Lecture 9 — Chevalley's Theorem

Prof. Victor Kac *Scribes: Gregory Minton, Aaron Potechin*

In the last lecture we defined Cartan subalgebras and gave a construction using regular elements. In this lecture we will show that this construction is essentially unique by proving Chevalley's Theorem on conjugacy of Cartan subalgebras.

To state the theorem, we need the notion of the exponential of a nilpotent operator.

Definition 9.1. Let A be a nilpotent operator on a vector space V over a field $\mathbb{F}$ of characteristic 0. Define $e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots$. (As A is nilpotent, this is a *finite* sum.)

Exercise 9.1. If A and B are commuting nilpotent operators, then we know that $A+B$ is also nilpotent, so e^{A+B}, e^A, and e^B are all defined. Show that $e^{A+B} = e^A e^B$. In particular, deduce that $e^A e^{-A} = I$, so e^A is always an invertible operator.

Solution: As A, B, and $A+B$ are nilpotent, we can find an integer $N \geq 0$ such that $A^{N+1} = B^{N+1} = (A+B)^{(2N+1)} = 0$. Because A and B commute, we can use the binomial theorem to get

$$e^{A+B} = \sum_{n=0}^{2N} \frac{(A+B)^n}{n!} = \sum_{n=0}^{2N} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k} = \sum_{\substack{k,\ell \geq 0 \\ k+\ell \leq 2N}} \frac{1}{k!\ell!} A^k B^\ell.$$

Every term above with $k > N$ or $\ell > N$ vanishes, as either $A^k = 0$ or $B^\ell = 0$. Conversely, every term with $k \leq N$ and $\ell \leq N$ is present in the sum. Thus

$$e^{A+B} = \sum_{k,\ell \leq N} \frac{1}{k!\ell!} A^k B^\ell = \left(\sum_{k=0}^{N} \frac{A^k}{k!} \right) \left(\sum_{\ell=0}^{N} \frac{B^\ell}{\ell!} \right) = e^A e^B,$$

as desired. The "deduce" part is immediate, as A and $-A$ commute and $e^0 = I$. □

Exercise 9.2. Let $\mathfrak{g}$ be an arbitrary (not necessarily Lie) algebra over a field $\mathbb{F}$ of characteristic 0, and let D be a nilpotent derivation of $\mathfrak{g}$. Show that e^D is an automorphism of $\mathfrak{g}$.

Solution: We already know from Exercise 9.1 that e^D is a linear isomorphism; the content of this exercise is that it also respects multiplication in $\mathfrak{g}$. Let $\mu : \mathfrak{g} \otimes_F \mathfrak{g} \to \mathfrak{g}$ denote the multiplication map. Define linear transformations D_1 and D_2 on $\mathfrak{g} \otimes_F \mathfrak{g}$ by $D_1(a \otimes b) = D(a) \otimes b$ and $D_2(a \otimes b) = a \otimes D(b)$. It is clear that D_1 and D_2 commute, that they are both nilpotent, and that $e^{D_1}(a \otimes b) = e^D(a) \otimes b$ and $e^{D_2}(a \otimes b) = a \otimes e^D(b)$. Next, because D is a derivation, notice that

$$(\mu \circ (D_1 + D_2))(a \otimes b) = D(a) \cdot b + a \cdot D(b) = D(a \cdot b) = (D \circ \mu)(a \otimes b),$$

so $\mu \circ (D_1 + D_2) = D \circ \mu$. It follows easily that $\mu \circ e^{D_1+D_2} = e^D \circ \mu$. Applying Exercise 9.1,

$$e^D(a \cdot b) = (e^D \mu)(a \otimes b) = (\mu e^{D_1+D_2})(a \otimes b) = (\mu e^{D_1} e^{D_2})(a \otimes b) = e^D(a) \cdot e^D(b),$$

as desired. □

We are now ready to state the main result.

Theorem 9.1 (Chevalley)**.** *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra of an algebraically closed field $\mathbb{F}$ of characteristic 0. Denote by G the subgroup of the group of automorphisms of $\mathfrak{g}$ which is generated by automorphisms of the form $e^{\operatorname{ad} a}$ for $a \in \mathfrak{g}$ such that $\operatorname{ad} a$ is nilpotent. Then any two Cartan subalgebras $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are conjugate by G, i.e. there exists $\sigma \in G$ such that $\sigma(\mathfrak{h}_1) = \mathfrak{h}_2$.*

Remark 9.1. The group G in Chevalley's Theorem is almost (but not quite) the Lie group associated to the Lie algebra $\mathfrak{g}$.

