Math 126 Lecture 10

Conjugacy of Cartan subalgebras.

Motivation from the unitary groups.

It is a standard theorem in linear algebra that any unitary matrix can be diagonalized (by conjugation by unitary matrices). On the other hand, it is easy to check that the subgroup $T \subset U(n)$ consisting of all unitary matrices is a maximal commutative subgroup: any matrix which commutes with all diagonal unitary matrices must itself be diagonal; indeed if A is a diagonal matrix with distinct entries along the diagonal, any matrix which commutes with A must be diagonal. Notice that T is a product of circles, i.e. a torus.

Motivation from compact Lie groups.

This theorem has an immediate generalization to compact Lie groups: Let G be a compact Lie group, and let T and T' be two maximal tori. (So T and T' are connected commutative subgroups (hence necessarily tori) and each is not strictly contained in a larger connected commutative subgroup). Then there exists an element $a \in G$ such that $aT'a^{-1} = T$. To prove this, choose one parameter subgroups of T and T' which are dense in each. That is, choose x and x' in the Lie algebra $\mathbf{g}$ of G such that the curve $t \mapsto \exp tx$ is dense in T and the curve $t \mapsto \exp tx'$ is dense in T'. If we could find $a \in G$ such that the

$$a(\exp tx')a^{-1} = \exp t \operatorname{Ad}_a x'$$

commute with all the $\exp sx$, then $a(\exp tx')a^{-1}$ would commute with all elements of T, hence belong to T, and by continuity, $aT'a^{-1} \subset T$ and hence $= T$. So we would like to find and $a \in G$ such that

$$[Ad_a x', x] = 0.$$

Proof via a variational principle.

Put a positive definite scalar product $(\ , \ )$ on $\mathbf{g}$, the Lie algebra of G which is invariant under the adjoint action of G. This is always possible by choosing any positive definite scalar product and then averaging it over G.

Choose $a \in G$ such that $(\mathrm{Ad}_a x', x)$ is a maximum. Let

$$y := \mathrm{Ad}_a x'.$$

We wish to show that

$$[y, x] = 0.$$

For any $z \in \mathbf{g}$ we have

$$([z, y], x) = \frac{d}{dt} (\mathrm{Ad}_{\exp tz}\, y, x)_{|t=0} = 0$$

by the maximality. But

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

by the invariance of $(\ , \ )$, hence $[y, x]$ is orthogonal to all $\mathbf{g}$ hence 0.

Goal

We want to give an algebraic proof of the analogue of this theorem for Lie algebras over the complex numbers. In contrast to the elementary proof given above for compact groups, the proof in the general Lie algebra case will be quite involved, and the flavor of the proof will by quite different for the solvable and semi-simple cases. Nevertheless, some of the ingredients of the above proof (choosing "generic elements" analogous to the choice of x and x' for example) will make their appearance. The proofs in this chapter follow Humphreys.

Derivations.

Let δ be a derivation of the Lie algebra $\mathbf{g}$. this means that

$$\delta([y,z]) = [\delta(y), z] + [y, \delta(z)] \quad \forall\ y, z \in \mathbf{g}.$$

Then, for $a, b \in \mathbf{C}$

$$\begin{array}{rcl}
(\delta - a - b)[y,z] & = & [(\delta - a)y, z] + [y, (\delta - b)z] \\
(\delta - a - b)^2[y,z] & = & [(\delta - a)^2 y, z] + 2[(\delta - a)y, (\delta - b)z] + [y, (\delta - b)^2 z \\
(\delta - a - b)^3[y,z] & = & [(\delta - a)^3 y, z] + 3[(\delta - a)^2 y, (\delta - b)^z)] + \\
 & & \quad 3[(\delta - a)y, (\delta - b)^2 z] + [y, (\delta - b)^3 z] \\
\vdots & & \vdots \\
(\delta - a - b)^n[y,z] & = & \sum \left(\begin{array}{c} n \\ k \end{array} \right) [(\delta - a)^k y, (\delta - b)^{n-k} z].
\end{array}$$

Consequences of

$$(\delta - a - b)^n[y, z] \;=\; \sum \binom{n}{k} [(\delta - a)^k y, (\delta - b)^{n-k} z].$$

- Let $\mathbf{g}_a = \mathbf{g}_a(\delta)$ denote the generalized eigenspace corresponding to the eigenvalue a, so $(\delta - a)^k = 0$ on $\mathbf{g}_a$ for large enough k. Then
$$[\mathbf{g}_a, \mathbf{g}_b] \subset \mathbf{g}_{[a+b]}. \tag{1}$$

- Let $s = s(\delta)$ denote the diagonizable (semi-simple) part of δ, so that $s(\delta) = a$ on $\mathbf{g}_a$. Then, for $y \in \mathbf{g}_a, z \in \mathbf{g}_b$
$$s(\delta)[y, z] = (a + b)[y, z] = [s(\delta)y, z] + [y, s(\delta)z]$$
so s and hence also $n = n(\delta)$, the nilpotent part of δ are both derivations.

