First, let A be a generalised $\mathbb{F}$–algebra, and denote $m(a, b) = a \odot b$.

Definition 1.3.3. A ***derivation*** of the generalised algebra A is a linear map $\delta\colon A \to A$ satisfying the following property.

$$\delta(a \odot b) = \delta(\alpha) \odot b + a \odot \delta(b)$$

Definition 1.3.4.

$$\mathrm{Der}(A) \coloneqq \{\delta \in \mathrm{End}(A) \mid \delta \text{ is a derivation.}\}$$

Remark 1.3.5. $\mathrm{Der}(A)$ is clearly a linear sub-space of $\mathrm{End}(A)$. Now, if δ_1 and δ_2 are derivations, $\delta_1 \circ \delta_2$ is *not* a derivation, usually. But, $[\delta_1, \delta_2] = \delta_1 \circ \delta_2 - \delta_2 \circ \delta_1$ *is* in fact a derivation.

Conclusion. $\mathrm{Der}(A)$, with the bracket inherited from $\mathrm{End}(A)$ is a Lie algebra.

Proof. We compute the following.

$$\begin{aligned}
[\delta_1, \delta_2](a \odot b) &= (\delta_1 \circ \delta_2 - \delta_2 \circ \delta_1)(a \odot b) \\
&= \delta_1 \circ \delta_2 (a \odot b) - \delta_2 \circ \delta_1 (a \odot b) \\
&= \delta_1 (\delta_2(a) \odot b + a \odot \delta_2(b)) - \delta_2(\delta_1(a) \odot b + a \odot \delta_1(b)) \\
&= \delta_1\delta_2(a) \odot b + \delta_2(a) \odot \delta_1(b) + \delta_1(a) \odot \delta_2(b) + a \odot \delta_1\delta_2(b) \\
&- (\delta_2\delta_1(a) \odot b + \delta_1(a) \odot \delta_2(b) + \delta_2(a) \odot \delta_1(b) + a \odot \delta_2\delta_1(b)) \\
&= (\delta_1\delta_2 - \delta_2\delta_1)(a) \odot b + a \odot (\delta_1\delta_2 - \delta_2\delta_1)(b)
\end{aligned}$$ ■

Example.

1. If A is an associative algebra, then $\mathrm{Der}(A)$ is a Lie algebra, $\mathrm{Der}(A) \subseteq \mathrm{End}(A)$. $\mathrm{Der}(A)$ is a sub-Lie-algebra of $\mathrm{End}(A)$ under bracket of linear transformations.
2. A Lie algebra is a generalised algebra and so $\mathrm{Der}(L)$ is another Lie algebra.

Fact 1.3.6 (important). *There is a very natural collection of derivations of any Lie algebras. For each $x \in L$, let us define a linear transformation denoted $\mathrm{ad}(x)\colon L \to L$ (this stands for "adjoint") via $\mathrm{ad}(x)(y) = [x, y]$. (This is linear from the bi-linearity of the bracket) In fact, $\mathrm{ad}(x)$ is a derivation of L. Namely,*

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

Indeed,

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

which is an identity as a consequence of the Jacobi identity.

Conclusion. The set $\{\mathrm{ad}(x) \mid x \in L\} \subseteq \mathrm{Der}(L)$ is a sub-algebra. We have the map $x \mapsto \mathrm{ad}(x)$ which is obviously linear (from bi-linearity of the bracket). So, $\mathrm{ad}(L) := \{\mathrm{ad}(x) \mid x \in L\}$ is a linear sub-space. In fact it is a Lie sub-algebra of $\mathrm{Der}(L)$.

Proof. We have to show that $[\mathrm{ad}(x), \mathrm{ad}(y)] = \mathrm{ad}(x)\mathrm{ad}(y) - \mathrm{ad}(y)\mathrm{ad}(x)$ is in the space $\mathrm{ad}(L)$. But, actually $[\mathrm{ad}(x), \mathrm{ad}(y)] = \mathrm{ad}\,[x, y]$, as the following proposition states.

Proposition 1.3.7. $\mathrm{ad}\colon L \to \mathrm{Der}(L)$ *is a Lie algebra homomorphism.*

Proof. Let us compute.

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

where the $\star$ is given from the Jacobi identity. ∎

In conclusion, $\mathrm{Der}(L)$ is a Lie sub-algebra of $\mathrm{End}(L)$ under bracket, and $\mathrm{ad}\colon L \to \mathrm{Der}(L) \subseteq \mathrm{End}(L)$ is a linear representation of the Lie algebra L with the image being $\mathrm{ad}(L) = \{\mathrm{ad}(x) \mid x \in L\}$. ∎

Example. Given $L_0 \subseteq L$ a sub-space. Then $N\,(L_0) = \{x \mid [x, L_0] \subseteq L_0\}$ is the set of elements x such that the linear transformation $\mathrm{ad}\,(x)$ leaves the subspace L_0 invariant. $N_L\,(L_0)$ is a Lie sub-algebra, and if L_0 is a Lie sub-algebra, then L_0 is an ideal of $N_L\,(L_0)$.

Example. The condition $[X, Y] = 0$ means $Y \in \ker\,(\mathrm{ad}(x))$ or equivalently $x \in \ker\,(\mathrm{ad}(y))$. Therefore

$$\begin{aligned}
Z\,(L_0) &= \{x \in L \mid [x, L_0] = 0\} \\
&= \{x \in L \mid L_0 \subseteq \ker\,(\mathrm{ad}(x))\}\,.
\end{aligned}$$

$Z\,(L_0)$ is a Lie sub-algebra of L, the Lie sub-algebra of elements commuting with every $x \in L_0$.

Remark 1.3.8. If $L_0 \subseteq L$ is a Lie sub-algebra, then $N\,(L_0)$ is the largest sub-algebra such that L_0 is is an ideal in it.

Remark 1.3.9. $Z_L(L)$ is the center of L, and an ideal.Indeed, if $z \in Z\,(L)$, and $x \in L$, then $\mathrm{ad}\,[x, z] = \mathrm{ad}\,x\,\mathrm{ad}\,z - \mathrm{ad}\,z - \mathrm{ad}\,x$ and $L \subseteq \ker\mathrm{ad}\,z$, so $L \subseteq \ker\mathrm{ad}\,[x, z]$, so $[x, z] \in Z\,(L)$ and $Z(L)$ is an ideal.

1.4 Sub-algebras and ideals

Remark 1.4.1.

1. If L_1 and L_2 are Lie sub-algebras, then $L_1 + L_2$ generally *is not*!
2. Suppose $I = L_1$ is an ideal and L_2 a sub-algebra. Then $I + L_2$ is a sub-algebra.
3. If $L_1 = I$ and $L_2 = J$ are ideals, then the Lie sub-algebra $I + J$ is an ideal. Indeed $[x, i] \in I$ and $[x, j] \in J$ for all j, so $[x, I + J] \subseteq I + J$.

Definition 1.4.2. The ***commutator*** of two sub-algebras L_1, L_2 is defined to be

$$\mathrm{Span}\,\{[X, Y] \mid X \in L_1, Y \in L_2\}\,.$$

Remark 1.4.3. The commutator of two sub-algebras *is not* in general a sub-algebra. Generally $[[X, Y]\,, [X', Y']]$ isn't in $[L_1, L_2]$ if $X, X' \in L_1$ and $Y, Y' \in L_2$. Let

$$\sum_{i=1}^{n} [X_i, Y_i] \in [L_1, L_2]$$

and

$$\sum_{j=1}^{m} \left[X'_j, Y'_j\right] \in [L_1, L_2]\,.$$

Then

$$\left[\sum_{i=1}^{n} [X_i, Y_i] \sum_{j=1}^{m} \left[X'_j, Y'_j\right]\right] = \sum_{\substack{i=1\\j=1}}^{\substack{n\\m}} \left[[X_i, Y_i]\,, \left[X'_j, Y'_j\right]\right]. \tag{1.1}$$

1. If $L_1 = I$ is an ideal, then $[I, L_2] \subseteq I$, is a sub-space of I.
2. If $L_1 = I$ and $L_2 = J$ are ideals, then $[I, J] \subseteq I \cap J$, and it is an ideal of L. Equation 1.1 shows that $[I, J]$ is indeed a sub-algebra. Now, let $[i, j] \in [I, J]$, and let $x \in L$. We should show that $[x, [i, j]] \in [I, J]$ which is sufficient for the span. Now

$$[x, [i, j]] \overset{\text{Jacobin identity}}{=} [[x, i], j] + [i, [x, j]] = [i', j] + [i, j'] \in [I, J]$$

as required.

Conclusion. $I + J$ and $[I, J]$ are ideals if I and J are.

Remark 1.4.4. In general $[I, J] \subseteq I \cap J$, but the inclusion may be strict.

Examples.
1. Take L an abelian Lie algebra and I, J any two sub-spaces which are both sub-algebras, and ideals. Then $[I, J] = 0$, but $I \cap J$ may be large.
2. Take L a Lie algebra of upper-triangular matrices, and $I = J$ the ideal of strict upper-triangular matrices. Then $[I, I]$ contains matrices that have zero entries in the diagonal above the main diagonal, hence $[I, I] \subsetneq I \cap J = I$.

Definition 1.4.5. If $[I, J] = 0$, we say that I and J ***commute***.

Remark 1.4.6. L is an ideal of itself, so $[L, L] = \operatorname{Span}\{[X, Y] \mid X, Y \in L\}$ is also an ideal, ***the commutator ideal of*** L.

Definition 1.4.7. L is ***abelian*** if $[L, L] = 0$.

Definition 1.4.8. L is ***perfect*** if $[L, L] = L$.

Definition 1.4.9. L is called a ***simple Lie-algebra*** if $\dim L > 1$ and L has no non-trivial ideals.

Exercise. A simple Lie algebra is in particular perfect.

Proposition 1.4.10. *If $\varphi \colon L \to L'$ is a Lie-algebra homomorphism, then $\ker \varphi$ is an ideal.*

Definition 1.4.11. For any ideal $I \lhd L$, the factor vector space $L/I = \{\ell + I \mid \ell \in L\}$ has a structure of a Lie algebra, given by the following.

$$[x + I, y + I]_{L/I} := [x, y]_L + I$$

Remark 1.4.12. The above is well defined since

$$[x + i, y + i'] = [x, y] + [i, y] + [x, i'] + [i, i'] \equiv [x, y] \pmod{I}.$$

The identities for Lie algebras follow immediately from those on L.

Theorem 1.4.13 (1st homomorphism theorem).

$$\begin{aligned} \pi \colon L &\to L/I \\ x &\mapsto x + I \end{aligned}$$

is a surjective Lie-algebra homomorphism, and $\ker \pi = I$.

Theorem 1.4.14 (2nd homomorphism theorem). *If I and J are ideals of L, and $I \subset J$, then the map*

$$\begin{aligned} \varphi \colon L/I &\to L/J \\ x + I &\mapsto x + J \end{aligned}$$

is a well-defined Lie-algebra epimorphism. We have from the first homomorphism theorem that

$$\left(L/I\right)/\ker \varphi \cong L/J$$

and

$$\ker \varphi = J/I$$

therefore

$$\left(L/I\right)/\left(J/I\right) \cong L/J.$$

Theorem 1.4.15 (3rd homomorphism theorem). *Given any two ideals I, J, their intersection $I \cap J$ is an ideal of L and we have a map*

$$\begin{aligned} \psi\colon I &\to {}^{I+J}\!/\!_{J} \\ i &\mapsto i + J \quad . \end{aligned}$$

This is a Lie-algebra homomorphism which is obviously surjective, with kernel $I \cap J$, hence

$$^{I}\!/\!_{I \cap J} \cong {}^{I+J}\!/\!_{J}$$

with Lie-algebra homomorphism induced by ψ.

Remark 1.4.16. If L_0 is an arbitrary Lie sub-algebra of L, and $J \triangleleft L$, then $J \cap L_0 \triangleleft L_0$ and $J \triangleleft L_0 + J$, and the Lie algebras ${}^{L_0+J}\!/\!_{J}$ and ${}^{L_0}\!/\!_{L_0 \cap J}$ are isomorphic under the canonical map ψ.

Chapter 2

Structure of Lie algebras

2.1 Nilpotent Lie algebras

Definition 2.1.1. The commutator ideal $[L, L]$ is denoted $L^{(1)}$. Similarly we denote $L^{(n)} = \left[L^{(n-1)}, L\right]$, which is an ideal of L.

Remark 2.1.2. The above gives a descending chain

$$L = L^{(0)} \supseteq L^{(1)} \supseteq L^{(2)} \supseteq \dots$$

and since $\dim L < \infty$, this sequence has to stabilise. It is however possible that $[L, L] = 0$, if L is abelian, or that $[L, L] = L$, if L is perfect.