Before proving Chevalley's Theorem, we give a corollary that addresses the question with which we opened the lecture.

Corollary 9.2. *Let $\mathbb{F}$ be an algebraically closed field of characteristic 0 and let $\mathfrak{g}$ be a finite-dimensional Lie algebra over $\mathbb{F}$. Then any Cartan subalgebra $\mathfrak{h}$ in $\mathfrak{g}$ is of the form $\mathfrak{g}_0^a$ for some regular element $a \in \mathfrak{g}$, and in particular $\dim \mathfrak{h} = \operatorname{rank} \mathfrak{g}$. Also, all such subalgebras $\mathfrak{g}_0^a$ are isomorphic.*

Proof. Fix a regular element $a \in \mathfrak{g}$. By Chevalley's Theorem, any Cartan subalgebra $\mathfrak{h}$ is conjugate to $\mathfrak{g}_0^a$, say $\mathfrak{h} = \sigma(\mathfrak{g}_0^a)$. Hence $\dim \mathfrak{h} = \dim \mathfrak{g}_0^a = \operatorname{rank} \mathfrak{g}$. But because σ is an algebra automorphism, it is easy to check that $\sigma(\mathfrak{g}_0^a) = \mathfrak{g}_0^{\sigma(a)}$. The dimensionality of this tells us that $\sigma(a) \in \mathfrak{g}$ is a regular element. Finally, the last claim is immediate as conjugate subalgebras are isomorphic. □

To prepare for the proof of Chevalley's Theorem, we first prove two lemmas.

Lemma 9.3. *Let $\mathfrak{h} \subseteq \mathfrak{g}$ be a Cartan subalgebra and suppose there is a regular element $a \in \mathfrak{g}$ which is in $\mathfrak{h}$. Then $\mathfrak{h} = \mathfrak{g}_0^a$.*

Proof. Being a Cartan subalgebra, $\mathfrak{h}$ is nilpotent. Thus $\operatorname{ad} a|_{\mathfrak{h}}$ is nilpotent, and so $\mathfrak{h} \subseteq \mathfrak{g}_0^a$. Now we know from the last lecture that $\mathfrak{g}_0^a$ is a Cartan subalgebra, so in particular it is nilpotent. We also know from the last lecture that $\mathfrak{h}$, being a Cartan subalgebra, must be a maximal nilpotent subalgebra. Hence $\mathfrak{h} \subseteq \mathfrak{g}_0^a$ implies $\mathfrak{h} = \mathfrak{g}_0^a$, as desired. □

The second lemma is a special case of a result from algebraic geometry.

Lemma 9.4. *Let $f : \mathbb{F}^m \to \mathbb{F}^m$ be a polynomial map, i.e.*

$$f(x_1, \dots, x_m) = (f_1(x_1, \dots, x_m), \dots, f_m(x_1, \dots, x_m))$$

where the f_i's are polynomials. Suppose that for some $a \in \mathbb{F}^m$ the linear map $df|_{x=a} : \mathbb{F}^m \to \mathbb{F}^m$ is nonsingular. Then the image $f(\mathbb{F}^m)$ contains a nonempty Zariski open subset of $\mathbb{F}^m$.

Exercise 9.3. Prove Lemma 9.4 by the following steps.

1. Note that $\mathrm{d}f|_{x=a}$ is the linear map $\mathbb{F}^m \to \mathbb{F}^m$ given by the matrix $\left(\frac{\partial f_i}{\partial x_j}\right)_{i,j=1}^m$.
2. Suppose for contradiction that $F(f_1, \dots, f_m) = 0$ identically for some nonzero polynomial F. Show that $\det \left(\frac{\partial f_i}{\partial x_j}\right) = 0$. (Thus the polynomials f_i are algebraically independent.)
3. Given algebraically independent elements $y_1, \dots, y_m \in \mathbb{F}[x_1, \dots, x_m]$, show that the extension of fields

$$\mathbb{F}(x_1, \dots, x_m) \supseteq \mathbb{F}(y_1, \dots, y_m)$$

is finite, i.e. each x_i satisfies a nonzero polynomial equation over $\mathbb{F}(y_1, \dots, y_m)$.