- $[\delta, \operatorname{ad} x] = \operatorname{ad}(\delta x)]$. Indeed, $[\delta, \operatorname{ad} x](u) = \delta([x, u]) - [x, \delta(u)] = [\delta(x), u]$. In particular, the space of inner derivations, $\operatorname{Inn} \mathbf{g}$ is an ideal in $\operatorname{Der} \mathbf{g}$.

Derivations of semi-simple Lie algebras.

- If $\mathbf{g}$ is semisimple then Inn $\mathbf{g}$ = Der $\mathbf{g}$. Indeed, split off an invariant complement to Inn $\mathbf{g}$ in Der $\mathbf{g}$ (possible by Weyl's theorem on complete reducibility). For any δ in this invariant complement, we must have $[\delta, \mathrm{ad}\, x] = 0$ since $[\delta, \mathrm{ad}\, x] = \mathrm{ad}\, \delta x$. This says that δx is in the center of $\mathbf{g}$. Hence $\delta x = 0 \ \forall x$ hence $\delta = 0$.

- Hence any $x \in \mathbf{g}$ can be uniquely written as $x = s + n$, $s \in \mathbf{g}$, $n \in \mathbf{g}$ where $\mathrm{ad}\, s$ is semisimple and $\mathrm{ad}\, n$ is nilpotent. This is known as the decomposition into semi-simple and nilpotent parts for a semi-simple Lie algebra.

- (Back to general $\mathbf{g}$.) Let $\mathbf{k}$ be a subalgebra containing $\mathbf{g}_0(\operatorname{ad} x)$ for some $x \in \mathbf{g}$. Then x belongs $\mathbf{g}_0(\operatorname{ad} x)$ hence to $\mathbf{k}$, hence $\operatorname{ad} x$ preserves $N_{\mathbf{g}}(\mathbf{k})$ (by Jacobi's identity). We have

$$x \in \mathbf{g}_0(\operatorname{ad} x) \subset \mathbf{k} \subset N_{\mathbf{g}}(\mathbf{k}) \subset \mathbf{g}$$

  all of these subsepaces being invariant under $\operatorname{ad} x$. Therefore, the characteristic polynomial of $\operatorname{ad} x$ restricted to $N_{\mathbf{g}}(\mathbf{k})$ is a factor of the charactristic polyinomial of $\operatorname{ad} x$ acting on $\mathbf{g}$. But all the zeros of this characteristic polynomial are accounted for by the generalized zero eigenspace $\mathbf{g}_0(\operatorname{ad} x)$ which is a subspace of $\mathbf{k}$. This means that $\operatorname{ad} x$ acts on $N_{\mathbf{g}}(\mathbf{k})/\mathbf{k}$ without zero eigenvalue.

  On the other hand, $\operatorname{ad} x$ acts trivially on this quotient space since $x \in \mathbf{k}$ and hence $[N_{\mathbf{g}}\mathbf{k}, x] \subset \mathbf{k}$ by the definition of the normalizer. Hence

$$N_{\mathbf{g}}(\mathbf{k}) = \mathbf{k}. \qquad (2)$$

A key lemma.

Lemma 1 *Let* $\mathbf{k} \subset \mathbf{g}$ *be a subalgebra. Let* $z \in \mathbf{k}$ *be such that* $\mathbf{g}_0(\operatorname{ad} z)$ *does not strictly contain any* $\mathbf{g}_0(\operatorname{ad} x)$, $x \in \mathbf{k}$. *Suppose that*

$$\mathbf{k} \subset \mathbf{g}_0(\operatorname{ad} z).$$

Then

$$\mathbf{g}_0(\operatorname{ad} z) \subset \mathbf{g}_0(\operatorname{ad} y) \quad \forall\ y \in \mathbf{k}.$$

