Math 128 Lecture 8

Engel's and Lie's theorem.

Engel's theorem

Define a Lie algebra $\mathbf{g}$ to be **nilpotent** if:

$$\exists n |\ [x_1, [x_2, \ldots x_{n+1}] \ldots] = 0 \quad \forall\ x_1, \ldots, x_{n+1} \in \mathbf{g}.$$

Example: $\mathbf{n}^+ := \mathbf{n}^+(gl(d)) :=$ all strictly upper triangular matrices. Notice that the product of any $d+1$ such matrices is zero.

The claim is that all nilpotent Lie algebras are essentially like $\mathbf{n}^+$.

We can reformulate the definition of nilpotent as saying that the product of any n operators $\operatorname{ad} x_i$ vanishes. One version of Engel's theorem is

Theorem 1 $\mathbf{g}$ *is nilpotent if and only if* $\operatorname{ad} x$ *is a nilpotent operator for each* $x \in \mathbf{g}$.

Another version of Engel's theorem.

Theorem 1 $\mathbf{g}$ *is nilpotent if and only if* $\mathrm{ad}\, x$ *is a nilpotent operator for each* $x \in \mathbf{g}$.

This follows (taking $V = \mathbf{g}$ and the adjoint representation) from

Theorem 2 Engel *Let* $\rho : \mathbf{g} \to \mathrm{End}(V)$ *be a representation such that* $\rho(x)$ *is nilpotent for each* $x \in \mathbf{g}$. *Then there exists a basis in terms of which* $\rho(\mathbf{g}) \subset \mathbf{n}^+(gl(d))$, *i.e. becomes strictly upper triangular. Here* $d = \dim V$.

A third version of Engel's theorem.

Theorem 2 Engel *Let* $\rho : \mathbf{g} \to \mathrm{End}(V)$ *be a representation such that* $\rho(x)$ *is nilpotent for each* $x \in \mathbf{g}$. *Then there exists a basis in terms of which* $\rho(\mathbf{g}) \subset \mathbf{n}^+(gl(d))$, *i.e. becomes strictly upper triangular. Here* $d = \dim V$.

Given a single nilpotent operator, we can always find a non-zero vector, v which it sends into zero. Then on $V/\{v\}$ a non-zero vector which the induced map sends into zero etc. So in terms of such a flag, the corresponding matrix is strictly upper triangular. The theorem asserts that we are can find a single flag which works for all $\rho(x)$. In view of the above proof for a single operator, Engel's theorem follows from the following simpler looking statement:

Theorem 3 *Under the hypotheses of Engel's theorem, if* $V \neq 0$, *there exists a non-zero vector* $v \in V$ *such that* $\rho(x)v = 0 \ \forall x \in \mathbf{g}$.

Proof of Theorem 3 in seven easy steps.

- Replace $\mathbf{g}$ by its image, i.e. assume that $\mathbf{g} \subset \operatorname{End} V$.

- Then $(\operatorname{ad} x)y = L_x y - R_x y$ where L_x is the linear map of $\operatorname{End} V$ into itself given by left multiplication by x, and R_x is given by right multiplication by x. Both L_x and R_x are nilpotent as operators since x is nilpotent. Also they commute. Hence by the binomial formula $(\operatorname{ad} x)^n = (L_x - R_x)^n$ vanishes for sufficiently large n.

- We may assume (by induction) that for any Lie algebra, $\mathbf{m}$, of smaller dimension than that of $\mathbf{g}$ (and any representation) there exists a $v \in V$ such that $xv = 0\ \forall x \in \mathbf{m}$.

Step 4.

- Let $\mathbf{k} \subset \mathbf{g}$ be a subalgebra, $\mathbf{k} \neq \mathbf{g}$, and let

$$N = N(\mathbf{k}) := \{x \in \mathbf{g} | (\operatorname{ad} x)\mathbf{k} \subset \mathbf{k}\}$$

be its normalizer. The claim is that

1 $N(\mathbf{k})$ is strictly larger than $\mathbf{k}$.

