Lie Algebras

Shlomo Sternberg

March 8, 2004

Chapter 5

Conjugacy of Cartan subalgebras

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, an matrix which commutes with A must be diagonal. Notice that T is a product of circles, i.e. a torus.

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

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

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 $(\operatorname{Ad}_a x', x)$ is a maximum. Let

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

We wish to show that

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

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

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

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. QED

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.

5.1 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 $a, b \in \mathbf{C}$

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

Consequences:

- 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

$$[\mathbf{g}_a, \mathbf{g}_b] \subset \mathbf{g}_{[a+b]}. \tag{5.1}$$

- 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, Inn $\mathbf{g}$ is an ideal in Der $\mathbf{g}$.

- 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 $[\delta, \operatorname{ad} x] = 0$ since $[\delta, \operatorname{ad} x] = \operatorname{ad} \delta x$. 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 $\operatorname{ad} s$ is semisimple and $\operatorname{ad} n$ 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

  $$N_{\mathbf{g}}(\mathbf{k}) = \mathbf{k}. \tag{5.2}$$

We now come to the 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

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

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. Since 0 is not an eigenvalue of $\mathrm{ad}\,z$ on $\mathbf{g}/\mathbf{g}_0(\mathrm{ad}\,z)$, 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,

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

By the minimality, this forces

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

for these values of c. 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.

5.2 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.*

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

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

Notice that every element of $\mathcal{N}(\mathbf{g})$ is nilpotent and that $\mathcal{N}(\mathbf{g})$ is stable under $\mathrm{Aut}(\mathbf{g})$. As any $x \in \mathcal{N}(\mathbf{g})$ is nilpotent, $\exp \mathrm{ad}\,x$ 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.

As a first step we give an alternative characterization of a **CSA**.

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

Proof. Suppose $\mathbf{h} = \mathbf{g}_0(\operatorname{ad} z)$ which is minimal in the sense of the proposition. Then we know by (5.2) 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

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 $\mathbf{k} = \operatorname{Ker} \phi$ 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 $\mathbf{h} = \mathbf{g}_0(\operatorname{ad} z)$ for some minimal such z, and as an algebra containing a $\mathbf{g}_0(\operatorname{ad} z)$, $\mathbf{h}+\mathbf{k}$ is self-normalizing. So $x \in \mathbf{h} + \mathbf{k}$. QED

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

5.3 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

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*

$$\begin{tikzcd}
\mathbf{g} \arrow[r, "\phi"] \arrow[d, "\sigma"'] & \mathbf{g}' \arrow[d, "\sigma'"] \\
\mathbf{g} \arrow[r, "\phi"'] & \mathbf{g}'
\end{tikzcd}$$

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 theorem in the solvable case. 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:

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

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

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

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 $z = [x, z']$, $z' \in \mathbf{a}_*(\operatorname{ad} x)$. 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

5.4 Toral subalgebras and Cartan 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. QED

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.*

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

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

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

- $[\mathbf{g}_\alpha, \mathbf{g}_\beta] \subset \mathbf{g}_{\alpha+\beta}$ (by Jacobi) so
- $\operatorname{ad} x$ 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.*

Our next intermediate step is to prove:

Proposition 3

$$\mathbf{h} = \mathbf{g}_0 \tag{5.3}$$

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

Proceed according to the following steps:

$$x \in \mathbf{g}_0 \ \Rightarrow\ x_s \in \mathbf{g}_0 \ \ x_n \in \ \mathbf{g}_0. \tag{5.4}$$

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 $\mathbf{h} \to 0$.

$$x \in \mathbf{g}_0,\ x \text{ semisimple } \Rightarrow x \in \mathbf{h}. \tag{5.5}$$

Indeed, such an x commutes with all of $\mathbf{h}$. As the sum of commuting semisimple 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 now show that

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 $\kappa(h,x) = 0\ \forall$ semisimple $x \in \mathbf{g}_0$. 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

Next observe that

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

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

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

We next prove

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 (5.3). 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 (5.3). QED