Definition 2.1.3. If $L^{(n)} = 0$ for some n, L is called a ***nilpotent Lie algebra*** .If $L^{(n)} = 0$ and $L^{(n-1)} \neq 0$, we call $n - 1$ the ***index of nilpotency***.

Note 2.1.4. In some books n itself is called the index of nilpotency.

Definition 2.1.5. The sequence of ideals $L^{(n)}$ is called ***the descending central series*** of L.

Remark 2.1.6. $L^{(k)} \triangleleft L$ and hence $L^{(l)} \triangleleft L^{(k-1)}$. Also ${L^{(k-1)}}\big/{L^{(k)}}$ is an *abelian* algebra since $L^{(k)} = \left[L^{(k-1)}, L\right] \supseteq \left[L^{(k-1)}, L^{(k-1)}\right]$ and in general an ideal $I \triangleleft M$ is such that ${M}/{I}$ is abelian if and only if $I \supseteq [M, M]$.

Proposition 2.1.7. *Let $\varphi \colon L_1 \to L_2$ be an epimorphism of Lie algebras. Then $\varphi\left(L_1^{(n)}\right) = L_2^{(n)}$.*

Exercise. Prove the above proposition. For $n = 1$, we have $\varphi\left([L_1, L_1]\right) \subseteq [L_2, L_2]$, but in fact equality holds (**Exercise!**). Similarly prove for any $n \in \mathbb{N}$.

Proposition 2.1.8. *Let L be a nilpotent Lie algebra.*

1. *Every Lie sub-algebra and every factor Lie algebra are also nilpotent.*
2. *For M a Lie algebra, if ${M}\big/{Z(M)}$ is nilpotent, so is M.*
3. *$Z(L) \neq 0$.*

Proof. 1. **Sub-algebras:** If $L_0 \subset L$ is a Lie sub-algebra, then clearly $L_0^{(k)} \subseteq L^{(k)}$. So if $L^{(n)} = 0$ then $L_0^{(n)} = 0$ and the index of nilpotency of L_0 is bounded by that of L.

Factor algebras: Let $\bar{L} = \varphi(L) = {L}/{I}$ be an epimorphic image of L. Then $L^{(k)} = \varphi\left(L^{(k)}\right)$, so if $L^{(k)} = 0$, then $\bar{L}^{(k)} = 0$. We similarly have a bound on the nilpotency index of the factor algebra.

2. Suppose $\bar{L} = {L}/{Z}$ is nilpotent. Then $\bar{L}^{(n)} = \bar{0}$ for some n. So

$$\varphi\left(L^{(n)}\right) = \bar{L}^{(n)} = \bar{0}.$$

Then

$$\varphi\left(L^{(n)}\right) = \bar{L}^{(n)} = 0$$

and therefore $L^{(n)} \subseteq Z(L) = \ker \varphi$. Therefore $\left[L^{(n)}, L\right] \in [Z(L), L] = 0$. So $L^{(n+1)} = 0$, and so the index of nilpotency may increase by 1.

3. By definition,
$$L^{(0)} \supseteq L^{(1)} \supseteq \ldots \supseteq L^{(n-1)} \supsetneq L^{(n)} = 0$$
for some $n \in \mathbb{N}$. Now $\left[L^{(n-1)}, L\right] = L^{(n)} = 0$, so certainly $L^{(n-1)} \subseteq Z(L)$ and $Z(L) \neq 0$. ■

Exercise.
$$\mathfrak{n}(n, \mathbb{F}) = \left\{ \begin{pmatrix} 0 & & a_{i,j} \\ & \ddots & \\ 0 & & 0 \end{pmatrix} \middle| \, a_{ij} \in \mathbb{F} \right\}$$
is a nilpotent Lie sub-algebra of $M_n(\mathbb{F})$.

Example. In $\mathfrak{n}(2, \mathbb{F})$, the commutator of any two elements is zero.
$$\left[\begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & y \\ 0 & 0 \end{pmatrix} \right] = 0$$
Therefore $\mathfrak{n}(3, \mathbb{F})$ is a one-dimensional abelian algebra. For $\mathfrak{n}(3, \mathbb{F})$, the commutator of an element $\begin{pmatrix} 0 & 0 & x \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ with any other element is zero. However,
$$\left[\begin{pmatrix} 0 & u & w \\ 0 & 0 & v \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & x & y \\ 0 & 0 & z \\ 0 & 0 & 0 \end{pmatrix} \right] = \begin{pmatrix} 0 & 0 & uz - vx \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$
Then $[L, L] \neq 0$. However, $L^{(2)} = [[L, L], L] = 0$. Hence $\mathfrak{n}(3, \mathbb{F})$ is nilpotent of index 1.[1] Hence $L \big/ [L, L]$ is abelian of dimension 2, and $[L, L]$ abelian of dimension 1, and it is central (contained in the center). Being of dimension 1, we conclude $[L, L] = Z(L)$.
$\mathfrak{n}(3, \mathbb{F})$ is isomorphic to the ***first Heisenberg algebra*** denoted $\mathfrak{H}_1$.

Proposition 2.1.9. *For every $n \geq 2$ and field $\mathbb{F}$, $\mathfrak{n}(n, \mathbb{F})$ is a nilpotent Lie algebra of nilpotency index $n - 2$.*

Definition 2.1.10. An element $x \in L$ is called ***ad-nilpotent*** if $\operatorname{ad} x$ is a nilpotent linear transformation on L. Namely, $\exists k \in \mathbb{N}\colon (\operatorname{ad} x)^k = 0$.

Remark 2.1.11. In general, in a Lie algebra L which is nilpotent of index at most $n - 1$, $L^{(n)} = 0$, or equivalently
$$[[\ldots[[[x_1, x_2], x_3], x_4] \ldots, x_n], x_{n+1}] = 0$$
for all $x_1, \ldots, x_{n+1}$. Equivalently the product (in any order) of the linear transformations $\operatorname{ad} x_2, \ldots, \operatorname{ad} x_{n+1}$ is zero.

Theorem 2.1.12 (Engel). *Let L be a a Lie algebra such that every element of L is ad-nilpotent. Then L is a nilpotent Lie algebra.*

For the proof we shall develop some properties of nilpotent linear Lie algebras, namely Lie sub-algebras of $\operatorname{End}(V)$.

Proposition 2.1.13. *Let $X \in \operatorname{End}(V)$ be a nilpotent linear transformation on V. Then $\operatorname{ad}(X)$ is a nilpotent linear transformation on $\operatorname{End}(V)$, in particular $\operatorname{ad}(X) \in \operatorname{End}(\operatorname{End}(V))$.*

[1] Similarly once can show that $\mathfrak{n}(n, \mathbb{F})$ is nilpotent of index $n - 2$.

Proof. Define for each $X \in \mathrm{End}\,(V)$ two linear maps on $\mathrm{End}\,(V)$:

$$\lambda_X\,(Y) = XY$$
$$\rho_X\,(Y) = YX$$

Clearly if $X^k = 0$ then $\rho_X^k = \lambda_X^k = 0$. Furthermore, ρ_X commutes with λ_X (as linear maps on $\mathrm{End}\,(V)$). I.e. $[\lambda_X, \rho_X] = 0$. This is obvious because $(XY)\,X = X\,(YX)$. In general, in any associative algebra, (or any ring) the sum or the difference of two commuting nilpotent elements is also a nilpotent element.
We have

$$(\lambda_X - \rho_X)\,(Y) = XY - YX = [X, Y] = \mathrm{ad}\,(X)\,(Y)$$

so it suffices to prove the last claim, since this implies $\mathrm{ad}\,X$ is nilpotent.
By the binomial formula,

$$(a-b)^N = \sum_{j=0}^{N} \binom{N}{j} a^j\,(-b)^{N-j}\,.$$

If $a^k = b^k = 0$, then for large N s.t. $\min\{j, N-j\} \geq k$, the sum vanishes. ■

Remark 2.1.14. $\mathrm{ad}\,X\colon L \to L$ is a nilpotent linear transformation with index of nilpotency being $n-1$.

Remark 2.1.15. We saw that $X \in \mathrm{End}\,(V)$ is nilpotent, $\mathrm{ad}\,X \in \mathrm{End}\,(\mathrm{End}\,(V))$ is nilpotent. The converse *is not* true. For example take $X = I + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ which is *not* nilpotent, but is ad-nilpotent.

Theorem 2.1.16. *Let $L \subseteq \mathfrak{gl}_n\,(V)$ all of whose elements are nilpotent linear transformations. Then there exists $v \neq 0$ such that $\forall X \in L\colon Xv = 0$. Namely, a sub-algebra of $\mathfrak{gl}_n\,(V)$ consisting of nilpotent elements has a non-trivial joint kernel.*

Proof. Let us prove the theorem by induction on $\dim L$.

Induction Basis: The theorem is clearly true if $\dim L = 1$. Then $L = \mathbb{F}x$ and x is nilpotent, so there's $v \neq 0$ such that $xv = 0$.

Induction Step: (I) Assume the statement of the theorem for all linear Lie algebra of dimension less than $n \geq 2$. Let L have dimension $n \geq n$ and let $L_0 \subseteq L$ be a sub-algebra of strictly smaller dimension. (e.g. the span of a single matrix) Consider the linear maps $\mathrm{ad}\,x$ where $x \in L_0$. We have $L_0 \subseteq L \subseteq \mathrm{End}\,(V)$. Now $\mathrm{ad}\,x$ leaves both the linear sub-spaces L_0 and L invariant. In fact L is $\mathrm{ad}\,y$ invariant for any $y \in L$. (since L is closed under brackets) So $\mathrm{ad}\,x\,(L) \subseteq L$ and $\mathrm{ad}\,x\,(L_0) \subseteq L_0$. Therefore $\mathrm{ad}\,X$ also acts on ${}^{L}\!/_{L_0}$[2] via

$$\overline{\mathrm{ad}}x\,(y + L_0) = \mathrm{ad}\,x\,(y) + L_0.$$

Now,

$$\dim\{\mathrm{ad}\,x \mid x \in L_0\} \overset{\star}{\leq} \dim L_0 < L$$

where $\star$ is true because ad is linear on L_0, and cannot expand the dimension. But, $\overline{\mathrm{ad}}\,(L_0)$ is in fact a linear Lie algebra consisting of linear transformations of $\mathcal{U} := {}^{L}\!/_{L_0}$, because we saw that ad is in fact a Lie-algebra homomorphism. Now, each $\overline{\mathrm{ad}}x$ with $x \in L_0$, is a nilpotent linear transformation on the factor ${}^{L}\!/_{L_0}$, since x and hence $\mathrm{ad}\,x$ are nilpotent linear maps. Furthermore, $\dim \overline{\mathrm{ad}}\,(L_0) < \dim L$, so by the induction hypothesis, $\overline{\mathrm{ad}}\,(L_0)$ has a non-trivial vector in the joint kernel. I.e. $\exists y + L_0 \neq L_0$ such that $\overline{ad}\,(x)\,(y + L_0) = 0 + L_0$ for all $x \in L_0$. Namely, $[x, y] + L_0 = 0 + L_0$ for all $x \in L_0$, or equivalently $[x, y] \in L_0$ for all $x \in L_0$, so y normalises the sub-algebra L_0. So, $\mathrm{span}\,(L_0, y)$ is a Lie sub-algebra of L[3], containing L_0 strictly.[4]

(II) Let now L_0 be a sub-algebra of L such that $L_0 \subsetneq L$ and it's maximal with this property. Applying the previous argument to L_0, we have $N_L\,(L_0) \supsetneq L_0$ and therefore $N_L\,(L_0) = L$. So such an L_0 is an ideal of L.

(III) Consider ${}^{L}\!/_{L_0}$, which is a Lie algebra. If $x_0 + L_0 \neq L_0$ then $\mathbb{F}\,(x_0 + L_0)$ is a Lie sub-algebra of ${}^{L}\!/_{L_0}$. Its inverse image in L (under the canonical Lie-algebra homomorphism $L \to {}^{L}\!/_{L_0}$) is a Lie sub-algebra of L, containing L_0. But, having chosen L_0 to be maximal, and because $x_0 \notin L_0$, we have $\mathbb{F}x_0 + L_0 \supsetneq L_0$. So $L = \mathbb{F}x_0 + L_0$, namely L_0 is an ideal of co-dimension 1. So our sub-algebra L_0 which has $L_0 \subsetneq L$, and maximal with this property, turns out to be an ideal of co-dimension 1.

Remark 2.1.17. $N_1 := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $N_2 := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ are both nilpotent matrices, but $A := [N_1, N_2] = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ *is not.* These matrices are a linear basis of $\mathfrak{sl}_2 := \{z \in \mathfrak{gl}_2 \mid \mathrm{tr}\,z = 0\}$.

[2] This isn't necessarily a Lie algebra.
[3] We know that the normaliser is a sub-algebra.
[4] Since $y \notin L_0$.