Proof. Choose z as in the lemma, and let x be an arbitrary element of $\mathbf{k}$. By hypothesis, $x \in \mathbf{g}_0(\operatorname{ad} z)$ and we know that $[\mathbf{g}_0(\operatorname{ad} z), \mathbf{g}_0(\operatorname{ad} z)] \subset \mathbf{g}_0(\operatorname{ad} z)$. Therefore $[x, \mathbf{g}_0(\operatorname{ad} z)] \subset \mathbf{g}_0(\operatorname{ad} z)$ and hence

$$\operatorname{ad}(z + cx)\mathbf{g}_0(\operatorname{ad} z) \subset \mathbf{g}_0(\operatorname{ad} z)$$

for all constants c. Thus $\mathbf{g}_0(\operatorname{ad}(z + cx))$ acts on the quotient space $\mathbf{g}/\mathbf{g}_0(\operatorname{ad} z)$. We can factor the characteristic polynomial of $\operatorname{ad}(z + cx)$ acting on $\mathbf{g}$ as

$$P_{\operatorname{ad}(z+cx)}(T) = f(T, c)g(T, c)$$

We can factor the characteristic polynomial of $\operatorname{ad}(z+cx)$ acting on $\mathbf{g}$ as

$$P_{\operatorname{ad}(z+cx)}(T) = f(T,c)g(T,c)$$

where f is the characteristic polynomial of $\operatorname{ad}(z+cx)$ on $\mathbf{g}_0(\operatorname{ad} z)$ and g is the characteristic polynomial of $\operatorname{ad}(z+cx)$ on $\mathbf{g}/\mathbf{g}_0(\operatorname{ad} z)$. Write

$$\begin{aligned} f(T,c) &= T^r + f_1(c)T^{r-1} + \cdots f_r(c) \quad r = \dim \mathbf{g}_0(\operatorname{ad} z) \\ g(T,c) &= T^{n-r} + g_1(c)T^{n-r-1} + \cdots + g_{n-r}(c) \qquad n = \dim \mathbf{g}. \end{aligned}$$

The f_i and the g_i are polynomials of degree at most i in c. Since 0 is not an eigenvalue of $\operatorname{ad} z$ on $\mathbf{g}/\mathbf{g}_0(\operatorname{ad} z)$, we see that $g_{n-r}(0) \neq 0$. So we can find $r+1$ values of c for which $g_{n-r}(c) \neq 0$, and hence for these values,

$$\mathbf{g}_0(\operatorname{ad}(z+cx)) \subset \mathbf{g}_0(\operatorname{ad} z).$$

By the minimality, this forces

$$\mathbf{g}_0(\operatorname{ad}(z+cx)) = \mathbf{g}_0(\operatorname{ad} z)$$

for these values of c.

$$P_{\mathrm{ad}(z+cx)}(T) = f(T,c)g(T,c)$$

where f is the characteristic polynomial of $\mathrm{ad}(z+cx)$ on $\mathbf{g}_0(\mathrm{ad}\, z)$ and g is the characteristic polynomial of $\mathrm{ad}(z+cx)$ on $\mathbf{g}/\mathbf{g}_0(\mathrm{ad}\, z)$. Write

$$\begin{aligned} f(T,c) &= T^r + f_1(c)T^{r-1} + \cdots f_r(c) \quad r = \dim \mathbf{g}_0(\mathrm{ad}\, z) \\ g(T,c) &= T^{n-r} + g_1(c)T^{n-r-1} + \cdots + g_{n-r}(c) \quad n = \dim \mathbf{g}. \end{aligned}$$

The f_i and the g_i are polynomials of degree at most i in c.

we can find $r+1$ values of c for which $\mathbf{g}_0(\mathrm{ad}(z+cx)) = \mathbf{g}_0(\mathrm{ad}\, z)$

This means that $f(T,c) = T^r$ for these values of

c, so each of the polynomials $f_1, \ldots, f_r$ has $r+1$ distinct roots, and hence is identically zero. Hence

$$\mathbf{g}_0(\mathrm{ad}(z+cx)) \supset \mathbf{g}_0(\mathrm{ad}\, z)$$

for all c. Take $c = 1, x = y - z$ to conclude the truth of the lemma.

Cartan subalgebras.

A Cartan subalgebra (**CSA**) is defined to be a nilpotent subalgebra which is its own normalizer. A Borel subalgebra (**BSA**) is defined to be a maximal solvable subalgebra. The goal is to prove

Theorem 1 *Any two* **CSA***'s are conjugate. Any two* **BSA***'s are conjugate.*

Conjugacy.

Here the word **conjugate** means the following: Define

$$\mathcal{N}(\mathbf{g}) = \{x|\ \exists y \in \mathbf{g},\ a \neq 0,\ \text{with } x \in \mathbf{g}_a(\operatorname{ad} y)\}.$$

Notice that every element of $\mathcal{N}(\mathbf{g})$ is nilpotent and that $\mathcal{N}(\mathbf{g})$ is stable under $\operatorname{Aut}(\mathbf{g})$. As any $x \in \mathcal{N}(\mathbf{g})$ is nilpotent, $\exp \operatorname{ad} x$ is well defined as an automorphism of $\mathbf{g}$, and we let

$$\mathcal{E}(\mathbf{g})$$

denote the group generated by these elements. It is a normal subgroup of the group of automorphisms. Conjugacy means that there is a $\phi \in \mathcal{E}(\mathbf{g})$ with $\phi(\mathbf{h}_1) = \mathbf{h}_2$ where $\mathbf{h}_1$ and $\mathbf{h}_2$ are **CSA**'s. Similarly for **BSA**'s.