To see this, observe that each $x \in \mathbf{k}$ acts on $\mathbf{k}$ and on $\mathbf{g}/\mathbf{k}$ by nilpotent maps, and hence there is an $0 \neq \hat{y} \in \mathbf{g}/\mathbf{k}$ killed by all $x \in \mathbf{k}$. But then $y \notin \mathbf{k}$, and $[y, x] = -[x, y] \in \mathbf{k}$ for all $x \in \mathbf{k}$. So $y \in N(\mathbf{k}),\ y \notin \mathbf{k}$.

Steps 5 and 6.

- If $\mathbf{g} \neq 0$, there is an ideal $\mathbf{i} \subset \mathbf{g}$ such that $\dim \mathbf{g}/\mathbf{i} = 1$. Indeed, let $\mathbf{i}$ be a maximal proper subalgebra of $\mathbf{g}$. Its normalizer is strictly larger, hence all of $\mathbf{g}$, so $\mathbf{i}$ is an ideal. The inverse image in $\mathbf{g}$ of a line in $\mathbf{g}/\mathbf{i}$ is a subalgebra, and is strictly larger than $\mathbf{i}$. Hence it must be all of $\mathbf{g}$.

- Choose such an ideal, $\mathbf{i}$. The subspace

$$W \subset V, \quad W = \{v | xv = 0, \ \forall x \in \mathbf{i}\}$$

  is invariant under $\mathbf{g}$. Indeed, if $y \in \mathbf{g}$, $w \in W$ then $xyw = yxw + [x, y]w = 0$.

End of proof.

- Choose such an ideal, $\mathbf{i}$. The subspace

$$W \subset V, \quad W = \{v | xv = 0, \ \forall x \in \mathbf{i}\}$$

  is invariant under $\mathbf{g}$. Indeed, if $y \in \mathbf{g}$, $w \in W$ then $xyw = yxw + [x, y]w = 0$.

- $W \neq 0$ by induction. Take $y \in \mathbf{g}$, $y \notin \mathbf{i}$. It preserves W and is nilpotent. Hence there is a non-zero $v \in W$ with $yv = 0$. Since y and $\mathbf{i}$ span $\mathbf{g}$, we have $xv = 0 \ \ \forall x \in \mathbf{g}$. QED

No assumptions about the ground field went into this.

Solvable Lie algebras.

Let $\mathbf{g}$ be a Lie algebra. $D^n\mathbf{g}$ is defined inductively by

$$D^0\mathbf{g} := \mathbf{g},\ D^1(\mathbf{g}) := [\mathbf{g},\mathbf{g}], \ldots,\quad D^{n+1}\mathbf{g} := [D^n\mathbf{g}, D^n\mathbf{g}].$$

If we take $\mathbf{b}$ to consist of all upper triangular $n \times n$ matrices, then $D^1\mathbf{b} = \mathbf{n}^+$ consists of all strictly triangular matrices and then successive brackets eventually lead to zero. We claim that the following conditions are equivalent and any Lie algebra satisfying them is called **solvable**.

1. $\exists n \ \ | D^n\mathbf{g} = 0.$

2. $\exists n$ such that for every family of 2^n elements of $\mathbf{g}$ the successive brackets of brackets vanish; e.g for $n = 4$ this says

$$[[[[x_1, x_2], [x_3, x_4]], [[x_5, x_6], [x_7, x_8]]], [[[x_9, x_{10}], [x_{11}, x_{12}]], [[x_{13}.x_{14}], [x_{15}, x_{16}]]]] = 0.$$

3. There exists a sequence of subspaces $\mathbf{g} := \mathbf{i}_1 \supset \mathbf{i}_2 \supset \cdots \supset \mathbf{i}_n = 0$ such each is an ideal in the preceding and such that the quotient $\mathbf{i}_j/\mathbf{i}_{j+1}$ is abelian, i.e. $[\mathbf{i}_j, \mathbf{i}_j] \subset \mathbf{i}_{j+1}$.

Equivalence of the three definitions.

