1 Engel's theorem

We briefly review the main items from the last lecture.

Let L be a Lie algebra over a field $\mathbb{F}$.

If $I, J \subseteq L$ then define $[I, J] = \mathbb{F}\text{-span}\{[X, Y] : X \in I \text{ and } Y \in J\}$.

Some definitions from last time:

- L is *solvable* if $L^{(n)} = 0$ for some $n \gg 0$ where $L^{(0)} = L$ and $L^{(n+1)} = [L^{(n)}, L^{(n)}]$.
- L is *nilpotent* if $L^n = 0$ for some $n \gg 0$ where $L^0 = L$ and $L^{n+1} = [L, L^n] = [L^n, L]$.
- L is *semisimple* if its unique maximal solvable ideal $\mathsf{Rad}(L)$ is zero.

We will only discuss semisimple Lie algebras that are finite-dimensional.

Our proof of the existence of the unique maximal solvable ideal $\mathsf{Rad}(L)$ required $\dim(L) < \infty$.

A linear map $f : V \to V$ is *nilpotent* if $f^n = 0$ for some $n \gg 0$.

An element $X \in L$ is ad-*nilpotent* if the linear map $\mathsf{ad}_X = [X, \cdot] : L \to L$ is nilpotent.

Easy-to-prove fact: if L is a nilpotent Lie algebra then every $X \in L$ is ad-nilpotent.

Harder-to-prove converse:

Theorem (Engel's theorem)**.** If $\dim(L) < \infty$ and every $X \in L$ is ad-nilpotent then L is nilpotent.

One more result from last time, proved on the way to Engel's theorem.

Suppose V is a finite-dimensional $\mathbb{F}$-vector space and $L \subseteq \mathfrak{gl}(V)$ is a Lie subalgebra.

Recall that if $v_1, v_2, \dots, v_n$ is a basis for V and $X \in \mathfrak{gl}(V)$ then the corresponding matrix of X is the square array $[X_{ij}]_{1 \le i,j \le n}$ where $X_{ij} \in \mathbb{F}$ is such that $Xv_j = \sum_{i=1}^n X_{ij} v_i$.

Theorem. If all elements of L are nilpotent linear maps $V \to V$ then there exists a basis of V relative to which the matrices of all elements of L are strictly upper-triangular.

2 Lie's theorem

We now assume $\mathbb{F}$ is an algebraically closed field of characteristic zero (like the complex numbers $\mathbb{C}$).

We also assume that L is a Lie algebra over $\mathbb{F}$ of finite dimension.

Theorem. Assume L is a solvable Lie subalgebra of $\mathfrak{gl}(V)$ for some vector space V with $0 < \dim(V) < \infty$.

Then there exists a nonzero vector $0 \neq v \in V$ that is an eigenvector for every $X \in L$, meaning that

$$Xv = \lambda(X)v \quad \text{for all } X \in L$$

for some linear map $\lambda : X \to \mathbb{F}$.

Proof. One can prove this result by the following steps, similar to the proof of the previous theorem:

(0) Since $\mathbb{F}$ is algebraically closed, any given $X \in L$ has an eigenvector in V.

Hence the theorem holds if $\dim(L) \le 1$. Assume $\dim(L) > 1$.

(1) Find an ideal $K \subseteq L$ with $\dim(L) = \dim(K) + 1$.

This can be done by letting K be the preimage in L of any codimension one subspace of $L/[L, L]$.

This works since $L/[L, L]$ is nonzero (as L is solvable) and abelian (by the definition of $[L, L]$).

(2) By induction K has at least one common eigenvector in V, say with eigenvalue function $\lambda : K \to \mathbb{F}$.

(3) Let $W = \{v \in V : Yv = \lambda(Y)v \text{ for all } Y \in K\}$ be the common λ-eigenspace for K in V.

Then we can check that L stabilizes this nonzero subspace. This is the hard part of the argument.

(4) Finally choose any $Z \in L$ with $Z \notin K$. Step (3) shows that Z restricts to a linear map $W \to W$.

Since $\mathbb{F}$ is algebraically closed, Z has an eigenvector $v \in W$.

Since $\dim(L) = \dim(K) + 1$ we have $L = K \oplus \mathbb{F}Z$ and v is a common eigenvector for all of L.

The eigenvalue map $L \to \mathbb{F}$ is the linear extension of $\lambda : K \to \mathbb{F}$ sending Z to its eigenvalue for v.

Apart from some from minor details (see Humphrey's textbook) only step (3) requires more explanation. This is where we use the $\mathrm{char}(\mathbb{F}) = 0$ hypothesis.