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

5.5 Roots.

We have proved that the restriction of κ to $\mathbf{h}$ is non-degenerate. This allows us to associate to every linear function ϕ on $\mathbf{h}$ the unique element $t_\phi \in \mathbf{h}$ given by

$$\phi(h) = \kappa(t_\phi, h).$$

The set of $\alpha \in \mathbf{h}^*$, $\alpha \neq 0$ for which $\mathbf{g}_\alpha \neq 0$ is called the set of **roots** and is denoted by Φ. We have

- Φ spans $\mathbf{h}^*$ for otherwise $\exists h \neq 0 : \alpha(h) = 0 \ \forall \alpha \in \Phi$ implying that $[h, \mathbf{g}_\alpha] = 0 \ \forall \alpha$ so $[h, \mathbf{g}] = 0$.
- $\alpha \in \Phi \Rightarrow -\alpha \in \Phi$ for otherwise $\mathbf{g}_\alpha \perp \mathbf{g}$.
- $x \in \mathbf{g}_\alpha, y \in \mathbf{g}_{-\alpha}, \alpha \in \Phi \Rightarrow [x, y] = \kappa(x, y) t_\alpha$. Indeed,

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

- $[\mathbf{g}_\alpha, \mathbf{g}_{-\alpha}]$ is one dimensional with basis t_α. This follows from the preceding and the fact that $\mathbf{g}_\alpha$ can not be perpendicular to $\mathbf{g}_{-\alpha}$ since otherwise it will be orthogonal to all of $\mathbf{g}$.

- $\alpha(t_\alpha) = \kappa(t_\alpha, t_\alpha) \neq 0$. Otherwise, choosing $x \in \mathbf{g}_\alpha, y \in \mathbf{g}_{-\alpha}$ with $\kappa(x, y) = 1$, we get
$$[x, y] = t_\alpha, [t_\alpha, x] = [t_\alpha, y] = 0.$$
So x, y, t_α span a solvable three dimensional algebra. Acting as ad on $\mathbf{g}$, it is superdiagonizable, by Lie's theorem, and hence ad t_α, which is in the commutator algebra of this subalgebra is nilpotent. Since it is ad semisimple by definition of $\mathbf{h}$, it must lie in the center, which is impossible.

- Choose $e_\alpha \in \mathbf{g}_\alpha, f_\alpha \in \mathbf{g}_{-\alpha}$ with
$$\kappa(e_\alpha, f_\alpha) = \frac{2}{\kappa(t_\alpha, t_\alpha)}.$$
Set
$$h_\alpha := \frac{2}{\kappa(t_\alpha, t_\alpha)} t_\alpha.$$
Then $e_\alpha, f_\alpha, h_\alpha$ span a subalgebra isomorphic to $sl(2)$. Call it $sl(2)_\alpha$. We shall soon see that this notation is justified, i.e that $\mathbf{g}_\alpha$ is one dimensional and hence that $sl(2)_\alpha$ is well defined, independent of any "choices" of e_α, f_α but depends only on α.

- Consider the action of $sl(2)_\alpha$ on the subalgebra $\mathbf{m} := \mathbf{h} \oplus \bigoplus \mathbf{g}_{n\alpha}$ where $n \in \mathbf{Z}$. The zero eigenvectors of h_α consist of $\mathbf{h} \subset \mathbf{m}$. One of these corresponds to the adjoint representation of $sl(2)_\alpha \subset \mathbf{h}$. The orthocomplement of $h_\alpha \in \mathbf{h}$ gives dim $\mathbf{h} - 1$ trivial representations of $sl(2)_\alpha$. This must exhaust all the even maximal weight representations, as we have accounted for all the zero weights of $sl(2)_\alpha$ acting on $\mathbf{g}$. In particular, dim $\mathbf{g}_\alpha = 1$ and no integer multiple of α other than $-\alpha$ is a root. Now consider the subalgebra $\mathbf{p} := \mathbf{h} \oplus \bigoplus \mathbf{g}_{c\alpha},\ c \in \mathbf{C}$. This is a module for $sl(2)_\alpha$. Hence all such c's must be multiples of $1/2$. But $1/2$ can not occur, since the double of a root is not a root. Hence the $\pm\alpha$ are the only multiples of α which are roots.