An alternative definition of a **CSA**.

Recall: Let $\mathbf{k}$ be a subalgebra containing $\mathbf{g}_0(\operatorname{ad} x)$ then

$$N_{\mathbf{g}}(\mathbf{k}) = \mathbf{k}. \qquad (2)$$

Proposition 1 $\mathbf{h}$ *is a* **CSA** *if and only if* $\mathbf{h} = \mathbf{g}_0(\operatorname{ad} z)$ *where* $\mathbf{g}_0(\operatorname{ad} z)$ *contains no proper subalgebra of the form* $\mathbf{g}_0(\operatorname{ad} x)$.

Proof. Suppose $\mathbf{h} = \mathbf{g}_0(\operatorname{ad} z)$ which is minimal in the sense of the proposition. Then we know by (2) that $\mathbf{h}$ is its own normalizer. Also, by the lemma, $\mathbf{h} \subset \mathbf{g}_0(\operatorname{ad} x)\ \forall x \in \mathbf{h}$. Hence $\operatorname{ad} x$ acts nilpotently on $\mathbf{h}$ for all $x \in \mathbf{h}$. Hence, by Engel's theorem, $\mathbf{h}$ is nilpotent and hence is a **CSA**.

Suppose that $\mathbf{h}$ is a **CSA**. Since $\mathbf{h}$ is nilpotent, we have $\mathbf{h} \subset \mathbf{g}_0(\operatorname{ad} x)$, $\forall\, x \in \mathbf{h}$. Choose a minimal z. By the lemma,

$$\mathbf{g}_0(\operatorname{ad} z) \subset \mathbf{g}_0(\operatorname{ad} x) \quad \forall x \in \mathbf{h}.$$

Thus $\mathbf{h}$ acts nilpotently on $\mathbf{g}_0(\operatorname{ad} z)/\mathbf{h}$. If this space were not zero, we could find a non-zero common eigenvector with eigenvalue zero by Engel's theorem. This means that there is a $y \notin \mathbf{h}$ with $[y, \mathbf{h}] \subset \mathbf{h}$ contradicting the fact $\mathbf{h}$ is its own normalizer. QED

The image of a **CSA** under a surjection is a **CSA**.

Lemma 2 *If* $\phi : \mathbf{g} \to \mathbf{g}'$ *is a surjective homomorphism and* $\mathbf{h}$ *is a* **CSA** *of* $\mathbf{g}$ *then* $\phi(\mathbf{h})$ *is a* **CSA** *of* $\mathbf{g}'$.

Clearly $\phi(\mathbf{h})$ is nilpotent. Let $\mathbf{k} = \mathrm{Ker}\ \phi$ and identify $\mathbf{g}' = \mathbf{g}/\mathbf{k}$ so $\phi(\mathbf{h}) = \mathbf{h}+\mathbf{k}$. If $x+\mathbf{k}$ normalizes $\mathbf{h}+\mathbf{k}$ then x normalizes $\mathbf{h}+\mathbf{k}$. But $\mathbf{h} = \mathbf{g}_0(\mathrm{ad}\, z)$ for some minimal such z, and as an algebra containing a $\mathbf{g}_0(\mathrm{ad}\, z)$, $\mathbf{h}+\mathbf{k}$ is self-normalizing. So $x \in \mathbf{h}+\mathbf{k}$. QED

The inverse image of a **CSA**.

Lemma 3 *$\phi : \mathbf{g} \to \mathbf{g}'$ be surjective, as above, and $\mathbf{h}'$ a* **CSA** *of $\mathbf{g}'$. Any* **CSA** $\mathbf{h}$ *of $\mathbf{m} := \phi^{-1}(\mathbf{h}')$ is a* **CSA** *of $\mathbf{g}$.*

$\mathbf{h}$ is nilpotent by assumption. We must show it is its own normalizer in $\mathbf{g}$. By the preceding lemma, $\phi(\mathbf{h})$ is a Cartan subalgebra of $\mathbf{h}'$. But $\phi(\mathbf{h})$ is nilpotent and hence would have a common eigenvector with eigenvalue zero in $\mathbf{h}'/\phi(\mathbf{h})$, contradicting the selfnormalizing property of $\phi(\mathbf{h})$ unless $\phi(\mathbf{h}) = \mathbf{h}'$. So $\phi(\mathbf{h}) = \mathbf{h}'$. If $x \in \mathbf{g}$ normalizes $\mathbf{h}$, then $\phi(x)$ normalizes $\mathbf{h}'$. Hence $\phi(x) \in \mathbf{h}'$ so $x \in \mathbf{m}$ so $x \in \mathbf{h}$. QED

Conjugacy of **CSA**s in the solvable case.