(IV) Now consider the action of $\mathrm{ad}(L_0)$ on V. Now $\dim L_0 < \dim L$ and so by the induction hypothesis there's $v_0 \in V$ such that $v_0 \neq 0$ and $xv_0 = 0$ for all $x \in L_0$. We must find $v \neq 0$ such that $xv = 0$ for all $x \in L$. Let now

$$W = \{w \in V \mid \forall x \in L_0 \colon xw = 0\}$$

be the common kernel of non-zero elements in L_0. We claim that $W \subseteq V$ is invariant under the transformations in L. This finishes the proof, because it follows that x_0 leaves W invariant, and since x_0 is nilpotent, it must have a non-zero vector $v \in W$ such that $x_0 v = 0$. This v satisfies that $x_0 v = 0$ and $xv = 0$ for all $x \in L_0$, and therefore $xv = 0$ for all $x \in L$.

(V) We have to show that indeed W is invariant under L. Let $y \in L$ and let $w \in W$. We should show that $yw \in W$. So, we must show that for all $x \in L_0$, we have that $x(yw) = 0$. We shall prove this. We have

$$x(yw) = y(xw) + [x, y](w).$$

Now $xw = 0$ since $x \in L_0$ and $w \in W$, and $[x, y]w = 0$ since $[x, y] \in L_0$ and $w \in W$ (for L_0 is an ideal). Therefore $x(yw) = 0$ as required. ■

We remind Engel's theorem, for which we proved the above.

Theorem 2.1.18 (Engel). *Let L be a a Lie algebra such that every element of L is ad-nilpotent. Then L is a nilpotent Lie algebra.*

Proof. Consider $\mathrm{ad}\colon L \to \mathrm{End}(L)$. $\mathrm{ad}(L)$ is a linear Lie algebra consisting of linear maps on L. By assumption, $\mathrm{ad}(L)$ consists of nilpotent linear maps and by the previous theorem, there is $z \in L \setminus \{0\}$ which is the common kernel of all $\mathrm{ad}\, x$ with $x \in L$. Namely,

$$\exists z \neq 0 \forall x \in L\colon\ [x, z] = \mathrm{ad}\, x(z) = 0.$$

So, $z \in Z(L)$ and consider ${L}/{Z(L)} = \bar{L}$ which is a Lie algebra of dimension strictly less than $\dim L$. But, $\bar{L}$ is also ad-nilpotent since

$$\overline{ad}(x) : \bar{L} \to \bar{L}$$

is a transformation obtained from $\mathrm{ad}\, x$ by passing to a factor space. So by induction on the dimension, ${L}/{Z}$ is nilpotent, and so L is nilpotent since we saw that if ${L}/{Z}$ is nilpotent (Z the center) then L is nilpotent.
This proves that L is nilpotent.

2.1.1 Flags

Let V be a vector space over $\mathbb{F}$. A ***full flag*** in V is a sequence of linear subspaces

$$V_0 = 0 \subsetneq V_1 \subsetneq V_2 \subsetneq \ldots \subsetneq V_n = V$$

such that $\dim V_i = i$ for all $0 \leq i \leq n$.
A ***partial flag*** is any sequence

$$W_1 \subsetneq W_2 \subsetneq \ldots \subsetneq W_k$$

of nested subspaces. ■

Definition 2.1.19. Given a full flag, a linear transformation $T\colon V \to V$ is said to ***stabilise the flag*** if $TV_i \subseteq V_i$ for $0 \leq i \leq n$.

Let us choose $e_1 \in V_1$, $e_1, e_2 \in V_2$ a basis, etc. such that $e_1, \ldots, e_i$ is a basis of V_i. The matrix representing T in this basis is upper-triangular (because $TV_i \subseteq V_i$). Conversely, a linear transformation S represented in this basis by an upper triangular matrix, stabilises the flag.
Similarly, given for example a partial flag $V_0 \subsetneq W_1 \subsetneq W_2 \subseteq V$, with $\mathrm{rank}\, W_i = k_i$, we can choose a basis of W_1, complement it to a basis of W_2, then to a basis of V. A transformation U stabilises the partial flag if and only if it's represented by a block upper-triangular matrix with blocks of sizes $k_1, k_2 - k_1, n - k_2$. In our linear theorem we that every linear Lie algebra L consisting of nilpotent linear maps, has a non-zero vector in the common kernel. In follows that L stabilises a full flag, and in a basis adapted to this flag (as we chose before) *all* linear transformations in our algebra have a common upper-triangulation, with zeroes on the main diagonal. We want to show that indeed L stabilises a full flag.

Claim 2.1.20. L stabilises a full flag.

Proof. There's $v \in V$ non-zero such that for all $x \in L$, $xv = 0$. Let $V_1 = \mathbb{F}v$ and consider ${}^{V}\!/_{V_1}$. Now V_1 is invariant under *all* $x \in L_1$, so $xV_1 \subseteq V$, since $xv = 0$. So x defines a transformation $\bar{x}\colon {}^{V}\!/_{V_1} \to {}^{V}\!/_{V_1}$. This collection $\{\bar{x} \mid x \in L\}$ is a nilpotent linear Lie algebra. Therefore, $\bar{L}$ has a vector v_2 such that $x(v_2 + V_1) = 0 + V_1$, and where $v_2 + v_1 \neq 0 + V_1$. So, if $V_2 = \operatorname{span}\{v_1, v_2\}$ then $xV_2 \subseteq V_2$. Furthermore, $xV_1 = 0$ and $xv_2 \in V_1$. More generally, by induction, $\bar{L}$ stabilises a full flag in ${}^{V}\!/_{V_!}$, and its inverse image in V, together with V_1 is a full flag in V, which is invariant under all $x \in L$. Also, in the basis associated to this flag, the representing matrix has 0 on the diagonal. So every linear nilpotent Lie algebra stabilises a flag, with representing matrices as described.
Now, $\mathfrak{n}(n, \mathbb{F}) \subseteq \mathfrak{t}(n, \mathbb{F})$ is a nilpotent Lie algebra. ∎

Conclusion. Every linear nilpotent Lie algebra has a basis in which it is represented by a sub-algebra of $\mathfrak{n}(n, \mathbb{F})$.

Corollary 2.1.21. *Let L be a nilpotent algebra.* $\operatorname{ad} L$ *must have an invariant flag. This flag gives a sequence of ideals*

$$0 = I_0 \lneq I_1 \lneq I_2 \lneq \ldots \lneq I_n$$

where each I_j is an ideal of L and they have dimension $\dim I_j = j$.

2.2 Solvable Lie algebras

Definition 2.2.1 (Derived sequence of ideals). Let L be a Lie algebra. Denote $D_1(L) = L^{(1)} = [L, L]$ and similarly $D_k(L) = [D_{k-1}(L), D_{k-1}(L)]$ for all k. $(D_k)_{k \in \mathbb{N}_+}$ is the ***derived sequence of ideals for*** L.

Definition 2.2.2. L is ***solvable*** if $D_k(L) = 0$ for some k.

Remark 2.2.3. Every nilpotent Lie algebra is solvable. $L^{(k)} = 0$ implies $D_k(L) = 0$.

Definition 2.2.4. If $D_k(L) = 0$ and $D_{k-1}(L) \neq 0$ where $k \in \mathbb{N}_+$, we say L is ***solvable of index*** $k - 1$.

Example. The simplest solvable non-nilpotent algebra is the 2-dimensional algebra of 2×2 matrices generated by $X = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $Y = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Here $[X, Y] = XY - YX = Y$. We have $D_1(L) = [L, L] = \operatorname{Span}(Y)$. We have $D_2(L) = 0$ since $[L, L]$ is abelian. But, $[[L, L], L] = [L, L]$, so L is not nilpotent.

Example. For every n and $\mathbb{F}$, $\mathfrak{t}(n, \mathbb{F})$ is a solvable algebra. $D_1(\mathfrak{t}(n, \mathbb{F})) \subseteq \mathfrak{n}(n, \mathbb{F})$ and $\mathfrak{n}(n, \mathbb{F})$ is nilpotent, so $\mathfrak{t}(n, \mathbb{F})$ is solvable.

Proposition 2.2.5 (properties of solvable algebras).

1. *Every sub-algebra and quotient algebra of L is also solvable.*
2. *If I is an ideal in L and both I and ${}^{L}\!/_{I}$ are solvable, then L is solvable.*
3. *If I and J are solvable ideals, then $I + J$ is also solvable.*

Proof. First, if $\varphi\colon L \to L'$ is an epimorphism of Lie algebras, then $\varphi([L, L]) = [L', L']$, and in general

$$\varphi(D_k(L)) = D_k(\varphi(L)) = D_k(L').$$

1. Clearly, if $L_0 \subseteq L$ then $D_k(L_0) \subseteq D_k(L)$, so $D_k(L)$ implies $D_k(L_0)$, and L_0 is solvable of index at most that of L.
   Similarly, if $L' = {}^{L}\!/_{I}$ is a quotient algebra, and $\varphi\colon L \to {}^{L}\!/_{I}$ is the canonical epimorphism, then $D_k(L') = \varphi(D_k(L))$ and $D_k(L) = 0$ implies L' is solvable of index at most that of L.
2. Suppose that $\bar{L} = {}^{L}\!/_{I}$ is solvable. Then $D_k(\bar{L}) = \bar{0}$ and equivalently $D_k(L) \subseteq I$. Now, if I is a solvable ideal, then $D_l(I) = 0$ for some l. Then $D_{k+l}(L) \subseteq D_l(I) = 0$. So, L is solvable of index at most $l + k$.
3. If I, J are solvable ideals, consider ${}^{I+J}\!/_{J} \cong {}^{I}\!/_{I \cap J}$. Since I is solvable, so is ${}^{I}\!/_{I \cap J}$. So then ${}^{I+J}\!/_{J}$ is solvable. Since J is solvable, we get by (2). ∎

Proposition 2.2.6. *Every Lie algebra L has a unique maximal solvable ideal, containing all other solvable ideals.*

Proof. Let R be a solvable ideal, maximal with this property. If I is any solvable ideal, $R + I \supseteq R$ is solvable. Hence from maximality $R + I = R$, and hence $I \subseteq R$. ■

Remark 2.2.7. We obtained that R is the sum of all solvable ideals.

Definition 2.2.8. R is called the ***solvable radical*** of L, denoted $\mathfrak{R} = \mathrm{Rad}\,(L)$.

Question 2.2.9. We say that if ${}^{L}\!/_{I}$ and I are solvable, then L is solvable. IS it true that if ${}^{L}\!/_{I}$ and I are nilpotent then L is nilpotent?

The answer is no. Take $L = \left\{ X = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right\}$. Then $I = [L, L] \leq L$ is $\mathrm{Span}\,(Y)$. Both I and ${}^{L}\!/_{I}$ are abelian, hence nilpotent. However, L is not!

Definition 2.2.10. A Lie algebra L is called ***semi-simple*** if its radical vanishes.

Exercise. Let L be a Lie algebra. Prove that ${}^{L}\!\big/_{\mathrm{Rad}\,(L)}$ is semi-simple. Namely, $\mathrm{Rad}\left({}^{L}\!\big/_{\mathrm{Rad}\,(L)}\right)$.

Theorem 2.2.11 (Lie's theorem on solvable algebras). *Let $\mathbb{F}$ be an algebraically-closed field such that* $\mathrm{char}\,(\mathbb{F}) = 0$*. Let $L \subseteq \mathrm{End}\,(V)$ be a solvable Lie algebra. Let V be a vector space over F.*

1. *There's v non-zero which is a joint eigenvector of all $x \in L$.*
2. *L stabilises a full flag*
$$0 = V_0 \subseteq V_1 \subseteq V_2 \subseteq \ldots \subseteq V_n = V.$$
3. *In a basis adaptd to the flag, $e_1, \ldots, e_n$ such that* $\mathrm{Span}\,\{e_1, \ldots, e_i\} = V_i$, all *linear transformations $X \in L$ are represented by upper triangular matrices.*

Example. $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M_2\,(\mathbb{R})$ has eigenvalues $\pm i$. Take $L = \mathbb{R} \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, which is abelian and hence solvable. This has no upper-triangulisation because the eigenvalues are in $\mathbb{R}$.
The conclusion of Lie's theorem doesn't hold for Lie algebras over $\mathbb{R}$.

Proof. We prove the theorem by induction on $\dim\,(L)$. The theorem has a similar conclusion to Engel's theorem and the proof follows similar lines. It suffices to prove 1 as the rest follows by modding out eigenvectors.

Basis: If $\dim L = 1$, then $L = F \cdot X$ and the theorem holds since every linear transformation is conjugate to an upper-triangular one over an algebraically-closed field.