1. $\exists n \ \ | D^n \mathbf{g} = 0.$

2. $\exists n$ such that for every family of 2^n elements of $\mathbf{g}$ the successive

   brackets of brackets vanish;

3. There exists a sequence of subspaces $\mathbf{g} := \mathbf{i}_1 \supset \mathbf{i}_2 \supset \cdots \supset \mathbf{i}_n = 0$ such each is an ideal in the preceding and such that the quotient $\mathbf{i}_j / \mathbf{i}_{j+1}$ is abelian, i.e. $[\mathbf{i}_j, \mathbf{i}_j] \subset \mathbf{i}_{j+1}$.

Proof of the equivalence of these conditions. $[\mathbf{g}, \mathbf{g}]$ is always an ideal in $\mathbf{g}$ so the $D^j \mathbf{g}$ form a sequence of ideals demanded by 3), and hence 1) $\Rightarrow$ 3). We also have the obvious implications 3) $\Rightarrow$ 2) and 2) $\Rightarrow$ 1). So all these definitions are equivalent.

Lie's theorem.

Theorem 4 [Lie.] *Let* $\mathbf{g}$ *be a solvable Lie algebra over an algebraically closed field* k *of characteristic zero, and* (ρ, V) *a finite dimensional representation of* $\mathbf{g}$*. Then we can find a basis of* V *so that* $\rho(\mathbf{g})$ *consists of upper triangular matrices.*

By induction on dim V this reduces to

Theorem 5 [Lie.] *Under the same hypotheses, there exists a (non-zero) common eigenvector* v *for all the* $\rho(y)$*, i.e. there is a vector* $v \in V$ *and a function* $\chi : \mathbf{g} \to k$ *such that*

$$\rho(y)v = \chi(y)v \quad \forall\ y \in \mathbf{g}. \tag{1}$$

A lemma for Lie's theorem.

$$\rho(y)v = \chi(y)v \quad \forall\ y \in \mathbf{g}. \tag{1}$$

Lemma 1 *Suppose that* $\mathbf{i}$ *is an ideal of* $\mathbf{g}$ *and (1) holds for all* $y \in \mathbf{i}$. *Then*

$$\chi([x,h]) = 0, \quad \forall\ x \in \mathbf{g} \quad h \in \mathbf{i}.$$

Proof of lemma. For $x \in \mathbf{g}$ let V_i be the subspace spanned by $v, xv, \ldots, x^{i-1}v$ and let $n > 0$ be minimal such that $V_n = V_{n+1}$. So V_n is finite dimensional and $xV_n \subset V_n$. Also $V_n = V_{n+k}\ \forall k$.

Also, for $h \in \mathbf{i}$, (dropping the ρ) we have:

$$\begin{aligned} hv &= \chi(h)v \\ hxv &= xhv - [x,h]v \\ &\equiv \chi(h)xv \mod V_1 \end{aligned}$$

Continuing:

$$\begin{aligned}
hv &= \chi(h)v \\
hxv &= xhv - [x,h]v \\
&\equiv \chi(h)xv \mod V_1 \\
hx^2v &= xhxv + [h,x]xv \\
&\equiv \chi(h)x^2v + uxv, \mod V_1\ u \in I \\
&\equiv \chi(h)x^2v + \chi(u)xv \mod V_1 \\
&= \chi(h)x^2v \mod V_2 \\
&\vdots \quad\quad \vdots \\
hx^iv &\equiv \chi(h)x^iv \mod V_i.
\end{aligned}$$

Thus V_n is invariant under $\mathbf{i}$ and for each $h \in \mathbf{i}$, $\mathrm{tr}_{|V_n} h = n\chi(h)$. In particular both x and h leave V_n invariant and $\mathrm{tr}_{|V_n}[x,h] = 0$ since the trace of any commutator is zero. This proves the lemma.

Proof of Lie's theorem.