4. For each $i = 1, 2, \ldots, m$, take a polynomial equation satisfied by x_i over $\mathbb{F}(f_1, \ldots, f_m)$, clear denominators to get a polynomial over $\mathbb{F}[f_1, \ldots, f_m]$, and let $p_i(f_1, \ldots, f_m)$ be the leading coefficient of this polynomial. Show that the set of points $\{y \in \mathbb{F}^m : p_i(y) \neq 0, i = 1, 2, \ldots, m\}$ satisfies Lemma 9.4.

Solution: Step 1 is completely standard. For step 2, suppose for contradiction that the f_i's are algebraically dependent, and let $F(y_1, \ldots, y_m)$ be a nonzero polynomial of minimal degree such that $F(f_1, \ldots, f_m) = 0$ identically. For any $j = 1, 2, \ldots, m$, apply the partial derivative $\partial/\partial x_j$ to the identity $F(f_1, \ldots, f_m) = 0$. By the chain rule, this gives the equation

$$\sum_{i=1}^{m} \frac{\partial F}{\partial y_i}\bigg|_{(f_1,\ldots,f_m)} \cdot \frac{\partial f_i}{\partial x_j} = 0.$$

Let $J(x)$ be the Jacobian matrix $\left(\frac{\partial f_i}{\partial x_j}\Big|_x \right)_{i,j=1}^m$ and let $\vec{v}(y)$ be the row vector $\left(\frac{\partial F}{\partial y_i}\Big|_{(y_1,\ldots,y_m)} \right)_{i=1}^m$. Then the above equations yield $\vec{v}(f(x))J(x) = 0$.

Let $\mathcal{L}$ be the locus of points x such that $\det(J(x)) \neq 0$ and let $\mathcal{L}'$ be the locus of points x such that $\vec{v}(f(x)) \neq 0$. At any point of $\mathcal{L}'$, the matrix J has a nonzero null vector, so it is singular: thus $\mathcal{L} \cap \mathcal{L}' = \emptyset$. Both $\mathcal{L}$ and $\mathcal{L}'$ are Zariski open subsets of $\mathbb{F}^m$, and $\mathcal{L}$ is nonempty by assumption. As the base field $\mathbb{F}$ is infinite, the only way for the nonempty open set $\mathcal{L}$ not to intersect the open set $\mathcal{L}'$ is to have $\mathcal{L}' = \emptyset$. But this means that the polynomials in $\vec{v}$, i.e. the partial derivatives $\partial F/\partial y_i$, are polynomials in $y_1, \cdots, y_n$ which vanish identically upon substitution of the f_i's. As $\deg(\partial F/\partial y_i) < \deg F$, our minimality assumption on $\deg F$ then implies $\partial F/\partial y_i = 0$. Over a field of characteristic zero, a polynomial whose partial derivatives all vanish identically must be constant. But F is not constant; contradiction.

For step 3, note that both $\mathbb{F}(x_1, \ldots, x_m)$ and $\mathbb{F}(y_1, \ldots, y_m)$ have transcendence degree m over $\mathbb{F}$, and an extension of fields with the same transcendence degree must be algebraic. Hence each x_i is algebraic over $\mathbb{F}(y_1, \ldots, y_m)$, and in particular they generate a finite extension.

For step 4, note that the stated set is certainly a nonempty Zariski open set, as it is the nonvanishing locus of the nonzero polynomial $p_1 \cdots p_m$. We need to show that it is in the image of f, so choose any y such that $p_1(y), \ldots, p_m(y) \neq 0$. Because the elements f_i are algebraically independent, we can define an evaluation homomorphism

$$e_y : \mathbb{F}[f_1, \ldots, f_m] \to \mathbb{F}$$

mapping each f_i to y_i. By our assumption on y, e_y maps each p_i to a nonzero field element. Thus e_y extends uniquely to a map on the localization, $\hat{e}_y : \mathbb{F}[f_1, \ldots, f_m, p_1^{-1}, \ldots, p_m^{-1}] \to \mathbb{F}$. For each i, it follows from the definition of p_i that x_i is integral over the ring $\mathbb{F}[f_1, \ldots, f_m, p_1^{-1}, \ldots, p_m^{-1}]$. We now use the following fact from commutative algebra:

Given an integral extension of rings $R \subseteq R'$ and a ring homomorphism from R to an algebraically closed field, there exists an extension of that ring homomorphism to R'.