Step: Since L is solvable, $[L, L] \subsetneq L$, but ${}^{L}\!\big/_{[L, L]}$ is abelian, and in an abelian algebra, every sub-space is an ideal. Choose any subspace of codimension 1 in ${}^{L}\!\big/_{[L, L]}$ and take its pre-image. This is an *ideal* K of $K \leq L$ of codimension 1. By induction, since K is solvable, we have a joint eigenvector v of K. For $Y \in K$ denote $\lambda\,(Y)$ the eigenvalue of v under Y. Clearly

$$(\alpha Y_1 + \beta Y_2)\, v = \alpha Y_1 v + \beta Y_2 v = \alpha \lambda\,(Y_1)\, v + \beta \lambda\,(Y_2)\, v = \lambda\,(\alpha Y_1 + \beta Y_2)\, v$$

so $\lambda \colon K \to \mathbb{F}$ is a linear functional. Let us define the *lambda-**characteristic sub-space*** $W_\lambda \subseteq V$ by

$$W_\lambda = \{v \in V \mid \forall Y \in K \colon Yv = \lambda\,(Y)\, v\}$$

This is the sub-space consisting of all joint eigenvectors of K with joint eigenvalue λ.

Note 2.2.12. If we show that W_λ is L-invariant, the proof is complete, since $L = \mathbb{F}X + K$ for some X, and in particular, W_λ is X-invariant, and X has an eigenvector $u \in W_\lambda$ (since $\mathbb{F} = \bar{\mathbb{F}}$). So, u is a joint eigenvector of L.

Lemma 2.2.13. *W_λ is L-invariant.*

Proof. Write $L = K + \mathbb{F}X_0$ with some X_0.

I) We need to show that for $w \in W_\lambda$ and $X \in L$ we have $Xw \in W_\lambda$. So we need to show that $Y(Xw) = \lambda(Y) Xw$ for all $Y \in K$, by definition of W_λ. Now, $Y(Xw) = XYw - [X,Y]w$. Recalling that K is an ideal in L, we have $[X,Y] \in K$ for all $X \in L, Y \in K$. So

$$\begin{aligned} Y(Xw) &= Xyw - [X,Y]w \\ &= \lambda(Y) Xw - \lambda([X,Y]) w \end{aligned}$$

since $w \in W_\lambda$ and $[X,Y] \in K$. So we have to prove that

$$\forall X \in L \forall Y \in K \colon \lambda([X,Y]) = 0. \tag{2.1}$$

for all $X \in L$ and $Y \in K$.

II) To show (2.1), fix $X \in L$ and fix $w \in W$. Consider the sequence

$$w, Xw, X^2w, \ldots, X^{n-1}w, X^n w$$

where n is the least positive integers such that the sequence is linearly dependant. So, if we define

$$U_i = \operatorname{Span}\left\{w, Xw, \ldots, X^{i-1}w\right\}$$

then $\dim U_i = i$ for $1 \le i \le n$. Also $U_n = U_{n+1} = U_{n+2} = \ldots$.

III) **Claim 2.2.14.** Each U_i for $1 \le i \le n$ is invariant under K. Namely $YU_i \subseteq U_i$ for all $Y \in K$.

Proof. We prove this claim inductively. First Let's see that U_1 is K-invariant.

(i) U_1 is K-invariant for $yw = \lambda(Y) w$ for all $Y \in K$.

(ii) U_2 is K-invariant. Write $U_2 = \mathbb{F}w + \mathbb{F}$. We've seen

$$YXw = \lambda(Y) Xw - \lambda([X,Y]) w \in \mathbb{F}Xw + \mathbb{F}w.$$

So, K leaves U_2 invariant, but in fact we know more:

$$YXw \equiv \lambda(Y) Xw \pmod{U_1}$$

since $YXw = \lambda(Y) Xw + cw$. So

$$YXw - \lambda(Y) Xw \in U_1.$$

(iii) We claim that in general,

$$\forall 1 \le i \le n-1 \; \forall Y \in K \colon YX^iw \equiv \lambda(Y) X^i w \pmod{U_i}. \tag{2.2}$$

To see that, compute again.

$$\begin{aligned} YX^iw &= YX\left(X^{i-1}w\right) \\ &= XYX^{i-1}w - [X,Y] X^{i-1}w \end{aligned}$$

- By the induction hypothesis, $YX^{i-1}w = \lambda(Y) X^{i-1} + w'$ where $w' \in U_{i-1}$. So

$$X\left(YX^{i-1}w\right) = \lambda(Y) X^i + Xw'.$$

But by definition, $XU_{i-1} \subseteq U_i$. Hence

$$XYX^{i-1} = \lambda(Y) X^i + w''$$

where $w'' \in U_i$.

- The second summand

$$[X,Y] X^{i-1}w = \lambda([X,Y]) X^{i-1}w + w'''$$

where $w''' \in U_{i-1}$ by the induction hypothesis. This means

$$[X,Y] X^{i-1}w \in U_i + U_{i-1} \subseteq U_i.$$

The net conclusion is that

$$YX^iw = \lambda(Y) X^i w + w''''$$

with $w'''' \in U_i$. So $YX^iw \equiv \lambda(Y) X^i w \pmod{U_i}$ for all $1 \le i \le n$. ∎

IV) We have proved (2.2). Formulated otherwise is says that in the basis of $U_n = \{w, Xw, \ldots, X^{n-1}w\}$ given by the sequence, the representing matrix of *every* $Y \in K$ is upper triangular (that statement follows immediately from the fact the we proved $KU_i \subseteq U_i$) and in fact, the diagonal has only the entry $\lambda(Y)$. So, $\operatorname{tr} Y|_{U_n} = n\lambda(Y)$ for *every* $Y \in K$. In particular, this is true for elements $Y \in K$ which are of the form $[X, Y]$ with $Y \in K$. So $\operatorname{tr}[X,Y]|_{U_n} = n\lambda([X,Y])$.
We expect the trace of $[X,Y]$ to vanish, and that is true here since both X *and* Y preserve U_n. The fact that U_n is X-invariant is obvious, and we *saw* that U_n is also invariant under *every* $Y \in K$. So

$$[X,Y]|_{U_n} = \left[X|_{U_n}, Y|_{U_n}\right]$$

and it follows that

$$\operatorname{tr}[X,Y]|_{U_n} = 0 = n\lambda([X,Y]).$$

Now[5] $\lambda([X,Y]) = 0$ for all $Y \in K$ and $X \in L$. So we are done. ■

Remark 2.2.15. For every vector space V over a field $\mathbb{F}$,we can consider the spaces of flags over V.

Example. Consider the space of all lines in V. Namely

$$\mathrm{Gr}_1(V) \coloneqq \{\ell \subseteq V \mid \dim \ell = 1\}$$

(Grasmann 1, also known as the projective space over V). Similarly we can take

$$\mathrm{Gr}_k = \{W \subseteq V \mid \dim W = k\}$$

the Grasmann variety of k-vector-spaces in V. We can look more generally at any configuration

$$\mathrm{Gr}_{k_1,\ldots,k_m} \coloneqq \{\ell_1 \subsetneq \ell_n \subsetneq \ldots \subsetneq \ell_m \mid \dim \ell_i = k_i, \ell_i \subseteq V\}.$$

Preservence of this flag corresponds to an existence of a basis such that the matrices have a certain upper-block-triangular form.

Corollary 2.2.16. *Let L be a solvable algebra over $\mathbb{F}$, where $\mathbb{F} = \bar{\mathbb{F}}$ and $\mathrm{char}\mathbb{F} = 0$. There is a full flag of ideals in L, namely*

$$0 \subsetneq L_1 \subsetneq L_2 \subsetneq \ldots \subsetneq L_{n-1} \subsetneq L$$

where $n = \dim L$.

Proof. Consider $ad\colon L \to \mathrm{End}(L)$. $\mathrm{ad}(L)$ is a solvable Lie algebra, so it stabilises a full flag by Lie's theorem (2). The corresponding sub-spaces are ideals: they satisfy $\mathrm{ad}(L)(L_i) \subseteq L_i$ and $[L, L_i]\, L_i$, so $L_i \triangleleft L$. ■

Corollary 2.2.17. *If L is a solvable Lie algebra (with $\mathbb{F}$ as above) then the commutator ideal $[L,L]$ is nilpotent.*

Proof. Consider again the adjoint representation. We show that every element $X \in [L,L]$ is ad-nilpotent as a linear transformation on L.[6]
So, $\mathrm{ad}(L)$ is a linear Lie algebra, solvable, and has a basis in which all linear transformations in $\mathrm{ad}(L)$ are upper-triangular. But, the (usual Lie) commutator of two upper-triangular matrices is a nilpotent matrix (as a strictly upper-triangular matrix). Hence

$$[\mathrm{ad}(L), \mathrm{ad}(L)] \subseteq \{\text{upper triangular matrices with 0 on the diagonal}\}.$$

Since we have $[\mathrm{ad}(L), \mathrm{ad}(L)] = \mathrm{ad}[L,L]$, (since ad is a Lie-algebra homomorphism) so every $X \in [L,L]$ is ad-nilpotent. ■

[5] For that we use $\mathrm{char}(\mathbb{F}) = 0$ and the proof wouldn't work otherwise
[6] It suffices to show it is ad-nilpotent when acting on $[L,L]$.

Chapter 3

Jordan-Chevalley decomposition

3.1 The Chinese remainder theorem

Theorem 3.1.1 (Chinese remainder theorem). *Let R be a commutative unital ring, and let I, J be two ideals in R such that $I + J = R$. Then, given any $a, binR$ there exists $X \in R$ such that $X \equiv a \pmod{I}$ and $X \equiv b \pmod{J}$.*

Proof. Consider $\pi\colon R \twoheadrightarrow {}^{R}\!/_{I}$ the canonical homomorphism. Since $R = I + J$ clearly $\pi\colon I \to {}^{R}\!/_{J}$ is also surjective. So for all $a \in R$, $\pi\colon I + a \to {}^{R}\!/_{J}$ is also surjective. So there is $x \in I + a$ such that $\pi(x) = b + J$. So for any chosen $b \in R$ we have x such that $x \equiv a \pmod{I}$ and $x \equiv b \pmod{J}$. ∎

Theorem 3.1.2 (Chinese remainder theorem (more general)). *More generally, let $I_1, \ldots, I_n$ be ideals in R such that $I_i + \cap_{j \neq i} I_j = R$ for any $i \in [n]$. Then, given $a_1, \ldots, a_n$ arbitrary, there is $x \in R$ such that $x \equiv a_i \pmod{I_i}$ for all $i \in [n]$.*

Proof. By the Chinese remainder theorem[1], for each i we can choose x_i such that $x_i \equiv 1 \pmod{I_i}$ and $x_i \equiv 0 \pmod{I_j}$ for $j \neq i$. Finally $x = \sum_{i=1}^{n} x_i a_i$ satisfies $x \equiv a_i \pmod{I_i}$ for all $i \in [n]$. ∎

Example. Look in particular at the polynomial ring $\mathbb{F}[x]$. That is a Euclidean ring, hence a PID. So, every ideal $I \triangleleft \mathbb{F}[x]$ is of the form $p\mathbb{F}[x]$. What does it mean that $I + J = \mathbb{F}[x]$? It means that if $J = q\mathbb{F}[x]$, that $p\mathbb{F}[x] + q\mathbb{F}[x] = \mathbb{F}[x]$, so p and q are coprime. I.e. for some $u(x), v(x)$ we have $p(x)u(x) + q(x)v(x) = 1$. Conversely, if p, q are co-prime polynomials, then there are such $u(x)$ and $v(x)$ such that $p(x)u(x) + q(x)v(x) = 1$. So $p\mathbb{F}[x] + q\mathbb{F}[x] = \mathbb{F}[x]$.

Remark 3.1.3. If $p_1, \ldots, p_n$ are pairwise co-prime, then

$$\bigcap_{j \neq i} p_j \mathbb{F}[x] = \left(\prod_{j \neq i} p_j \right) \mathbb{F}[x].$$

Conclusion. The Chinese remainder theorem, applied to $\mathbb{F}[x]$, implies that given pairwise co-prime polynomials $p_1, \ldots, p_n$, and arbitrary $a_1, \ldots, a_n$, there is a polynomial p such that $p \equiv a_i \pmod{p_i \mathbb{F}[x]}$ for all $i \in [n]$.

3.2 Decomposition of vector spaces

Proposition 3.2.1. *Let T be a linear transformation on a vector space over $\mathbb{F}$ (arbitrary). Let f_T be the characteristic polynomial, and write $f_T = p_1 p_2$ where p_1, p_2 are co-prime. Then V decomposes to the direct sum of two T-invariant subspaces $V = V_1 \oplus V_2$, and more precisely $V_1 = \ker p_1(T)$ and $V_2 = \ker p_2(T)$.*

Proof. Start by writing $u_1 p_1 + u_2 p_2 = 1$ for some polynomials u_i. Consider the ring homomorphism $\mathbb{F}[x] \to \mathbb{F}[T]$ given by $x \mapsto T$ and deduce that

$$I = u_1(T) p_1(T) + u_2(T) p_2(T).$$

Writing that again for all $v \in V$, we get

$$v = u_1(T) p_1(T) v + u_2(T) p_2(T) v. \tag{3.1}$$

(i) First, $\ker p_1(T) \cap \ker p_2(T) = 0$, by (3.1).