In this case a Borel subalgebra is all of $\mathbf{g}$ so we must prove conjugacy for **CSA**'s. In case $\mathbf{g}$ is nilpotent, we know that any **CSA** is all of $\mathbf{g}$, since $\mathbf{g} = \mathbf{g}_0(\operatorname{ad} z)$ for any $z \in \mathbf{g}$. So we may proceed by induction on dim $\mathbf{g}$. Let $\mathbf{h_1}$ and $\mathbf{h_2}$ be Cartan subalgebras of $\mathbf{g}$. We want to show that they are conjugate. Choose an abelian ideal $\mathbf{a}$ of smallest possible positive dimension and let $\mathbf{g}' = \mathbf{g}/\mathbf{a}$. By Lemma 2 the images $\mathbf{h}'_1$ and $\mathbf{h}'_2$ of $\mathbf{h_1}$ and $\mathbf{h_2}$ in $\mathbf{g}'$ are **CSA**'s of $\mathbf{g}'$ and hence there is a $\sigma' \in \mathcal{E}(\mathbf{g}')$ with $\sigma'(\mathbf{h}'_1) = \mathbf{h}'_2$. We claim that we can lift this to a $\sigma \in \mathcal{E}(\mathbf{g})$. That is, we claim

Recall that $\mathcal{N}(\mathbf{g}) = \{x \mid \exists y \in \mathbf{g},\ a \neq 0,\ \text{with } x \in \mathbf{g}_a(\operatorname{ad} y)\}$
As any $x \in \mathcal{N}(\mathbf{g})$ is nilpotent, $\exp \operatorname{ad} x$ is well defined as an automorphism of $\mathbf{g}$, and we let

$$\mathcal{E}(\mathbf{g})$$

denote the group generated by these elements.

Lemma 4 *Let $\phi : \mathbf{g} \to \mathbf{g}'$ be a surjective homomorphism. If $\sigma' \in \mathcal{E}(\mathbf{g}')$ then there exists a $\sigma \in \mathcal{E}(\mathbf{g})$ such that the diagram*

$$\begin{array}{ccc} \mathbf{g} & \xrightarrow{\phi} & \mathbf{g}' \\ \sigma \downarrow & & \downarrow \sigma' \\ \mathbf{g} & \xrightarrow[\phi]{} & \mathbf{g}' \end{array}$$

commutes.

Proof of lemma. It is enough to prove this on generators. Suppose that $x' \in \mathbf{g}_a(y')$ and choose $y \in \mathbf{g}$, $\phi(y) = y'$ so $\phi(\mathbf{g}_a(y)) = \mathbf{g}_a(y')$, and hence we can find an $x \in \mathcal{N}(\mathbf{g})$ mapping on to x'. Then $\exp \operatorname{ad} x$ is the desired σ in the above diagram if $\sigma' = \exp \operatorname{ad} x'$. QED

Back to the proof of the conjugacy of **CSA**s in the solvable case by induction on the dimension.

Choose an abelian ideal $\mathbf{a}$ of smallest possible positive dimension and let $\mathbf{g}' = \mathbf{g}/\mathbf{a}$. By Lemma 2 the images $\mathbf{h}_1'$ and $\mathbf{h}_2'$ of $\mathbf{h_1}$ and $\mathbf{h_2}$ in $\mathbf{g}'$ are **CSA**'s of $\mathbf{g}'$ and hence there is a $\sigma' \in \mathcal{E}(\mathbf{g}')$ with $\sigma'(\mathbf{h}_1') = \mathbf{h}_2'$. we can lift this to a $\sigma \in \mathcal{E}(\mathbf{g})$.

Let $\mathbf{m}_1 := \phi^{-1}(\mathbf{h}_1')$, $\mathbf{m}_2 := \phi^{-1}(\mathbf{h}_2')$. We have a σ with $\sigma(\mathbf{m}_1) = \mathbf{m}_2$ so $\sigma(\mathbf{h}_1)$ and $\mathbf{h}_2$ are both **CSA**'s of $\mathbf{m}_2$. If $\mathbf{m}_2 \neq \mathbf{g}$ we are done by induction. So the one new case is where