We want to find v such that

$$\rho(y)v = \chi(y)v \quad \forall\ y \in \mathbf{g}. \tag{1}$$

Proof of theorem by induction on dim $\mathbf{g}$, which we may assume to be positive. Let $\mathbf{m}$ be any subspace of $\mathbf{g}$ with $\mathbf{g} \supset \mathbf{m} \supset [\mathbf{g}, \mathbf{g}]$. Then $[\mathbf{g}, \mathbf{m}] \subset [\mathbf{g}, \mathbf{g}] \subset \mathbf{m}$ so $\mathbf{m}$ is an ideal in $\mathbf{g}$. In particular, we may choose $\mathbf{m}$ to be a subspace of codimension 1 containing $[\mathbf{g}, \mathbf{g}]$. By induction we can find a $v \in V$ and a $\chi : \mathbf{m} \to k$ such that (1) holds for all elements of $\mathbf{m}$. Let

$$W := \{w \in V | hw = \chi(h)w\ \forall\ h \in \mathbf{m}\}.$$

If $x \in \mathbf{g}$, then

$$hxw = xhw - [x, h]w = \chi(h)xw - \chi([x, h])w = \chi(h)xw$$

since $\chi([x, h]) = 0$ by the lemma. Thus W is stable under all of $\mathbf{g}$.

$$W := \{w \in V | hw = \chi(h)w \ \forall \ h \in \mathbf{m}\}.$$

$$W \text{ is stable under all of } \mathbf{g}.$$

Pick $x \in \mathbf{g}$, $x \notin \mathbf{m}$, and let $v \in W$ be an eigenvector of x with eigenvalue λ, say. Then v is a simultaneous eigenvector for all of $\mathbf{g}$ with χ extended as

$$\chi(h + rx) = \chi(h) + r\lambda. \ \text{QED}$$

We had to divide by n in the above argument. In fact, the theorem is not true over a field of characteristic 2, with $sl(2)$ as a counterexample.

Application of Lie's theorem to the adjoint representation.

Applied to the adjoint representation, Lie's theorem says that there is a flag of ideals with commutative quotients, and hence $[\mathbf{g}, \mathbf{g}]$ is nilpotent.

Facts about polynomials over a field.

The Euclidean algorithm.

This says that if f and g are polynomials with coefficients in a field, then there exist unique polynomials q and r such that

$$\deg r < \deg g$$

and

$$f = gq + r.$$

The proof is by long division: Write

$$\begin{aligned} f &= a_n X^n + \cdots + a_0 \qquad a_n \neq 0, \\ g &= b_d X^d + \cdots + b_0 \qquad b_d \neq 0 \end{aligned}$$

so that $n = \deg f$ and $d = \deg g$. The algorithm (and proof) is by induction on n. If $n = 0$ and $d > 0$ we must have $q = 0$ and $r = f$ and this works. If $n = 0$ and $d = 0$ then both f and g are constants so we must have $r = 0$ and $q = f/g$ a constant.

The euclidean algorithm, continued.

Assume that we have proved the theorem for all f of degree $< n$ where $n > 0$. If $d > n$ we must have $q = 0$ and $r = f$ and this works. So we may assume that $d \leq n$. Write

$$f = \frac{a_n}{b_d} X^{n-d} g + h \qquad \text{where} \quad \deg h < n.$$

By induction

$$h = pg + r$$

proving the result where $g = \frac{a_n}{b_d} X^{n-d} + p$. (The uniqueness is easy to prove.)

Every ideal in the ring of polynomials over a field is principal.

Let $\mathfrak{a}$ be such an ideal. If $\mathfrak{a} = \{0\}$ there is nothing to prove, so assume that $\mathfrak{a} \neq \{0\}$. Choose $g \in \mathfrak{a}$ to have smallest degree. Then applying the euclidean algorithm to any $f \in \mathfrak{a}$ gives $f = qg + r$ where $\deg r < \deg g$ and so $r = 0$. So every element of $\mathfrak{a}$ is a multiple of g. We write

$$\mathfrak{a} = (g).$$

A consequence is that if f and g do not have a common (non-constant) factor then the ideal

$$(f) + (g)$$

is the entire ring of polynomials. In general, in a commutative ring A with unit, we say that two ideals $\mathfrak{a}$ and $\mathfrak{b}$ are **mutually prime** if

$$\mathfrak{a} + \mathfrak{b} = A.$$

This is the same as saying that there exist $a \in \mathfrak{a}$ and $b \in \mathfrak{b}$ such that

$$1 = a + b.$$

The Chinese remainder theorem.