We need to show that if $X \in L$ and $w \in W$ then $Xw \in W$, meaning that $YXw = \lambda(Y)Xw$ for all $Y \in K$. By the Jacobi identity we know that

$$YXw = XYw - [X, Y]w = \lambda(Y)Xw - \lambda([X, Y])w.$$

So it is enough to check that $\lambda([X, Y]) = 0$ for all $Y \in K$.

Proof of this claim. Let $n > 0$ be minimal such that w, Xw, X^2w, ..., X^nw are linearly dependent. Define $W_i = \mathbb{F}\text{-span}\{w,\ Xw,\ X^2w,\ \dots,\ X^{i-1}w\}$. Then $W_0 = 0$ and $\dim(W_i) = \min\{i, n\}$.

We argue that $YX^iw \in \lambda(Y)X^iw + W_i$ for all $Y \in K$. This is clear if $i = 0$ and if $i > 0$ then

$$YX^iw = YXX^{i-1}w = X(YX^{i-1}w) - [X, Y]X^{i-1}w.$$

This expression is in $\lambda(Y)X^iw + W_i$ when $Y \in K$, since then $[X, Y] \in K$ so by induction

$$YX^{i-1}w \in \lambda(Y)X^{i-1}w + W_{i-1} \quad \text{and} \quad [X, Y]X^{i-1}w \in \lambda([X, Y])X^{i-1}w + W_{i-1} \subseteq W_i.$$

Thus, relative to the basis w, Xw, X^2w, ..., $X^{n-1}w$ each $Y \in K$ operates on W_n as an upper-$\triangle$ matrix

$$\begin{bmatrix} \lambda(Y) & & * \\ & \ddots & \\ 0 & & \lambda(Y) \end{bmatrix}.$$

Therefore if $Y \in K$ then $\mathsf{trace}(Y|_{W_n}) = n\lambda(Y)$ and also $\mathsf{trace}([X, Y]|_{W_n}) = n\lambda([X, Y])$ as $[X, Y] \in K$. The notation $f|_{W_n}$ here means the linear map $W_n \to W_n$ obtained by restricting f.

But X and $Y \in K$ both preserve W_n so $\mathsf{trace}([X, Y]|_{W_n}) = \mathsf{trace}(X|_{W_n}Y|_{W_n}) - \mathsf{trace}(Y|_{W_n}X|_{W_n}) = 0$. Thus $n\lambda([X, Y]) = 0$ so since $\mathrm{char}(\mathbb{F}) = 0$ we conclude that $\lambda([X, Y]) = 0$. ■

Having checked this claim, we conclude that step (3) and the rest of the argument work as expected. □

Theorem (Lie's theorem)**.** Suppose $L \subseteq \mathfrak{gl}(V)$ is a solvable Lie subalgebra where $\dim(V) = n < \infty$. Then there is some basis of V relative to which the matrices of all elements of L are upper-triangular.

Proof. Choose $0 \neq v_1 \in V$ with $Xv_1 = \lambda(X)v_1$ for all $X \in L$ for some linear map $\lambda : L \to \mathbb{F}$.

Set $V_1 = \mathbb{F}v_1$ and consider the quotient space V/V_1.

Apply the theorem inductively to obtain a basis $v_2 + V_1$, $v_3 + V_1$, ..., $v_n + V_1$ for V/V_1 where each $v_i \in V$.

Then the desired basis for V is then $v_1, v_2, v_3, \dots, v_n$. □

Corollary. Suppose L is a solvable Lie algebra with $\dim(L) = n < \infty$.

Then there exists a chain of ideals $0 = L_0 \subseteq L_1 \subseteq L_2 \subseteq \cdots \subsetneq L_n = L$ with $\dim(L_i) = i$.

Proof. The Lie algebra $\mathsf{ad}(L) \subseteq \mathfrak{gl}(L)$ is a homomorphic image of L so is solvable.

Apply Lie's theorem to this Lie algebra with $V = L$.

This gives a basis $v_1, v_2, \dots, v_n \in L$ such that $\mathsf{ad}_X(v_i) \in \mathbb{F}\text{-span}\{v_1, v_2, \dots, v_i\}$ for all $X \in L$.

The subspaces $L_i = \mathbb{F}\text{-span}\{v_1, v_2, \dots, v_i\}$ are therefore ideals, and they form the desired chain. □

We can now generalize the fact that $\mathfrak{t}_n(\mathbb{F})$ is solvable and $\mathfrak{n}_n(\mathbb{F}) = [\mathfrak{t}_n(\mathbb{F}), \mathfrak{t}_n(\mathbb{F})]$ is nilpotent.