Applying this, we get an extension of $\hat{e}_y$ to a homomorphism $\hat{\hat{e}}_y : \mathbb{F}[x_1, \ldots, x_m] \to \mathbb{F}$. Unwrapping definitions, the fact that $\hat{\hat{e}}_y$ extends e_y tells us that $f(\hat{\hat{e}}_y(x_1), \ldots, \hat{\hat{e}}_y(x_m)) = e_y(f_1, \ldots, f_m) = y$, so that indeed y is in the image of f. □

Example 9.1. Consider the map $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$. The image is $f(\mathbb{R}) = \mathbb{R}_{\geq 0}$, which does not contain a nonempty Zariski open set (because nonempty Zariski open sets in one dimension are cofinite). Thus the algebraic closure assumption in Lemma 9.4 is necessary. (The characteristic assumption is not necessary, though.)

Example 9.2. It is also important in Lemma 9.4 that the map f be algebraic. There is a well-known example of a smooth (but not algebraic) map from $\mathbb{R}$ to the torus which does not include any nonempty open set in its image.

We conclude with the proof of the main result.

Proof of Chevalley's Theorem: Let $\mathfrak{h}$ be any Cartan subalgebra of $\mathfrak{g}$. As $\mathfrak{h}$ is nilpotent, we can define a corresponding generalized root space decomposition $\mathfrak{g} = \oplus_{\alpha\in\mathfrak{h}^*}\mathfrak{g}_\alpha^{\mathfrak{h}}$. For later use, we observe that $\mathfrak{g}_0^{\mathfrak{h}} = \mathfrak{h}$ by a proposition from last lecture.

We claim that for any $\alpha \neq 0$ and any $x \in \mathfrak{g}_\alpha^{\mathfrak{h}}$, $\operatorname{ad} x$ is nilpotent. To prove this, recall that the root space decomposition satisfies $[\mathfrak{g}_\beta^{\mathfrak{h}}, \mathfrak{g}_\gamma^{\mathfrak{h}}] \subseteq \mathfrak{g}_{\beta+\gamma}^{\mathfrak{h}}$ for all β, γ. Thus for any β and any $N \geq 1$, $(\operatorname{ad} x)^N \mathfrak{g}_\beta^{\mathfrak{h}} \subseteq \mathfrak{g}_{\beta+N\alpha}^{\mathfrak{h}}$. As $\alpha \neq 0$ and $\operatorname{char} F = 0$, $\{\beta + N\alpha : N \geq 1\}$ is an infinite set of distinct functionals. There are only finitely many nonzero root spaces $\mathfrak{g}_\gamma^{\mathfrak{h}}$, so for some N we must have $(\operatorname{ad} x)^N \mathfrak{g}_\beta^{\mathfrak{h}} = 0$. This holds for all root spaces $\mathfrak{g}_\beta$. Replacing N by its maximum over the finite number of nonzero root spaces, we have $(\operatorname{ad} x)^N \mathfrak{g} = 0$, which proves the claim.

Our next goal is to show that there is a Zariski open subset of $\mathfrak{g}$ consisting of images of elements of $\mathfrak{h}$ under the action of the group G. Let $\{\alpha_i\}$ be the set of nonzero functionals α such that $\mathfrak{g}_\alpha^{\mathfrak{h}} \neq 0$, and let $\{b_j\}_{j=1}^m$ be a basis for $\oplus_i \mathfrak{g}_{\alpha_i}^{\mathfrak{h}}$ consisting of a union of bases for each $g_{\alpha_i}^{\mathfrak{h}}$. Then any element of $\mathfrak{g}$ has a unique expansion of the form $h + \sum_{j=1}^m x_j b_j$ for some $h \in \mathfrak{h}$ and some scalars $x_j \in \mathbb{F}$. Define a map $f : \mathfrak{g} \to \mathfrak{g}$ by

$$f\left(h + \sum_{j=1}^m x_j b_j\right) = e^{x_1 \operatorname{ad} b_1} e^{x_2 \operatorname{ad} b_2} \cdots e^{x_m \operatorname{ad} b_m}(h).$$