Let A be a commutative ring with unit. Let $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ be pairwise mutually prime ideals, i.e.

$$\mathfrak{a}_i + \mathfrak{a}_j = A \quad \text{if} \quad i \neq j.$$

The theorem says that given any $x_1, \ldots, x_n \in A$ there exists an $x \in A$ such that

$$x \equiv x_i \mod \mathfrak{a}_i.$$

The proof is by induction on n. For $n = 1$ there is nothing to prove. Let $n = 2$ so

$$1 = a_1 + a_2, \qquad a_1 \in \mathfrak{a}_1, \qquad a - 2 \in \mathfrak{a}_2.$$

Take $x = x_2 a_1 + x_1 a_2$. Since $1 \equiv a_2 \mod x_1$ we have $x \equiv a_1 \mod \mathfrak{a}_1$ and similarly $x \equiv x_2 \mod \mathfrak{a}_2$.

Assume that we have proved the theorem for $n-1$ ideals. For each $i \geq 2$ choose $a_i \in \mathfrak{a}_1$ and $b_i \in \mathfrak{a}_i$ such that

$$a_i + b_i = 1 \quad \forall\ i \geq 2.$$

Then

$$\prod_{i=2}^{n}(a_i + b_i) = 1$$

and the only element in this product which does not belong to $\mathfrak{a}_1$ is the term $b_2 \cdots b_n$ so

$$\mathfrak{a}_1 + \prod_{i=2}^{n} \mathfrak{a}_i = A.$$

By the result we already know for $n = 2$ we can find a $y_1 \in A$ such that

$$y_1 \equiv 1 \mod \mathfrak{a}_1 \quad \text{and} \quad y_1 \equiv 0 \mod \prod_{i=2}^{n} \mathfrak{a}_i.$$

In particular,

$$y_1 \equiv 0 \mod \mathfrak{a}_1 \quad \text{and} \quad y_1 \equiv 0 \mod \mathfrak{a}_j, \quad j \neq 1.$$

Conclusion of the proof of the Chinese remainder theorem

In the same way we find elements $y_i \in A$ such that

$$y_i \equiv 1 \mod \mathfrak{a}_i \quad \text{and} \quad y_i \equiv 0 \mod \mathfrak{a}_j, \quad j \neq i.$$

Then

$$x = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

satisfies

$$x \equiv x_i \mod \mathfrak{a}_i.$$

Some facts from linear algebra.

Let V be a finite dimensional vector space over an algebraically closed field of characteristic zero, and let

$$\det(TI - u) = \prod (T - \lambda_i)^{m_i}$$

be the factorization of its characteristic polynomial where the λ_i are distinct. Let $S(T)$ be any polynomial satisfying

$$S(T) \equiv \lambda_i \bmod (T - \lambda_i)^{m_i}, \quad S(T) \equiv 0 \bmod T,$$

which is possible by the Chinese remainder theorem. For each i let $V_i :=$ the kernel of $(u - \lambda_i)^{m_i}$. Then $V = \bigoplus V_i$ and on V_i, the operator $S(u)$ is just the scalar operator $\lambda_i I$. In particular $s = S(u)$ is semi-simple (its eigenvectors span V) and, since s is a polynomial in u it commutes with u. So

$$u = s + n$$

where

$$n = N(u), \ N(T) = T - S(T)$$

Decomposition of a linear transformation into its semi-simple and nilpotent parts.