Corollary. Suppose L is a solvable Lie algebra of finite dimension. Then $[L, L]$ is nilpotent.

Proof. Choose a basis of L such that the matrices of $\mathsf{ad}_X \in \mathfrak{gl}(L)$ are all upper-triangular.

Then the matrix of $\mathsf{ad}_{[X,Y]} = [\mathsf{ad}_X, \mathsf{ad}_Y]$ is always strictly upper-triangular.

Hence ad_Z is nilpotent for all $Z \in [L, L]$, so Engel's theorem implies that $[L, L]$ is nilpotent. □

3 Remarks

Under the hypotheses that

- $\mathbb{F}$ is an algebraically closed field with $\mathrm{char}(\mathbb{F}) = 0$, and
- $L \subseteq \mathfrak{gl}(V)$ for an $\mathbb{F}$-vector space V with $\dim(V) = n < \infty$,

we have shown the following properties:

(1) If L is solvable that there exists a basis of V that makes all elements of L upper-triangular.

Equivalently, if L is solvable then there is an injective morphism $\iota : L \to \mathfrak{t}_n(\mathbb{F})$.

(2) If L is nilpotent, then L is also solvable.

But in this case L does not have to be strictly upper-triangular in the basis from (1).

In other words, we might have $\iota(L) \not\subseteq \mathfrak{n}_n(\mathbb{F})$.

Example: take $V = \mathbb{F}^n$ and let $L = \mathfrak{d}_n(\mathbb{F})$ be the abelian Lie algebra of diagonal matrices.

(3) The vector space V must have a basis that makes all elements of L strictly upper-triangular (meaning that there is an injective morphism $L \to \mathfrak{n}_n(\mathbb{F})$) when all elements L are nilpotent.

However, L can be a nilpotent Lie algebra without having this property.

Example: again let $V = \mathbb{F}^n$ and $L = \mathfrak{d}_n(\mathbb{F})$.

4 Jordan decomposition

In this section we do not require $\mathrm{char}(\mathbb{F}) = 0$ but still assume $\mathbb{F}$ is algebraically closed.

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

We say that $X \in \mathfrak{gl}(V)$ is *semisimple* if X is *diagonalizable*, meaning V has a basis of eigenvectors for X.

Some standard properties from linear algebra (which you may check as exercises):

(a) If $X, Y \in \mathfrak{gl}(V)$ are semisimple with $XY = YX$ then $aX + bY$ is semisimple for all $a, b \in \mathbb{F}$.

(b) If $X \in \mathfrak{gl}(V)$ is semisimple and X preserves a subspace $W \subseteq V$ then $X|_W \in \mathfrak{gl}(W)$ is semisimple.

We now quote some less obvious, but still standard results linear algebra.

We assume these facts as background and will not prove them ourselves.

Recall that $X \in \mathfrak{gl}(V)$ is *nilpotent* if $X^n = 0$ for some $n \gg 0$.

Proposition. For each $X \in \mathfrak{gl}(V)$, there are unique elements $X_{\mathrm{ss}}, X_{\mathrm{nil}} \in \mathfrak{gl}(V)$ such that

$$X_{\mathrm{ss}} \text{ is semisimple}, \quad X_{\mathrm{nil}} \text{ is nilpotent}, \quad X_{\mathrm{ss}}X_{\mathrm{nil}} = X_{\mathrm{nil}}X_{\mathrm{ss}}, \quad \text{and} \quad \boxed{X = X_{\mathrm{ss}} + X_{\mathrm{nil}}}.$$

The boxed formula is called the *Jordan decomposition* (or *Jordan–Chevalley decomposition*) of X.

Proof idea: if $V = \mathbb{F}^n$ and the Jordan canonical form of the matrix of X has blocks

$$\begin{bmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \lambda & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & \lambda \end{bmatrix}$$

then the Jordan canonical forms of X_{ss} and X_{nil} are respectively obtained by replacing these blocks by

$$\begin{bmatrix} \lambda & 0 & & & 0 \\ & \lambda & 0 & & \\ & & \lambda & \ddots & \\ & & & \ddots & 0 \\ 0 & & & & \lambda \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & 0 \end{bmatrix}.$$

Proposition. Let $X \in \mathfrak{gl}(V)$. There are polynomials $p(x), q(x) \in x\mathbb{F}[x]$ with $p(0) = q(0) = 0$ such that

$$X_{\mathrm{ss}} = p(X) \qquad \text{and} \qquad X_{\mathrm{nil}} = q(X).$$

Consequently, the operators X_{ss} and X_{nil} commute with any $Y \in \mathfrak{gl}(V)$ that has $XY = YX$.