$$\mathbf{g} = \mathbf{a} + \mathbf{h}_1 = \mathbf{a} + \mathbf{h}_2.$$

Write

$$\mathbf{h}_2 = \mathbf{g}_0(\operatorname{ad} x)$$

for some $x \in \mathbf{g}$. Since $\mathbf{a}$ is an ideal, it is stable under $\operatorname{ad} x$ and we can split it into its 0 and non-zero generalized eigenspaces:

$$\mathbf{a} = \mathbf{a}_0(\operatorname{ad} x) \oplus \mathbf{a}_*(\operatorname{ad} x).$$

$$\mathbf{g} = \mathbf{a} + \mathbf{h}_1 = \mathbf{a} + \mathbf{h}_2$$

$$\mathbf{h}_2 = \mathbf{g}_0(\operatorname{ad} x)$$

$$\mathbf{a} = \mathbf{a}_0(\operatorname{ad} x) \oplus \mathbf{a}_*(\operatorname{ad} x).$$

Since $\mathbf{a}$ is abelian, ad of every element of $\mathbf{a}$ acts trivially on each summand, and since $\mathbf{h}_2 = \mathbf{g}_0(\operatorname{ad} x)$ and $\mathbf{a}$ is an ideal, this decomposition is stable under $\mathbf{h}_2$, hence under all of $\mathbf{g}$. By our choice of $\mathbf{a}$ as a minimal abelian ideal, one or the other of these summands must vanish. If $\mathbf{a} = \mathbf{a}_0(\operatorname{ad} x)$ we would have $\mathbf{a} \subset \mathbf{h}_2$ so $\mathbf{g} = \mathbf{h}_2$ and $\mathbf{g}$ is nilpotent. There is nothing to prove. So the only case to consider is $\mathbf{a} = \mathbf{a}_*(\operatorname{ad} x)$. Since $\mathbf{h}_2 \subset \mathbf{g}_0(\operatorname{ad} x)$ we have

$$\mathbf{a} = \mathbf{g}_*(\operatorname{ad} x).$$

We are reduced to the case $\mathbf{g} = \mathbf{a} + \mathbf{h}_1 = \mathbf{a} + \mathbf{h}_2.$

$$\mathbf{a} = \mathbf{g}_*(\operatorname{ad} x).$$

$$\mathbf{h}_2 = \mathbf{g}_0(\operatorname{ad} x)$$

Since $\mathbf{g} = \mathbf{h}_1 + \mathbf{a}$, write

$$x = y + z, \quad y \in \mathbf{h}_1, \quad z \in \mathbf{g}_*(\operatorname{ad} x).$$

Since $\operatorname{ad} x$ is invertible on $\mathbf{g}_*(\operatorname{ad} x)$, write $z = [x, z'], \quad z' \in \mathbf{a}_*(\operatorname{ad} x)$. Since $\mathbf{a}$ is an abelian ideal, $(\operatorname{ad} z')^2 = 0$, so $\exp(\operatorname{ad} z') = 1 + \operatorname{ad} z'$. So

$$\exp(\operatorname{ad} z')x = x - z = y.$$

So $\mathbf{h} := \mathbf{g}_0(\operatorname{ad} y)$ is a **CSA** (of $\mathbf{g}$), and since $y \in \mathbf{h}_1$ we have $\mathbf{h}_1 \subset \mathbf{g}_0(\operatorname{ad} y) = \mathbf{h}$ and hence $\mathbf{h}_1 = \mathbf{h}$. So $\exp \operatorname{ad} z'$ conjugates $\mathbf{h}_2$ into $\mathbf{h}_1$. Writing z' as sum of its generalized eigencomponents, and using the fact that all the elements of $\mathbf{a}$ commute, we can write the exponential as a product of the exponentials of the summands. QED

Toral subalgebras.

The strategy is now to show that any two **BSA**'s of an arbitrary Lie algebra are conjugate, thus reducing the proof of the conjugacy theorem for **CSA**'s to that of **BSA**'s. Since the radical is contained in any **BSA**, it is enough to prove this theorem for semi-simple Lie algebras. So for this section the Lie algebra $\mathbf{g}$ will be assumed to be semi-simple.

Since $\mathbf{g}$ does not consist entirely of ad nilpotent elements, it contains some x which is not ad nilpotent, and the semi-simple part of x is a non-zero ad semi-simple element of $\mathbf{g}$. A subalgebra consisting entirely of semi-simple elements is called **toral**, for example, the line through x_s.

Lemma 5 *Any toral subalgebra* $\mathbf{t}$ *is abelian.*

Proof. The elements $\operatorname{ad} x,\ x \in \mathbf{t}$ can be each be diagonalized. We must show that $\operatorname{ad} x$ has no eigenvectors with non-zero eigenvalues in $\mathbf{t}$. Let y be an eigenvector so $[x, y] = ay$. Then $(\operatorname{ad} y)x = -ay$ is a zero eigenvector of $\operatorname{ad} y$, which is impossible unless $ay = 0$, since $\operatorname{ad} y$ annihilates all its zero eigenvectors and is invertible on the subspace spanned by the eigenvectors corresponding to non-zero eigenvalues.