[1] to which we shall henceforward sometimes refer to as CRT

(ii) $\ker p_1(T) + \ker p_2(T) = V$, since $u_1(T) p_1(T) v \in \ker p_2(T)$ and $u_2(T) p_2(T) v \in \ker p_1(T)$ as follows from $f_T = p_1 p_2$ and $f_T(T) = p_1(T) p_2(T) = 0$ by Cayley-Hamilton.

So every vector $v \in V$ is a sum of a vector in $\ker p_1(T)$ and a vector in $\ker p_2(T)$. $v \in \ker p_1(T)$ implies $p_1(T)(Tv) = 0$, so $Tv \in \ker p_1(T)$, and the kernel is an invariant sub-space. Similarly for $\ker p_2(T)$. ■

Proposition 3.2.2. *Let T be a linear transformation and assume its different eigenvalues $a_1, \ldots, a_n$ are all in $\mathbb{F}$. Write $f_T(x) = \prod_{i=1}^n (x - a_i)^{m_i}$. Then V decomposes to a direct sum $V = \bigoplus_{i=1}^n V_i$ of T-invariant sub-spaces where $V_i = \ker p_i(T)$ and $p_i(T) = (T - a_i)^{m_i}$.*

Proof. This follows immediately from the previous proposition, applied to $(x - a_i)^{m_i}$ and $\prod_{j \neq i} (x - a_j)^{m_j}$. ■

Theorem 3.2.3 (Jordan-Chevalley). *Let T be a linear transformation over $\mathbb{F}$ and assume that all of its eigenvalues are in $\mathbb{F}$. There exist two linear transformations T_s, T_n such that the following hold.*

(i) $T = T_s + T_n$

(ii) *T_n is nilpotent, and T_s is diagonalisable.*

(iii) *T_s and T_n commute.*

(iv) *T_s and T_n commute with T and with any other transformation that commutes with T.*

(v) *T_s and T_n are given as polynomials in T without constant terms.*

(vi) *If $A \subseteq B$ are two sub-spaces and $TB \subseteq A$, then T_s and T_n have the same property. $T_s B, T_n B \subseteq A$.*

(vii) *The first three properties determine the decomposition uniquely.*

Proof. (I) Write $f(x) = \prod_{i=1}^n (x - a_i)^{m_i}$ with $a_i \neq a_j$ for $i \neq j$. Then $V = \bigoplus_{i=1}^n V_i$ with $V_i = \ker p_i(T)$ where $p_i(T) = (T - a_i)^{m_i}$ as we saw. There exists a polynomial $p(x)$ such that $p(x) \equiv a_i \pmod{p_i}$ for $i \in [n]$ *and* $p(x) \equiv 0 \pmod{x}$. This follows from CRT as follows: If some $a_i = 0$, then the condition $p(x) \equiv 0 \pmod{x}$ is satisfied, and otherwise x is co-prime to each p_i, so that we can solve and find $p(x)$ as stated.

(II) Define $q(x) = x - p(x)$, so $p(x) + q(x) = x$, so $p(T) + q(T) = T$. Define $T_s = p(T)$ and $T_n = q(T)$. Clearly $T_s + T_n = T$, $T_s + T_n$ commute with T and with any other transformation that commute with T. In addition, p, q have no constant terms, by construction.

(III) We now restrict T_s and T_n to V_i, which is invariant under T, hence invariant under $p(T)$ and $q(T)$. Now $p(x) \equiv a_i \pmod{p_i}$. That is $p(x) - a_i = u_i(x)(x - a_i)^{m_i}$ for some polynomial $u_i(x)$. Here $p_i(x) = (x - a_i)^{m_i}$. But, $V_i = \ker p_i(T) = \ker (T - a_i)^{m_i}$, by definition. So obviously, for $v_i \in V_i$ we have

$$(p(T) - a_i) v_i = u_i(T) p_i(T) v_i = 0.$$

So $p(T)$ acts as the scalar a_i on V_i! So T_s is a diagonalisable transformation with the same eigenvalues as T, namely $a_1, \ldots, a_n$, each obtained $\dim V_i$ times.[2] We claim that the restriction of $q(T) = T_n$ to each V_i is nilpotent! Indeed, if $v_i \in V_i$, then

$$q(T) v_i = T_n v_i = (T - p(T)) v_i = Tv - a_i v_i = (T - a_i) v_i.$$

Since $v_i \in \ker p_i(T)$, it follows that

$$T_n^{m_i} v_i = (T - a_i)^{m_i} v_i = 0.$$

So, T_n is nilpotent in each V_i, hence nilpotent.
So, $T = T_s + T_n$ where T_s is diagonalisable, T_n is nilpotent, and they commute with each other and with every transformation commuting with T, and are given by polynomials in T without constant terms.

(IV) *Action on sub-spaces:* If $A \subseteq B$ and $TB \subseteq A$, then $T^2 B \subseteq TA \subseteq TB \subseteq A$, so it follows that any polynomial in T without constant terms satisfies $f(T)B \subseteq A$.

[2] The characteristic polynomial of $p(T)$ restricted to V_i is $p_i(x)$. This characteristic polynomial is co-prime to the characteristic polynomial of $p(T)$ on V_j. So, $f_{p_i(T)} \mid (x - a_i)^{m_i}$, but the product of all these (partial) characteristic polynomials (on the invariant subspaces V_i) is $f_{p(T)}$. This implies that $\dim V_i = m_i$.

(V) *Uniqueness:* Suppose that $T = Q_s + Q_n$ so that Q_s is diagonalisable, Q_n is nilpotent, and $Q_s = Q_n = Q_n = Q_s$. We show $Q_s = T_s$ and $Q_n = T_n$. But $T = T_s + T_n = Q_s + Q_n$. The fact that Q_s, Q_n commute implies that they commute with $T = Q_s + Q_n$. Hence, T_s and T_n commute with Q_s and Q_n, since they commute with every transformations commuting with T.[3] Consider $T_s - Q_s = Q_n - T_n$. T_n commutes with Q_n and the sum of two *commuting* nilpotent transformations is nilpotent by the binomial theorem. We now claim that $T_s - Q_s$ is diagonalisable, and then, since all of its eigenvalues are zero (as a nilpotent transformation) it must be the zero transformation, so $T_s = Q_s$ and $T_n = Q_n$. Indeed, T_s and Q_s are commuting diagonalisable transformations, so they have a common diagonalisation (when all the eigenvalues are in the field, because of diagonalisability). So, their sum or difference is also diagonalisable. ■

Exercise. Let $\mathcal{F} \subset \mathfrak{gl}(V)$ be any set of commuting diagonalisable matrices. Then there is a basis of common eigenvectors to all transformations in F. One can use induction on the dimension.

Remark 3.2.4. Write $f_T(x) = \prod_{i=1}^{n} (x - a_i)^{m_i} = \prod_{i=1}^{n} p_i(x)$. $T|_{V_i}$ has a characteristic polynomial $(x - a_i)^{m_i} = p_i(x)$. Indeed, $V_i = \ker p_i(T)$ by definition. So, $(T - a_i)^{m_i} = p_i(T)$ acts as 0 on V_i. So the characteristic polynomial of $T|_{V_i}$ has only a_i as a root. So it is $(x - a_i)^{k_i}$. But, this characteristic polynomial is co-prime to the characteristic polynomial of T on V_j when $j \neq i$. The product of all these partial characteristic polynomials on V_i is simply f_T. So, $m_i = k_i$ and the characteristic polynomial of $T|_{V_i}$ is $(x - a_i)^{m_i} = p_i(x)$.
Now, $p(T)$ leaves V_i invariant and acts on this m_i-dimensional space as a scalar. So, it has characteristic polynomial $(x - a_i)^{m_i}$.

Remark 3.2.5. $f_{T_s}(x) = f_T(x)$, but $m_{T_s}(x) = \prod_{i=1}^{n} (x - \lambda_i)$ where $\lambda_1, \ldots, \lambda_n$ are the distinct eigenvalues of T.

Definition 3.2.6. A linear transformation is called ***semi-simple*** if all the roots of its minimal polynomial have multiplicity 1.

Fact 3.2.7. *If the roots of $f_T(x)$ are in $\mathbb{F}$, then T is semi-simple if and only if it's diagonalisable.*

Proposition 3.2.8. *Let $T \colon V \to V$ be linear.*

1. *If S is diagonalisable, then so is* $\operatorname{ad} S \colon \operatorname{End} V \to \operatorname{End} V$.
2. *If S is nilpotent, then* $\operatorname{ad} S$ *is nilpotent on* $\operatorname{End} V$.
3. *If $T = T_s + T_n$ is a Jordan-Chevalley decomposition for T, then $\operatorname{ad} T = \operatorname{ad} T_s + \operatorname{ad} T_n$ is the Jordan-Chevalley decomposition of* $\operatorname{ad} T$.

Proof. 1. Let T be diagonalisable over $\mathbb{F}$ with eigenvectors $v_1, \ldots, v_n$ and eigenvalues $\lambda_1, \ldots, \lambda_n$.

We first show that if $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ is diagonal, then $\operatorname{ad} D$ is diagonalisable on $M_n(\mathbb{F})$. In fact, let $\mathfrak{a}$ be the algebra of diagonal matrices, and let $E_{i,j}$ be the matrices where $(E_{i,j})_{k,\ell} = \delta_{i,k}\delta_{j,\ell}$. Then $\mathfrak{a} \subseteq \ker \operatorname{ad} D$ since $\mathfrak{a}$ is a commutative Lie algebra, and $D \in \mathfrak{a}$. Each $E_{i,j}$ is an eigenvector of $\operatorname{ad} D$ and $\operatorname{ad} D(E_{i,j}) = DE_{i,j} - E_{i,j}D = (\lambda_i - \lambda_j) E_{i,j}$. So, $E_{i,i}$ and $E_{i,j}$ form a basis of eigenvectors of $\operatorname{ad} D$.[4]
Let T be now a general linear map. Write $D = PTP^{-1}$ and then $\operatorname{ad}(PTP^{-1}) E_{i,j} = (\lambda_i - \lambda_j) E_{i,j}$. We write $T = P^{-1}DP$ and so

$$\begin{aligned}
\operatorname{ad} T\left(P^{-1}E_{i,j}P\right) &= TP^{-1}E_{i,j}P - P^{-1}E_{i,j}PT \\
&= P^{-1}DE_{i,j}P - P^{-1}E_{i,j}DP \\
&= P^{-1}\left(DE_{i,j} - E_{i,j}\right)P \\
&= (\lambda_i - \lambda_j) P^{-1}E_{i,j}P
\end{aligned}$$

therefore $P^{-1}E_{i,j}P$ is an eigenvector of $\operatorname{ad} T$ with eigenvalue $\lambda_i - \lambda_j$.

[3] We constructed T_n and T_s, and they satisfy all the properties, Q_n and Q_s are currently more general.
[4] If the eigenvalues are distinct, we obtain that the kernel is generated by the $E_{i,i}$.

2. Consider $\operatorname{ad} S\,(X) = SX - XS = \lambda_S(X) - \rho_S(X)$ where $\lambda_S(X) = SX$ and $\rho_S(X) = XS$ and λ_S, ρ_S are two *commuting* nilpotent transformations on $\operatorname{End} V$ and so $\lambda_S - \rho_S$ is also nilpotent by the binomial theorem.

3. By our characterisation, $\operatorname{ad} T_s$ is diagonalisable, and T_n is nilpotent. So, $\operatorname{ad}[T_s, T_n] = 0 = [\operatorname{ad} T_s, \operatorname{ad} T_n]$. So $\operatorname{ad} T_s, \operatorname{ad} T_n$ commute, so they are the Jordan-Chevalley decomposition. ■

Chapter 4

Cartan's criterion for semi-simplicity

4.1 Preliminary results

Proposition 4.1.1. *Let $U \subseteq W \subseteq \operatorname{End} V$ be linear subspaces. Define*

$$M = \{X \in \operatorname{End} V \mid [X, W] \subseteq U\} = \{X \in \operatorname{End} V \mid \operatorname{ad} X (W) \subseteq U\}.$$

Assume that $\operatorname{char}\mathbb{F} = 0$ *and* $\mathbb{F} = \bar{\mathbb{F}}$.
Let $X \in M$, if $\operatorname{tr} XY = 0$ for all $Y \in M$ then X is nilpotent.