This is well-defined because each $\operatorname{ad} b_j$ is nilpotent by the result of the last paragraph. Moreover, it is a polynomial map. Looking to apply Lemma 9.4, we compute the differential $\mathrm{d}f$ at an arbitrary point $a \in \mathfrak{h}$, evaluated at a point $b + h$ (where $h \in \mathfrak{h}$ and $b = \sum x_j b_j$). This is

$$\mathrm{d}f|_{x=a}(b+h) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} [f(a + t(b+h))] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left[e^{tx_1 \operatorname{ad} b_1} e^{tx_2 \operatorname{ad} b_2} \cdots e^{tx_m \operatorname{ad} b_m}(a + th)\right].$$

To compute this derivative, it suffices to expand the function we are differentiating to first order in t. Doing so, we find

$$\mathrm{d}f|_{x=a}(b+h) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left[\left(\prod_{j=1}^m (I + tx_j \operatorname{ad} b_j)\right)(a + th)\right] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left[(a+th) + \sum_{j=1}^m tx_j(\operatorname{ad} b_j)(a)\right],$$

which equals $h + \sum_{j=1}^m x_j[b_j, a] = h + [b, a]$. Thus the linear operator $\mathrm{d}f|_{x=a}$ restricts to the identity on $\mathfrak{h}$ and to $-\operatorname{ad} a$ on $\oplus_i \mathfrak{g}_{\alpha_i}^{\mathfrak{h}}$. On each space $\mathfrak{g}_{\alpha_i}^{\mathfrak{h}}$, the only eigenvalue of $-\operatorname{ad} a$ is $-\alpha_i(a)$. Thus if we can find $a \in \mathfrak{h}$ such that $\alpha_i(a) \neq 0$ for each i, then $\mathrm{d}f|_{x=a}$ will act invertibly on each generalized root space and so it will be nonsingular as an operator on $\mathfrak{g}$.

It is easy to prove that this is possible. Indeed, for each i we can find $a_i \in \mathfrak{h}$ such that $\alpha_i(a_i) \neq 0$, just because the functional $\alpha_i \in \mathfrak{h}^*$ is nonzero. We identify the span of the a_i's with affine space via the association $f = (f_i) \mapsto a = \sum_i f_i a_i$. For each i, the condition $\alpha_i(a) \neq 0$ defines a nonempty Zariski open subset. The (finite) intersection of these over all i is then also nonempty, as desired.

We have now proven that the polynomial map f has the property that $\mathrm{d}f|_{x=a}$ is nonsingular at some point a. Then by Lemma 9.4, the image of f contains a nonempty Zariski open subset of $\mathfrak{g}$. Let $\Omega_{\mathfrak{h}}$ be such a subset. Notice that, by definition, the image of f (and thus the subset $\Omega_{\mathfrak{h}}$) consists of points of the form $\sigma(h)$ for some $\sigma \in G$ and $h \in \mathfrak{h}$.

The arguments so far hold for any Cartan subalgebra of $\mathfrak{g}$. Thus, letting $\mathfrak{h}_1$ and $\mathfrak{h}_2$ be two arbitrary Cartan subalgebras, we obtain corresponding nonempty Zariski open subsets $\Omega_{\mathfrak{h}_1}$ and $\Omega_{\mathfrak{h}_2}$ of $\mathfrak{g}$. Let Ω_r be the subset of all regular elements of $\mathfrak{g}$; this is also a nonempty Zariski open subset. The intersection $\Omega_{\mathfrak{h}_1} \cap \Omega \mathfrak{h}_2 \cap \Omega_r$ is then nonempty. Rephrasing this, there exists a regular element $x \in \mathfrak{g}$, elements $h_1 \in \mathfrak{h}_1$ and $h_2 \in \mathfrak{h}_2$, and automorphisms $\sigma_1, \sigma_2 \in G$ such that $\sigma_1(h_1) = x = \sigma_2(h_2)$. As x is regular and σ_1 is an algebra automorphism of $\mathfrak{g}$, $h_1 = \sigma_1^{-1}(x)$ is also regular. Thus by Lemma 9.3, $\mathfrak{h}_1 = \mathfrak{g}_0^{h_1}$. Similarly, $\mathfrak{h}_2 = \mathfrak{g}_0^{h_2}$. The automorphism $\sigma = \sigma_2^{-1}\sigma_1 \in G$ maps h_1 to h_2, and so (as in the proof of Corollary 9.2)

$$\sigma(\mathfrak{h}_1) = \sigma(\mathfrak{g}_0^{h_1}) = \mathfrak{g}_0^{\sigma(h_1)} = \mathfrak{g}_0^{h_2} = \mathfrak{h}_2.$$

This finishes the proof of Chevalley's Theorem. □