Corollary. If $A \subseteq B \subseteq V$ are subspaces and $X \in \mathfrak{gl}(V)$ has $XB \subseteq A$ then $X_{\mathrm{ss}}B \subseteq A$and $X_{\mathrm{nil}}B \subseteq A$.

We now state some Lie theoretic results.

Continue to let V be a finite-dimensional $\mathbb{F}$-vector space.

Proposition. If $X \in \mathfrak{gl}(V)$ is nilpotent then so is $\mathsf{ad}_X \in \mathfrak{gl}(\mathfrak{gl}(V))$.

Proof. If $X^n = 0$ then $(\mathsf{ad}_X)^{2n} = 0$ since $(\mathsf{ad}_X)^{2n}(Y) \in \mathbb{F}\text{-span}\{X^i Y X^{2n-i} : 0 \le i \le 2n\}$. □

Proposition. If $X \in \mathfrak{gl}(V)$ is semisimple then so is $\mathsf{ad}_X \in \mathfrak{gl}(\mathfrak{gl}(V))$.

Proof. Suppose $v_1, v_2, \dots, v_n$ is a basis of V and $a_1, a_2, \dots, a_n \in \mathbb{F}$ are such that $Xv_i = a_i v_i$.

Let $e_{ij} \in \mathfrak{gl}(V)$ be the linear map with $e_{ij}(v_k) = \begin{cases} v_i & \text{if } j = k \\ 0 & \text{if } j \neq k. \end{cases}$

Then $\{e_{ij} : 1 \le i, j \le n\}$ is a basis for $\mathfrak{gl}(V)$ and $\mathsf{ad}_X e_{ij} = (a_i - a_j) e_{ij}$ since

$$(\mathsf{ad}_X e_{ij}) v_k = [X, e_{ij}] v_k = X \underbrace{e_{ij} v_k}_{\in \mathbb{F} v_i} - e_{ij} X v_k = a_i e_{ij} v_k - a_k e_{ij} v_k = (a_i - a_j) e_{ij} v_k$$

where the last equality uses the fact that $a_j e_{ij} v_k = a_k e_{ij} v_k$ since both sides are zero if $j \neq k$.

Thus the basis $\{e_{ij}\}$ for $\mathfrak{gl}(V)$ consists of eigenvectors for ad_X which is therefore diagonalizable. □

Proposition. Suppose $X \in \mathfrak{gl}(V)$ has Jordan decomposition $X = X_{\text{ss}} + X_{\text{nil}}$.

Then the Jordan decomposition of $\mathsf{ad}_X \in \mathfrak{gl}(\mathfrak{gl}(V))$ is $\mathsf{ad}_X = \mathsf{ad}_{X_{\text{ss}}} + \mathsf{ad}_{X_{\text{nil}}}$.

In other words, we have $(\mathsf{ad}_X)_{\text{ss}} = \mathsf{ad}_{X_{\text{ss}}}$ and $(\mathsf{ad}_X)_{\text{nil}} = \mathsf{ad}_{X_{\text{nil}}}$.

Proof. We have already seen that $\mathsf{ad}_{X_{\text{ss}}}$ is semisimple and $\mathsf{ad}_{X_{\text{nil}}}$ is nilpotent.

The adjoint representation is linear so $\mathsf{ad}_{X_{\text{ss}}} + \mathsf{ad}_{X_{\text{nil}}} = \mathsf{ad}_{X_{\text{ss}} + X_{\text{nil}}} = \mathsf{ad}_X$.

Finally, since X_{ss} and X_{nil} commute we have $[X_{\text{ss}}, X_{\text{nil}}] = 0$.

Therefore $\mathsf{ad}_{X_{\text{ss}}}$ and $\mathsf{ad}_{X_{\text{nil}}}$ also commute since $[\mathsf{ad}_{X_{\text{ss}}}, \mathsf{ad}_{X_{\text{nil}}}] = \mathsf{ad}_{[X_{\text{ss}}, X_{\text{nil}}]} = 0$ by the Jacobi identity.

Thus $\mathsf{ad}_X = \mathsf{ad}_{X_{\text{ss}}} + \mathsf{ad}_{X_{\text{nil}}}$ has the defining properties of the unique Jordan decomposition of ad_X. □