Toral subalgebras and Cartan subalgebras.

One of the consequences of the considerations in this section will be:

Theorem 2 *A subalgebra* $\mathbf{h}$ *of a semi-simple Lie algebra* $\mathbf{g}$*is a* **CSA** *if and only if it is a maximal toral subalgebra.*

Roots.

To prove this we want to develop some of the theory of roots. So fix a maximal toral subalgebra $\mathbf{h}$. Decompose $\mathbf{g}$ into simultaneous eigenspaces

$$\mathbf{g} = C_{\mathbf{g}}(\mathbf{h}) \bigoplus \mathbf{g}_\alpha(\mathbf{h})$$

where

$$C_{\mathbf{g}}(\mathbf{h}) := \{x \in \mathbf{g} | [h, x] = 0 \ \forall \ h \in \mathbf{h}\}$$

is the centralizer of $\mathbf{h}$, where α ranges over non-zero linear functions on $\mathbf{h}$ and

$$\mathbf{g}_\alpha(\mathbf{h}) := \{x \in \mathbf{g} | [h, x] = \alpha(h)x \quad \forall \ h \in \mathbf{h}\}.$$

As $\mathbf{h}$ will be fixed for most of the discussion, we will drop the $(\mathbf{h})$ and write

$$\mathbf{g} = \mathbf{g}_0 \oplus \bigoplus \mathbf{g}_\alpha$$

where $\mathbf{g}_0 = C_{\mathbf{g}}(\mathbf{h})$.

Facts about roots.

$$\mathbf{g} = \mathbf{g}_0 \oplus \bigoplus \mathbf{g}_\alpha$$

- $[\mathbf{g}_\alpha, \mathbf{g}_\beta] \subset \mathbf{g}_{\alpha+\beta}$ (by Jacobi) so
- $\mathrm{ad}\, x$ is nilpotent if $x \in \mathbf{g}_\alpha,\ \alpha \neq 0$
- If $\alpha + \beta \neq 0$ then $\kappa(x, y) = 0\ \forall\ x \in \mathbf{g}_\alpha,\ y \in \mathbf{g}_\beta$.

The last item follows by choosing an $h \in \mathbf{h}$ with $\alpha(h) + \beta(h) \neq 0$. Then $0 = \kappa([h, x], y) + \kappa(x, [h, y]) = (\alpha(h) + \beta(h))\kappa(x, y)$ so $\kappa(x, y) = 0$. This implies that $\mathbf{g}_0$ is orthogonal to all the $\mathbf{g}_\alpha,\ \alpha \neq 0$ and hence the non-degeneracy of κ implies that

Proposition 2 *The restriction of κ to $\mathbf{g}_0 \times \mathbf{g}_0$ is non-degenerate.*

Toral subalgebras and **CSA**s.

Our next intermediate step is to prove:

Proposition 3

$$\mathbf{h} = \mathbf{g}_0 \tag{3}$$

if $\mathbf{h}$ *is a maximal toral subalgebra.*

Proceed according to the following steps:

$$x \in \mathbf{g}_0 \;\Rightarrow\; x_s \in \mathbf{g}_0 \;\; x_n \in \; \mathbf{g}_0. \tag{4}$$

Indeed, $x \in \mathbf{g}_0 \;\Leftrightarrow \operatorname{ad} x : \mathbf{h} \to 0$, and then $\operatorname{ad} x_s$, $\operatorname{ad} x_n$ also map $\mathbf{h} \to 0$.

$$x \in \mathbf{g}_0,\; x \text{ semisimple } \Rightarrow x \in \mathbf{h}. \tag{5}$$

Indeed, such an x commutes with all of $\mathbf{h}$. As the sum of commuting semi-simple transformations is again semisimple, we conclude that $\mathbf{h} + \mathbf{C}x$ is a toral subalgebra. By maximality it must coincide with $\mathbf{h}$.