Proof. $X = X_s + X_n$, so we show that $X_s = 0$. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of X_s. Consider the $\mathbb{Q}$-vector space defined by the linear span of the eigenvalues.

$$E \coloneqq \operatorname{span}_{\mathbb{Q}} \{\lambda_1, \dots, \lambda_n\}$$

Then $\dim_{\mathbb{Q}} E < \infty$. We want to show that $\dim_{\mathbb{Q}} E = 0$ and then $\lambda_i = 0$ for all i, and $T_s = 0$.
To show that $E = 0$, it is enough to show that $E^* = \operatorname{Hom}_{\mathbb{Q}}(E, \mathbb{Q}) = 0$. Let $f\colon E \to \mathbb{Q}$ be a $\mathbb{Q}$-linear functional, we want to show that $f(\lambda_i) = 0$ for all $i \in [n]$.
To do that, let Y be the linear transformation such that in the basis B of eigenvectors of X_s, it (Y) is with values $f(\lambda_i)$ on the diagonal. I.e. $Y v_i = f(\lambda_i) v_i$ for all $i \in [n]$. So, the eigenvalues of $\operatorname{ad} Y$ are $f(\lambda_i) - f(\lambda_j)$ for $i, j \in [n]$ as we saw. The eigenvalues of $\operatorname{ad} X_s$ are $\lambda_i - \lambda_j$ and there exists a polynomial p with no constant term such that $p(\lambda_i - \lambda_j) = f(\lambda_i) - f(\lambda_j)$.[1] Note that because f is linear, where $\lambda_i - \lambda_j = \lambda_k - \lambda_\ell$ then $f(\lambda_i) - f(\lambda_j) = f(\lambda_k) - f(\lambda_\ell)$. We can arrange p to have no constant term. If $\lambda_i - \lambda_j = 0$ then $f(\lambda_i) - f(\lambda_j) = 0$ so $p(0) = 0$. Otherwise, 0 is distinct from $\lambda_i - \lambda_j$ for $i \neq j$. So we can add it in. Given p, consider $p(\operatorname{ad} X_s)$. It is diagonalisable since $\operatorname{ad} X_s$ is, and its eigenvalues coincide with those of $\operatorname{ad} Y$ with the same multiplicities. So, $p(\operatorname{ad} X_s) = \operatorname{ad} Y$. Now, $\operatorname{ad} X_s$ is a polynomial *without* constant term in $\operatorname{ad} X$[2].
We conclude that $\operatorname{ad} Y$ is a polynomial without constant term in $\operatorname{ad} X$. Recall, that our assumptions are that

$$M = \{X \in \operatorname{End} V \mid [X, W] \subseteq U\} = \{X \in \operatorname{End} V \mid \operatorname{ad} X (W) \subseteq U\}.$$

But, we chose $X \in M$, so we conclude that Y also satisfies that $\operatorname{ad} Y (W) \subseteq U$[3]. So $Y \in M$. Our basic assumption was that $\operatorname{tr} XY = 0$ for all $Y \in M$.
We claim that $\operatorname{tr} XY = \sum_{i \in [n]} \lambda_i f(\lambda_i)$. Let's compute in the basis B of eigenvectors of X_s (*and also* Y).

$$XY v_i = f(\lambda_i) X v_i = f(\lambda_i)(X_s + X_n) v_i = \lambda_i f(\lambda_i) v_i + f(\lambda_i) X_b v_i.$$

X_n is nilpotent, and so the claim follows.
So we conclude that $\sum_{i=1}^{n} \lambda_i f(\lambda_i) = \operatorname{tr} XY = 0$. Now, $f(\lambda_i)$ are rational numbers because $f\colon E \to \mathbb{Q}$ is a $\mathbb{Q}$-linear functional we obtain

$$\begin{aligned} f\left(\sum_{i \in [n]} \lambda_i f(\lambda_i)\right) &= \sum_{i \in [n]} f(\lambda_i f(\lambda_i)) \\ &= \sum_{i \in [n]} f(\lambda_i) f(\lambda_i) \\ &= \sum_{i \in [n]} f(\lambda_i)^2 \\ &= 0. \end{aligned}$$

[1] via Lagrange's polynomial of interpolation
[2] since $\operatorname{ad} X_s$ is the semi-simple part of $\operatorname{ad} X$, and the semi-simple part is a polynomial in the transformation
[3] as a polynomial without constant term in $\operatorname{ad} X$, which has this property

Finally $f(\lambda_i) = 0$ for all $i \in [n]$. Then $f = 0$ so $E^* = \{0\}$ so $E = \{0\}$ so $X_s = 0$ so $X = X_n$ is nilpotent. ■

Theorem 4.1.2 (Lagrange's polynomial of interpolation). *Let $a_1, \ldots, a_m$ be distinct in any field $\mathbb{F}$ and let $b_1, \ldots, b_m \in \mathbb{F}$. There is a unique polynomial $p(x) \in \mathbb{F}[x]$ such that $p(a_i) = b_i$ for all $i \in [m]$. p is unique among polynomials of degree at most $m-1$.*

Proposition 4.1.3. *Let $L \subseteq \operatorname{End} V$ be a Lie algebra. Assume $\mathbb{F} = \bar{\mathbb{F}}$ and $\operatorname{char}\mathbb{F} = 0$. If $\operatorname{tr} XY = 0$ for every $X \in [L, L]$ and every $Y \in L$ then L is a solvable Lie algebra.*

Proof. We use the previous proposition. First, for all $Y \in L$, $\operatorname{ad}(Y) L \subseteq [L, L]$ by definition. Consider $[L, L] \subseteq \operatorname{End} V$, $U = [L, L]$, $W = L$ and let $M = \{Z \in \operatorname{End} V \mid [Z, L] \subseteq [L, L]\}$. Then $L \subseteq M$. We assume $\operatorname{tr} XY = 0$ for all $Y \in L$, but to use the previous proposition, we need to show that $\operatorname{tr} XZ = 0$ for all $Z \in M$. If we show that, then by the previous proposition, X is nilpotent, so $[L, L]$ consists only of nilpotent linear transformations, and by the result preceding Engel's theorem[4], $[L, L]$ is nilpotent and so L is solvable.
To show that $\operatorname{tr} XZ = 0$ write $X = [U, V]$ since $X \in [L, L]$. Now,

$$\begin{aligned}
\operatorname{tr}([U, V], Z) &= \operatorname{tr}((UV - VU) Z) \\
&= \operatorname{tr}(UVZ - VUZ) \\
&= \operatorname{tr}(UVZ) - \operatorname{tr}(V(UZ)) \\
&= \operatorname{tr}(UVZ) - \operatorname{tr}(UZV) \\
&= \operatorname{tr}(U(VZ - ZV)) \\
&= \operatorname{tr}(U[V, Z])
\end{aligned}$$

and $Z \in M$ so $[V, Z] = [Z, V] \in [L, L]$. Also, $U \in L$, hence

$$\operatorname{tr}(U[V, Z]) = \ldots = \operatorname{tr}(U[V, Z]) = \operatorname{tr}([V, Z] U) = 0$$

by assumption. So, every $X \in [L, L]$ is nilpotent. Then $[L, L]$ is nilpotent, and then L is solvable. ■

Remark 4.1.4. We just saw that for a linear Lie algebra, $L \subseteq \operatorname{End} V$, $\operatorname{tr}([L, L] L) = \{0\}$ implies L is solvable.

4.2 Cartan's criterion

Theorem 4.2.1. *Let L be any Lie algebra over a field $\mathbb{F}$ as above. If $\operatorname{tr} \operatorname{ad} x \operatorname{ad} y = 0$ for all $x \in [L, L]$ and $y \in L$, then L is solvable.*

Proof. Consider $\operatorname{ad} L \subseteq \operatorname{End} L$ and $\operatorname{ad}[L, L] = [\operatorname{ad} L, \operatorname{ad} L]$, and so $\operatorname{ad} L$ is solvable by the proposition for the linear case. So $\operatorname{ad}(L) = L/\ker\operatorname{ad} = L/Z(L)$ is solvable, and so L is solvable.[5] ■

[4] A Lie algebra of nilpotent transformations is nilpotent

[5] As an exercise, if $L/ZL)$ is solvable, so is L.

Chapter 5

Killing form

Let L be a Lie algebra and define a symmetric bilinear form

$$B_L(X,Y) := \operatorname{tr}(\operatorname{ad} X \operatorname{ad} Y).$$

This is called the ***Killing form***.

Proposition 5.0.1 (invariance of the Killing form).

$$B([x,y],z) = B(x,[y,z])$$

Proof. For any three linear transformations R, T, S ($\operatorname{ad} X$, $\operatorname{ad} Y$, $\operatorname{ad} Z$, respectively) we shall compute $\operatorname{tr}([T,S]R)$ and $\operatorname{tr}(T[S,R])$ and show equality.

$$\begin{aligned}\operatorname{tr}([T,S]R) &= \operatorname{tr}((TS-ST)R)\\ &= \operatorname{tr}(TSR) - \operatorname{tr}(RST)\\ &= \operatorname{tr}(TSR) - \operatorname{tr}(TRS)\\ &= \operatorname{tr}(T(SR-RS)) = \operatorname{tr}(T[S,R])\end{aligned}$$

hence there's equality. ■

Proposition 5.0.2 (more invariance of the Killing form).

$$B(\operatorname{ad} Y(X), Z) + B(X \operatorname{ad} Y(Z)) = 0$$

Proof. $B([X,Y],Z) = B(-\operatorname{ad} Y(X), Z)$ and use the previous proposition. ■

More generally, let $\pi\colon L \to \operatorname{End} V$ be *any* linear representation of L (namely, a Lie-algebra homomorphism into $\operatorname{End} V$). Define

$$B_\pi(x,y) = \operatorname{tr}(\pi(x)\pi(y)).$$

The proves of the above propositions stay the same, therefore

$$B_\pi([x,y],z) = B_\pi(x,[y,z]).$$

B_π is a symmetric bilinear form satisfying the symmetry condition.

Conclusion. The Cartan criterion can be formulated to say that $B_\pi(L,[L,L]) = 0$, thus L is solvable.

5.1 Some properties of the Killing form

Proposition 5.1.1. *Let $I \lhd L$ be an ideal of L. Then the Killing form K_I of I, as a Lie algebra on its own is $K_L|_{I\times I}$, namely the restriction of the Killing form of L to $I \times I$.*

Proof. ***Lemma 5.1.2.*** *Let $T\colon V \to V$ be a linear transformation. Assume that $TV \subseteq W$. We have a linear transformation $T|_W\colon W \to W$. Then $\operatorname{tr} T = \operatorname{tr} T|_W$.*

Proof. Indeed, if $w_1, \ldots, w_k$ is a basis of W, and $w_{k+1}, \ldots, w_n$ is a completion to a basis of V, then Then T is of the form $\begin{pmatrix} T|_W & * \\ 0 & 0 \end{pmatrix}$. ■

Now, let $x \in I$ and $y \in L$, and consider the linear transformation $\operatorname{ad} x \operatorname{ad} y \colon L \to L \to I$. So by the lemma, $\operatorname{tr} \operatorname{ad} x \operatorname{ad} y$ on L is equal to $\operatorname{tr} \operatorname{ad} x \operatorname{ad} y|_I$. So, $K_L(x,y) = \operatorname{tr}(\operatorname{ad} \operatorname{ad} y|_I)$. But now take $x \in I$ and *also* $y \in I$. Then

$$K_I(x,y) = \operatorname{tr}(\operatorname{ad} x)|_I \cdot (\operatorname{ad} y)|_I = \operatorname{tr}((\operatorname{ad} x \operatorname{ad} y)|_I).$$

$K_L(x,y) = K_I(x,y)$ if $x, y \in I$. ■

Definition 5.1.3. The ***radical*** of a symmetric bilinear form B is defined by the following.

$$\operatorname{Rad}(B) = \{y \in V \mid B(x,y) = 0 \forall x \in V\}$$

Remark 5.1.4. $\operatorname{Rad}(B) \neq 0$ if and only if the form is degenerate.

Proposition 5.1.5. *The radical of the killing form is an ideal. In fact, this is true for* any *symmetric bilinear form which satisfies the symmetry condition, namely* $B([x,y],z) = B(x,[y,z])$.

Proof. Let $x \in \operatorname{Rad}(B)$ and $y \in L$. We should show that $[x,y] \in \operatorname{Rad}(L)$. But, for all $z \in L$, $B([x,y],z) = B(x,[y,z]) = 0$ since $x \in \operatorname{Rad}(B)$.
Hence $[x,y]$ is B-orthogonal to all $z \in L$, hence $[x,y] \in \operatorname{Rad}(B)$. ■

Conclusion. Consider the ideal $I = \operatorname{Rad}(K_L)$. Then $K_I = K_L|_{I \times I}$ as we proved. This restriction is obviously identically zero.
So, $I = \operatorname{Rad}(L)$ has $K_I(I,[I,I]) = 0$ and hence by the Cartan criterion is solvable.
Namely, the radical of the Killing form is a solvable ideal.