We have shown that $$u = s + n$$

where

$$n = N(u),\ N(T) = T - S(T)$$

is nilpotent. Also

$$ns = sn.$$

We claim that these two elements are uniquely determined by

$$u = s + n, sn = ns,$$

with s semisimple and n nilpotent. Indeed, since $sn = ns$, $su = us$ so $s(u - \lambda_i)^k = (u - \lambda_i)^k s$ so $sV_i \subset V_i$. Since $s - u$ is nilpotent, s has the same eigenvalues on V_i as u does, i.e. λ_i. So s and hence n is uniquely determined.

If $P(T)$ is any polynomial with vanishing constant term, then if $A \subset B$ are subspaces with $uB \subset A$ then $P(u)B \subset A$. So, in particular, $sB \subset A$ and $nB \subset A$.

The derivation action on tensors.

Define

$$V_{p,q} := V \otimes V \otimes \cdots \otimes V \otimes V^* \otimes \cdots \otimes V^*$$

with p copies of V and q copies of V^*. Let $u \in \mathrm{End}(V)$ act on V^* by $-u^*$ and on V_{pq} by derivation, so , for example,

$$u_{12} = u \otimes 1 \otimes 1 - 1 \otimes u^* \otimes 1 - 1 \otimes 1 \otimes u^*.$$

Similarly, u_{11} acts on $V_{1,1} = V \otimes V^*$ by

$$u_{11}(x \otimes \ell) = ux \otimes \ell - x \otimes u^*\ell.$$

The derivation action on End(V).

Under the identification of $V \otimes V^*$ with $\mathrm{End}(V)$, the element $x \otimes \ell$ acts on $y \in V$ by sending it into

$$\ell(y)x.$$

So the element $u_{11}(x \otimes \ell)$ sends y to

$$\ell(y)u(x) - (u^*\ell)(y)x = \ell(y)u(x) - \ell(u(y))x.$$

This is the same as the commutator of the operator u with the operator (corresponding to) $x \otimes \ell$ acting on y. In other words, under the identification of $V \otimes V^*$ with $\mathrm{End}(V)$, the linear transformation u_{11} gets identified with $\mathrm{ad}\, u$.

The decomposition of the derivation action.

Proposition 1 *If* $u = s + n$ *is the decomposition of* u *then* $u_{pq} = s_{pq} + n_{pq}$ *is the decomposition of* u_{pq}.

Proof. $[s_{pq}, n_{pq}] = 0$ and the tensor products of an eigenbasis for s is an eigenbasis for s_{pq}. Also n_{pq} is a sum of commuting nilpotents hence nilpotent. The map $u \mapsto u_{pq}$ is linear hence $u_{pq} = s_{pq} + n_{pq}$. QED

Additive maps.

If $\phi : k \to k$ is a map, we define $\phi(s)$ by $\phi(s)_{|V_i} = \phi(\lambda_i)$. If we choose a polynomial such that $P(0) = 0$, $P(\lambda_i) = \phi(\lambda_i)$ then $P(u) = \phi(s)$.

Proposition 2 *Suppose that ϕ is additive. Then*

$$(\phi(s))_{pq} = \phi(s_{pq}).$$

Proof. Decompose V_{pq} into a sum of tensor products of the V_i or V_j^*. On each such space we have

$$\begin{aligned} \phi(s_{p,q}) &= \phi(\lambda_{i_1} + \cdots - \cdots) \\ &= \phi(\lambda_{i_1}) + \phi(..).. \\ &= (\phi(s))_{p,q} \end{aligned}$$

where the middle equation is just the additivity. QED

Two important propositions.

As an immediate consequence we obtain

Proposition 3 *Notation as above. If* $A \subset B \subset V_{p,q}$ *with* $u_{pq}B \subset A$ *then for any additive map,* $\phi(s)_{pq}B \subset A$

Proposition 4 (over $\mathbf{C}$) *Let* $u = s + n$ *as above. If* $\mathrm{tr}(u\phi(s)) = 0$ *for* $\phi(s) = \overline{s}$ *then* u *is nilpotent.*

Proof. $\mathrm{tr}\, u\phi(s) = \sum m_i \lambda_i \overline{\lambda_i} = \sum m_i |\lambda_i|^2$. So the condition implies that all the $\lambda_i = 0$. QED