We are proving

Proposition 3

$$\mathbf{h} = \mathbf{g}_0 \tag{3}$$

if $\mathbf{h}$ *is a maximal toral subalgebra.*

Lemma 6 *The restriction of the Killing form* κ *to* $\mathbf{h} \times \mathbf{h}$ *is non-degenerate.*

So suppose that $\kappa(h, x) = 0 \ \forall \ x \in \mathbf{h}$. This means that $\kappa(h, x) = 0 \ \forall$ semi-simple $x \in \mathbf{g}_0$. Suppose that $n \in \mathbf{g}_0$ is nilpotent. Since h commutes with n, $(\operatorname{ad} h)(\operatorname{ad} n)$ is again nilpotent. Hence has trace zero. Hence $\kappa(h, n) = 0$, and therefore $\kappa(h, x) = 0 \ \forall \ x \in \mathbf{g}_0$. Hence $h = 0$. QED

We are proving

Proposition 3

$$\mathbf{h} = \mathbf{g}_0 \tag{3}$$

if $\mathbf{h}$ *is a maximal toral subalgebra.*

Lemma 7 $\mathbf{g}_0$ *is a nilpotent Lie algebra.*

Indeed, all semi-simple elements of $\mathbf{g}_0$ commute with all of $\mathbf{g}_0$ since they belong to $\mathbf{h}$, and a nilpotent element is ad nilpotent on all of $\mathbf{g}$ so certainly on $\mathbf{h}$. Finally any $x \in \mathbf{g}_0$ can be written as a sum $x_s + x_n$ of commuting elements which are ad nilpotent on $\mathbf{g}_0$, hence x is. Thus $\mathbf{g}_0$ consists entirely of ad nilpotent elements and hence is nilpotent by Engel's theorem. QED

We are proving

Proposition 3

$$\mathbf{h} = \mathbf{g}_0 \tag{3}$$

if $\mathbf{h}$ *is a maximal toral subalgebra.*

Now suppose that $h \in \mathbf{h}$, $x, y \in \mathbf{g}_0$. Then

$$\begin{aligned} \kappa(h, [x, y]) &= \kappa([h, x], y) \\ &= \kappa(0, y) \\ &= 0 \end{aligned}$$

and hence, by the non-degeneracy of κ on $\mathbf{h}$, we conclude that

Lemma 8

$$\mathbf{h} \cap [\mathbf{g}_0, \mathbf{g}_0] = 0.$$

Completion of the proof of

Proposition 3

$$\mathbf{h} = \mathbf{g}_0 \tag{3}$$

if $\mathbf{h}$ *is a maximal toral subalgebra.*

Lemma 9 $\mathbf{g}_0$ *is abelian.*

Suppose that $[\mathbf{g}_0, \mathbf{g}_0] \neq 0$. Since $\mathbf{g}_0$ is nilpotent, it has a non-zero center contained in $[\mathbf{g}_0, \mathbf{g}_0]$. Choose a non-zero element $z \in [\mathbf{g}_0, \mathbf{g}_0]$ in this center. It can not be semi-simple for then it would lie in $\mathbf{h}$. So it has a non-zero nilpotent part, n, which also must lie in the center of $\mathbf{g}_0$, by the $B \subset A$ theorem we proved in our section on linear algebra. But then $\operatorname{ad} n \ \operatorname{ad} x$ is nilpotent for any $x \in \mathbf{g}_0$ since $[x, n] = 0$. This implies that $\kappa(n, \mathbf{g}_0) = 0$ which is impossible. QED

Completion of proof of (3). We know that $\mathbf{g}_0$ is abelian. But then, if $\mathbf{h} \neq \mathbf{g}_0$, we would find a non-zero nilpotent element in $\mathbf{g}_0$ which commutes with all of $\mathbf{g}_0$ (proven to be commutative). Hence $\kappa(n, \mathbf{g}_0) = 0$ which is impossible. This completes the proof of (3). QED

Proof of:

Theorem 2 *A subalgebra* $\mathbf{h}$ *of a semi-simple Lie algebra* $\mathbf{g}$*is a* **CSA** *if and only if it is a maximal toral subalgebra.*

So we have the decomposition

$$\mathbf{g} = \mathbf{h} \oplus \bigoplus_{\alpha \neq 0} \mathbf{g}_\alpha$$

which shows that any maximal toral subalgebra $\mathbf{h}$ is a **CSA**.

Conversely, suppose that $\mathbf{h}$ is a **CSA**. For any $x = x_s + x_n \in \mathbf{g}$, $\mathbf{g}_0(\operatorname{ad} x_s) \subset \mathbf{g}_0(\operatorname{ad} x)$ since x_n is an ad nilpotent element commuting with $\operatorname{ad} x_s$. If we choose $x \in H$ minimal so that $\mathbf{h} = \mathbf{g}_0(\operatorname{ad} x)$, we see that we may replace x by x_s and write $\mathbf{h} = \mathbf{g}_0(\operatorname{ad} x_s)$. But $\mathbf{g}_0(\operatorname{ad} x_s)$ contains some maximal toral algebra containing x_s, which is then a Cartan subalgebra contained in $\mathbf{h}$ and hence must coincide with $\mathbf{h}$. This completes the proof of the theorem. QED