Conclusion. We defined $\operatorname{Rad}(L)$ as the unique maximal solvable ideal[1] and so we conclude from the above that $\operatorname{Rad}(K_L) \subseteq \operatorname{Rad}(L)$.

Conclusion. If L is a semi-simple Lie algebra, then by definition $\operatorname{Rad}(L) = 0$ and so $\operatorname{Rad}(K_L) = 0$, so the Killing form is non-degenerate.

Theorem 5.1.6. *A Lie algebra L is a semi-simple if and only if the Killing form K_L is non-degenerate. Namely,* $\operatorname{Rad}(L) = 0$ *if and only if* $\operatorname{Rad}(K_L) = 0$.

Proof. We already saw that $\operatorname{Rad}(K_L) \subseteq \operatorname{Rad}(L)$, so we know that $\operatorname{Rad}(L) = 0$ implies $\operatorname{Rad}(K_L)$.
Now assume that K_L is a non-degenerate form, and we have to show that L has no non-trivial solvable ideals. To do this, we start with a lemma.

Lemma 5.1.7. *If L has a non-trivial solvable ideal, then L has a non-trivial abelian ideal.*

Proof. First, if M is any solvable algebra, then for some k, $D_k(M) = [D_{k-1}(M), D_{k-1}(M)] = 0$ with $D_{k-1}(M) \neq 0$. So, by definition, $D_{k-1}(M)$ is a non-trivial abelian ideal of M.
Now, let L be a general Lie algebra, and I a general ideal of L. Then $D_k(I)$ are obviously ideals of I. In fact, they are also ideals of L!

Claim 5.1.8. $D_k(I) \lhd L$.

Remark 5.1.9. *Not* every ideal of I is an ideal of L!

Proof. To show that $D_k(I)$ is an ideal of L, it is necessary and sufficient to show that is an invariant sub-space under each of the linear transformations $\operatorname{ad} x \colon L \to L$ for all $x \in L$.
Now, $\operatorname{ad} x$ leaves I invariant, since I is an ideal, and $\operatorname{ad} x|_I$ is a derivation of I. It satisfies for all $y, z \in L$ that

$$\operatorname{ad} x[y,z] = [\operatorname{ad} x(y), z] + [y, \operatorname{ad} x(z)]$$

by the Jacobi identity. So, $\operatorname{ad} x \colon L \to L$ is a derivation, it is a map $\delta \colon L \to L$ satisfying

$$\delta[y,z] = [\delta(y), z] + [y, \delta(z)].$$

Hence $\operatorname{ad} x|_I \colon I \to I$ is a derivation of I.
We conclude the proof[2] by saying that $D_1(I) = [I,I]$ is invariant under *all* derivations of I:

$$\delta([I,I]) \subseteq [\delta(I), I] \subseteq [I,I]$$

Similarly, $D_k(I)$ are invariant under *all* derivations (proof by induction). So, each $D_k(I)$ is invariant under all $\operatorname{ad} x$ for $x \in L$. ■

[1] which contains every other solvable ideal
[2] of the claim

This claim proves the lemma, since if I is a solvable ideal, then $D_{k-1}(I) \neq 0$ and $D_k(I) = 0$ for some k, so $D_{k-1}(I)$ is an abelian ideal of I which is an ideal of L. ■

Going back to the proof of the theorem, we have L which has a non-degenerate Killing form and we show it has no non-trivial abelian ideals. By the lemma, it has then a trivial solvable radical.
Let $I \lhd L$ be abelian, namely $[I, I] = 0$. Let $x \in I$ and $y \in L$. Then $\operatorname{ad} x \operatorname{ad} y\colon L \to I$. We claim that $(\operatorname{ad} x \operatorname{ad} y)^2 : L \to [I, I] = 0$. If we prove this, $\operatorname{ad} x \operatorname{ad} y$ is a nilpotent linear transformation. So $\operatorname{tr} \operatorname{ad} x \operatorname{ad} y = K_L(x, y) = 0$ for all $y \in L$. Because K_L, is non-degenerate, this gives $x = 0$, hence $I = 0$.
To finish the proof, let $z \in L$ and let $w = \operatorname{ad} x \operatorname{ad} y(z) \in I$. Now,

$$\operatorname{ad} x \operatorname{ad} y(w) = (\operatorname{ad} x \operatorname{ad} y)^2(z) = [x, [y, w]] \in [I, I] = 0$$

since $[y, w] \in I$ (because $w \in I$). ■

Theorem 5.1.10 (The decomposition of semi-simple algebras to simple ideals). *Let L be a semi-simple Lie algebra. Then there are simple ideals $I_i \lhd L$, for $i \in [k]$, such that $L = \bigoplus_{i\in[k]} I_i$. I_i are uniquely determined up to order, and every simple ideal of L coincides with one of them.*

We remind that if A, B are Lie algebras, then

$$A \oplus B \coloneqq \{(a, b) \mid a \in A, b \in B\}$$

is a Lie algebra under

$$[(a, b), (a', b')] = ([a, a'], [b, b'])$$

and then $[(A, 0), (0, B)] = 0$ so the ideals $(A, 0)$ and $(0, B)$ commute.

Remark 5.1.11. For a semi-simple Lie algebra L, then $K_{I_i} = (K_L)_{I_i \times I_i}$.

Proof. If L is simple, the statement is obvious, so assume L is not simple.
Let J be an ideal of L, and assume that J is proper, and a minimal ideal. So $\dim J < \dim L$ and J does not contain a non-trivial ideal of L. Consider now

$$J^{\perp} \coloneqq \{y \in L \mid K(J, y) = 0\} = \{y \in L \mid K(x, y) = 0 \forall x \in J\}.$$

We claim that $J^{\perp}$ is also an ideal. If $x \in J$, $y \in L$ and $z \in J^{\perp}$ then $[y, z] \in J^{\perp}$ since

$$K([z, y], x) = K(z, [y, x]) = 0$$

for $z \in J^{\perp}$ and $[y, x] \in J$ as $x \in J$. Therefore $[z, y] \in J^{\perp}$. So, $I \coloneqq J \cap J^{\perp}$ is also an ideal. So, $J \cap J^{\perp}$ is also an ideal, but $K_I = (K_l)_{I \times I}$ is zero identically by definition.
So, I is a solvable ideal by the Cartan criterion. Since we assumed that $\operatorname{Rad}(L) = 0$, we get $I = 0$ so $J \cap J^{\perp} = 0$. Therefore $L = J \oplus J^{\perp}$.
Now both J and $J^{\perp}$ have trivial radical because L has trivial radical by the following sentence. $\left[J, J^{\perp}\right] \subseteq J \cap J^{\perp} = 0$ and $\left[J, J^{\perp}\right]$ *commute* which means an ideal of J or $J^{\perp}$ is an ideal of L.
In fact, we have shown that when the Killing form is non-degenerate, every ideal J has a direct complement $J^{\perp}$ which is also an ideal. Both of these also have non-degenerate Killing forms, since they have trivial radicals. By induction on the dimension, both J and $J^{\perp}$ are a direct sum of simple ideals, which are also ideals of L. Therefore L is a direct sum of simple ideals

$$L = \bigoplus_{i\in[k]} I_i$$

and $[I_i, I_j] = 0$ for $i \neq j$.
Now, let J be a simple ideal of L, we need to show $J = I_{i_0}$ for some i_0. Clearly

$$brsJ, L = \left[J, \bigoplus_{i\in[k]} I_i\right] = \bigoplus_{i\in[k]} [J, I_i].$$

$[J, I_i]$ is an ideal of J[3], so $[J, I_i] = 0$ or $[J, I_i] = J$.
$[J, L] \neq 0$ for otherwise $Z(L) \supseteq J$, but $Z(L)$ is an abelian ideal. So, $[J, I_{i_0}] = J$ for precisely one i_0.[4]
We want to show $J = I_{i_0}$. We already know $J \subseteq I_{i_0}$ because $J = [J, I_{i_0}] \subseteq I_{i_0}$. I_{i_0} is simple, and J is an ideal, so $J = I_{i_0}$. ■

[3] $[J, [J, I_i]] \subseteq [J, I_i]$ since $[J, I_i] \subseteq I_i$.
[4] Otherwise, J has non-trivial ideals.

Conclusion. Let L e a semi-simple algebra. So $L = \bigoplus_{i\in[k]} I_i$ for uniquely determined simple ideals. Then

1. Every ideal $I \triangleleft L$ is a direct sum of some of the I_i. This is true since I itself is a semi-simple algebra, and every simple ideal of I is a simple ideal of L.

   Proof. K_L is non-degenerate, so $I \oplus I^{\perp} = L$, again $I \cap I^{\perp} 0$ so $\left[I, I^{\perp}\right] \subseteq I \cap I^{\perp} = 0$, $I, I^{\perp}$ commute. So every ideal of I is an ideal of L. So, every simple ideal of I is a simple ideal of L. ■

2. Every factor ${}^{L}\!/_{I}$, where $I \triangleleft L$ is also isomorphic to a direct sum of simple ideals isomorphic to some of the I_i.
   This is an immediate consequence of 1.

Chapter 6

Derivations of simple algebras

Remind our assumptions $\operatorname{char}(\mathbb{F}) = 0$ and $\mathbb{F} = \bar{\mathbb{F}}$.

Theorem 6.0.1. *For every $\delta \in \operatorname{Der}(L)$, $\operatorname{Der}(L)$ contains also δ_n, δ_s. In fact, this is true for every $\mathbb{F}$-algebra L: The Lie algebra $\operatorname{Der}(L)$ contains the diagonalizable and nilpotent Jordan-Chevalley components of its elements.*

Proposition 6.0.2. *Define*

$$U_i = \left\{ v \in V \,\middle|\, \exists k \in \mathbb{N} \colon \left(T - a_i\right)^k v = 0 \right\}.$$

Then $U_i = W_i$.

Proof. U_i is a T-invariant subspace. So, $T|_{U_i}$ has a characteristic polynomial, and a minimal polynomial. Obviously, for some large N, $\left(T - a_i\right)^N \equiv 0$ on U_i. The minimal polynomial of $T|_{U_i}$ divides $\left(x - a_i\right)^N$. On the other hand, $f_T(x) \coloneqq \prod_{i \in [r]} \left(x - a_i\right)^{m_i}$ acts as zero of V and on U_i. So, the minimal polynomial of $T|_{U_i}$ divides f_T. So that minimal polynomial is a power of $(x - a_i)$, and the power is bounded by m_i. In particular, $\left(x - a_i\right)^{m_i} v = 0$ for all $v \in U_i$. ∎

Definition 6.0.3. Let $\delta \colon L \to L$ be a derivation of an arbitrary $\mathbb{F}$-algebra, we define

$$L_a \coloneqq \left\{ x \in L \,\middle|\, \exists k \colon \left(\delta - a\right)^k (x) = 0 \right\}.$$

Remark 6.0.4. L_a is a sub-space of L, which is non-zero when $a = a_i$ for the transformation $T = \delta$.

Lemma 6.0.5. *For every a and b, we have*

$$L_a \cdot L_b \subseteq L_{a+b}.$$

Proof. This follows from the generalised Leibniz formula.[1]

$$\left(\delta - (a+b) I\right)^n (x \circ y) = \sum_{j=0}^{n} \binom{n}{j} \left(\delta - aI\right)^{n-j} (x) \left(\delta - bI\right)^j (y)$$

For sufficiently large n, at least one of $\left(\delta - a_I\right)^{n-j}(x)$ or $\left(\delta - bI\right)^j (y)$ vanishes for $x \in L_a, y \in L_b$, for every j. Then

$$\left(\delta - (a+b) I\right)^n (x \circ y) = 0$$

and so $x \circ y \in L_{a+b}$. ∎

Proof (theorem 6.0.1). Let $\delta \in \operatorname{Der} L$ and write $\delta = \delta_s + \delta_n$ in $\operatorname{End}(L)$. To show $\delta_n, \delta_s \in \operatorname{Der} L$, it is enough to show that $\delta_s \in \operatorname{Der} L$.
To show that, we use that fact that it acts as a scalar in each of the subspaces $L_a \neq 0$. Namely, the scalar a. We use lemma 6.0.5 as follows. Fix $x \in L_a$ and $y \in L_b$ and compute

$$\delta_s (x \circ y) = (a+b)(x \circ y)$$

since $x \circ y \in L_{a+b}$. But,

$$\delta_s(x) \circ y + x \circ \delta_s(y) = ax \circ y + x \circ by = (a+b)(x \circ y).$$

So,

$$\delta_s(x \circ y) = \delta_s(x) \circ y + x \circ \delta_s(y),$$

and δ_s is a derivation. ∎

[1] The proof for the formula is done by induction, although it is surprisingly long.

Theorem 6.0.6. *Let L be a semi-simple algebra. Then* $\operatorname{ad}(L) = \operatorname{Der}(L)$.
Equivalently, the derivations of a semi-simple algebra are ***inner derivations****, namely, given by* $y \mapsto [x, y] = \operatorname{ad} x\,(y)$.

Lemma 6.0.7. *Let L be a Lie algebra and let $\delta\colon L \to L$ be a derivation. Remind that each $x \in L$ defines a linear map* $\operatorname{ad} x$ *which is also a derivation.*
Then $[\delta, \operatorname{ad} x] = \operatorname{ad} \delta\,(x)$.

Proof. In the exercise sheets. ∎

Corollary 6.0.8. $\operatorname{ad}(L) \subseteq \operatorname{Der}(L)$ *is an ideal of the algebra of derivations.*

Proof (theorem 6.0.6). Let $M := \operatorname{ad} L$, be an ideal of $\operatorname{Der} L$. Furthermore, $L \cong \operatorname{ad} L = M$, since $\ker \operatorname{ad}$ is the center $Z(L)$ and a semi-simple algebra has a trivial center. So the Killing form of L is non-degenerate. So the Killing form of M, being its restriction, is non-degenerate.[2]
Now consider $I = M^{\perp}$ the orthogonal of the Killing form of $\operatorname{Der} L$. Now $I \cap M = \{0\}$ since the Killing form of M (which is the restriction from $\operatorname{Der} L$) is non-degenerate. In addition, the orthogonal of an ideal under the Killing form is also an ideal, as we saw. So I is an ideal, and $\operatorname{Der} L = I \oplus M$.
Finally,

$$[I, M] \subseteq I \cap M = 0$$

since they are ideals. Hence I and M commute, so $[\delta, \operatorname{ad} x] = 0$ if $\delta \in I$, and

$$[\delta, \operatorname{ad} x] = 0 = \operatorname{ad} \delta\,(x)$$

as we saw. Now, ad is injective on L, so $\delta(x) = 0$ for all $x \in L$, and so $\delta = 0$. Hence $I = 0$ and hence $M = \operatorname{Der} L$. ∎

Theorem 6.0.9 (Abstract Jordan Chevalley decomposition of semi-simple Lie algebras). *For $x \in L$ we can write $x = x_s + x_n$, with the properties*

$$[x_s, x_n] = [x, x_s] = [x, x_n] = 0$$

and $\operatorname{ad} x_s$ *is diagonalisable on L,* $\operatorname{ad} x_n$ *is nilpotent on L, and the decomposition is* unique.

Proof. Write $\operatorname{ad} x = T_n + T_s$ a Jordan-Chevalley decomposition in $\operatorname{End} L$. From $\operatorname{Der} L = \operatorname{ad} L \cong L$ so T_n, T_s come from elements x_n, x_s with $x = x_s + x_n$. The properties follow from those on $\operatorname{End} L$. ∎

Exercise. Assume $L \subseteq \mathfrak{gl}(V)$ is semi-simple.
Show that the abstract Jordan-Chevalley decomposition of L agree with that of L viewed as linear transformations on V.

---

[2]We've seen that the restriction of the Killing form to an ideal, is the Killing form of the ideal.

Chapter 7

Toral subalgebras

Assume S is semi-simple, $\mathrm{char}(\mathbb{F}) = 0$ and $\mathbb{F} = \bar{\mathbb{F}}$. If every $x \in L$ is ad-nilpotent, then L is a nilpotent Lie algebra. So for some $x \in L$, we have $x_s \neq 0$[1] where $x = x_s + x_n$ and $\mathrm{ad}\, x_s$ is diagonalisable. So, x_n is ad-nilpotent and x_s is ad-diagonalisable. $\mathbb{F} \cdot x_s$ is a Lie sub-algebra consisting of ad-diagonalisable elements, and any such a sub-algebra is called a ***toral*** sub-algebra.

Theorem 7.0.1. *A toral sub-algebra is abelian.*

Proof. Let T be toral, and $x \in T$. We want to show that $\mathrm{ad}_T(x) = 0$, i.e. $\mathrm{ad}\, x \colon T \to T$ is the zero map. So we want to show $[x, y] = 0$ for all $y \in T$.
Assume towards a contradiction that this isn't the case. Then since $\mathrm{ad}\, x$ is diagonalisable, there's $y \in T$ non-zero which is an eigenvector of $\mathrm{ad}_T(x)$ with non-zero eigenvalue.
$\mathrm{ad}\, x$ is diagonalisable on L, hence on the invariant subspace T, and if all its eigenvalues are 0 it is the zero map. So

$$[x, y] = ay \neq y$$

with $a \neq 0$. We should derive a contradiction.
Consider

$$\mathrm{ad}_T(y, x) = [y, x] = -ay.$$

Then $\mathrm{ad}_T(y)(x)$ is a vector which is an eigenvector of $\mathrm{ad}_T(y)$ with eigenvalue 0.
But, $\mathrm{ad}_T(y)$ is diagonalisable in T, so *every* element of T is a linear combination of eigenvectors of $\mathrm{ad}_T(y)$. So

$$x = \sum_i f_i x_i$$

with x_i being eigenvectors of $\mathrm{ad}_T(y)$ and $f_i \neq 0$.
Apply $\mathrm{ad}_T(y)$ to this equation, we get

$$-ay = \mathrm{ad}_T(y)(x) = \sum_j f_j x_j$$

where the sum is ranging of j such that x_j is an eigenvector of $\mathrm{ad}_T(y)$ with non-zero eigenvalue and $f_j' \neq 0$.
But, we have this way a representation of the non-zero vector $-ay$ as a linear combination of eigenvectors of $\mathrm{ad}_T(y)$ with non-zero eigenvalue. Because y itself is an eigenvector with eigenvalue 0 of $\mathrm{ad}_T y$, this is a contradiction. ■

Let L be a semi-simple Lie algebra and let $0 \neq H \leq L$ be a maximal toral sub-algebra. Then each $h \in H$ has the property that $\mathrm{ad}\, h \colon L \to L$ is diagonalisable, and since H is abelian, L can be decomposed to joint eigen-spaces of all $h \in H$.

$$L_\alpha = \{x \in L \mid [h, x] = \alpha(h)\, x\}$$

Here, $\alpha \colon H \to \mathbb{F}$ is a linear functional. If x is a joint eigenvector, we get the following where $h_1, h_2 \in H$.

$$[h_1, x] = \alpha(h_1)\, x, \qquad [h_2, x] = \alpha(h_2)\, x, \qquad [h_1 + h_2, x] = \alpha(h_1 + h_2)\, x$$

We can write

$$L = C_L(H) \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$$

[1] L is semi-simple, so it isn't nilpotent.

where

$$C_L(H) = L_0 = \{x \in L \mid \forall h \in H\colon\ [h,x] = 0\}$$

is the ***centraliser of*** H ***in*** L and where

$$\Phi = \{\alpha \in H^* = \operatorname{Hom} H, \mathbb{F} \mid L_\alpha \neq 0\} \setminus \{0\}$$

is the set of roots of L under the toral sub-algebra H, ***the root system of*** H ***with respect to*** L. We also write

$$\Phi = \Phi(L, H).$$

We list a few properties of the root system $\Phi(L,H)$ in general.

Property 1.

$$[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta}$$

Proof. Let $x \in L_\alpha, y \in L_\beta$. To see $[x,y] \in L_{\alpha+\beta}$, compute using the Jacobi identity:

$$\begin{aligned}\operatorname{ad} h\,[x,y] &= [\operatorname{ad} h(x), y] + [x, \operatorname{ad} h(y)]\\ &= \alpha(h)[x,y] + \beta(h)[x,y]\\ &= (\alpha+\beta)(h)[x,y]\end{aligned}$$

Hence $[x,y] \in L_{\alpha+\beta}$. ■

Property 2. If $\alpha + \beta \neq 0$, then $L_\alpha \wedge L_\beta$ under the Killing form on L.

Proof. By the symmetry property of K,

$$K([h,x], y) = \alpha(h) K(x,y) = -K(x,[x,y]) = -\beta(h) K(x,y).$$

Therefore

$$(\alpha+\beta)(h) K(x,y) = 0,$$

but if $\alpha + \beta(h) \neq 0$ for some $h \in H$, then $K(x,y) = 0$, so $L_\alpha \wedge L_\beta$. ■

Proposition 7.0.2. *If L is a semi-simple algebra, then the restriction of K to $C_L(H)$ is non-degenerate.*

Remark 7.0.3. $C_L(H)$ is *not* an ideal of L.

Proof. Let $x \in C_L(H)$. Then since for $\alpha \neq 0$, $\alpha + 0 \neq 0$, $L_0 \perp L_\alpha$ for all $\alpha \in \Phi$.
So, if $x \in L_0$ is orthogonal to L_0 under K, $K(x, L_0) = 0$, then $K(x, L) = 0$ by linearity.
But in a semi-simple algebra, K is non-degenerate, so $x = 0$. Hence $K|_{L_0}$ is non-degenerate. ■

Theorem 7.0.4. *If L is a semi-simple Lie algebra and H is a maximal toral sub-algebra, then $C_L(H) = H$.*

Proof. Let $C = C_L(H)$.

Step 1. C contains the Jordan-Chevalley components of its elements. This follows since the Jordan-Chevalley components are polynomials without constant term in the underlying element. Now, an element of C maps (under ad) H to 0, and hence the same is true for the Jordan-Chevalley components. So the Jordan Chevalley components are in $C_L(H) = C$.

Step 2. Every ad-diagonalisable element $w \in C$ is in fact already in H. Indeed, $H + F \cdot w$ is also toral, since w commutes with every $h \in H$, and so by maximality of H, $w \in H$.

Step 3. The restriction of K to H is non-degenerate (not just that of K to C). Suppose $K(h, H) = 0$ for some h. Now let $n \in C$ be ad-nilpotent. Then, $\operatorname{ad} h \operatorname{ad} n$ is nilpotent since

$$[\operatorname{ad} h, \operatorname{ad} n] = \operatorname{ad}[h,n] = 0$$

and $\operatorname{ad} n$ is nilpotent. Therefore $\operatorname{tr}(\operatorname{ad} h \operatorname{ad} n) = 0$, so $K(h,n) = 0$. Therefore, h is K-orthogonal to every diagonalisable element in C[2], and to every ad-nilpotent element in C, and so $K(h, C) = 0$.
K is non-degenerate on C, so $h = 0$. Hence K is non-degenerate on H.

[2]since such an element is already in H as we saw, and we assume $K(h,H) = 0$

Step 4. C is a nilpotent algebra, since every element in C is ad-nilpotent and by Engel's theorem. Indeed if $x \in C$ then $x = x_n + x_s$, and $\operatorname{ad} x = \operatorname{ad} x_n + \operatorname{ad} x_s$, where $\operatorname{ad} x_s$ is diagonalisable, hence $x_s \in H$. Now

$$\operatorname{ad} x = \operatorname{ad} x_n + \operatorname{ad} x_s|_C$$

and since $x_s \in H$, $\operatorname{ad} x_s$ is the zero operator on C.
So, on C $\operatorname{ad} x$ is equal to an ad-nilpotent operator, so x is ad-nilpotent on C.

Step 5. $H \cap [C, C] = 0$ since by the symmetry property of the Killing form

$$K(H, [C, C]) = K([H, C], C) = K(0, C) \qquad \blacksquare$$

which is in turn true for $C = C_L(H)$ means $[C, H] = 0$.
Since K is non-degenerate on H, $H \cap [C, C] = 0$[3].

Step 6. Assume towards a contradiction $[C, C] \neq 0$.

Exercise. In a nilpotent Lie algebra, every non-zero ideal intersects the center non-trivially.

Let $z \in Z(C) \cap [C, C] \setminus \{0\}$. z is *not* ad-diagonalisable, because in that case it would be in H, and as wee just saw $H \cap [C, C] = 0$.
So, z has a non-trivial ad-nilpotent part, denoted z_n. Since z takes C to 0 under ad, so do z_s, z_n[4]. So, z_n is itself in $Z(C)$. But then $[z_n, c] = 0$ for all c, and

$$\operatorname{tr} \operatorname{ad} z_n \operatorname{ad} c = 0$$

since $\operatorname{ad} z_n$ is nilpotent, and commutes with $\operatorname{ad} C$. So, $K(z_n, c) = 0$ for all c, and so K is degenerate on C, unless z_n is zero. This contradiction implies $[C, C] = 0$ and C is abelian.

Step 7. We should show $C_L(H) = C = H$. Do this as an exercise.

Corollary 7.0.5. *The killing form of L restricted to H is a non-degenerate symmetric form.*

[3]otherwise a non-zero vector in it is the radical of the form on H. $h \in H \cap [C, C]$ implies $K(h, H) = 0$ and this implies $h = 0$.
[4]by the Jordan-Chevalley properties