Now consider $\beta \in \Phi,\ \beta \neq \pm\alpha$. Let
$$\mathbf{k} := \bigoplus \mathbf{g}_{\beta + j\alpha}.$$

Each non-zero summand is one dimensional, and $\mathbf{k}$ is an $sl(2)_\alpha$ module. Also $\beta + i\alpha \neq 0$ for any i, and evaluation on h_α gives $\beta(h_\alpha) + 2i$. All weights differ by multiples of 2 and so $\mathbf{k}$ is irreducible. Let q be the maximal integer so that $\beta + q\alpha \in \Phi$, and r the maximal integer so that $\beta - r\alpha \in \Phi$. Then the entire string
$$\beta - r\alpha, \beta - (r-1)\alpha, \ldots \beta + q\alpha$$

are roots, and

$$\beta(h_\alpha) - 2r = -(\beta(h_\alpha) + 2q)$$

or

$$\beta(h_\alpha) = r - q \in \mathbf{Z}.$$

These integers are called the **Cartan integers**.

We can transfer the bilinear form κ from $\mathbf{h}$ to $\mathbf{h}^*$ by defining

$$(\gamma, \delta) = \kappa(t_\gamma, t_\delta).$$

So

$$\begin{aligned} \beta(h_\alpha) &= \kappa(t_\beta, h_\alpha) \\ &= \frac{2\kappa(t_\beta, t_\alpha)}{\kappa(t_\alpha, t_\alpha)} \\ &= \frac{2(\beta, \alpha)}{(\alpha, \alpha)}. \end{aligned}$$

So

$$\frac{2(\beta, \alpha)}{(\alpha, \alpha)} = r - q \in \mathbf{Z}.$$

Choose a basis $\alpha_1, \ldots, \alpha_\ell$ of $\mathbf{h}^*$ consisting of roots. This is possible because the roots span $\mathbf{h}^*$. Any root α can be written uniquely as linear combination

$$\beta = c_1\alpha_1 + \cdots + c_\ell\alpha_\ell$$

where the c_i are complex numbers. We claim that in fact the c_i are rational numbers. Indeed, taking the scalar product relative to (,) of this equation with the α_i gives the ℓ equations

$$(\beta, \alpha_i) = c_1(\alpha_1, \alpha_i) + \cdots + c_\ell(\alpha_\ell, \alpha_i).$$

Multiplying the i-th equation by $2/(\alpha_i, \alpha_i)$ gives a set of ℓ equations for the ℓ coefficients c_i where all the coefficients are rational numbers as are the left hand sides. Solving these equations for the c_i shows that the c_i are rational.

Let E be the *real* vector space spanned by the $\alpha \in \Phi$. Then (,) restricts to a real scalar product on E. Also, for any $\lambda \neq 0 \in E$,

$$\begin{aligned} (\lambda, \lambda) &:=: \kappa(t_\lambda, t_\lambda) \\ &:= \operatorname{tr}(\operatorname{ad} t_\lambda)^2 \\ &= \sum_{\alpha \in \Phi} \alpha(t_\lambda)^2 \\ &> 0. \end{aligned}$$

So the scalar product (,) on E is positive definite. E is a Euclidean space.

In the string of roots, β is q steps down from the top, so q steps up from the bottom is also a root, so

$$\beta - (r - q)\alpha$$

is a root, or

$$\beta - \frac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha \in \Phi.$$

But

$$\beta - \frac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha = s_\alpha(\beta)$$

where s_α denotes Euclidean reflection in the hyperplane perpendicular to α. In other words, for every $\alpha \in \Phi$

$$s_\alpha : \Phi \to \Phi. \tag{5.6}$$

The subgroup of the orthogonal group of E generated by these reflections is called the **Weyl group** and is denoted by W. We have thus associated to every semi-simple Lie algebra, and to every choice of Cartan subalgebra a finite subgroup of the orthogonal group generated by reflections. (This subgroup is finite, because all the generating reflections, s_α, and hence the group they generate, preserve the finite set of all roots, which span the space.) Once we will have completed the proof of the conjugacy theorem for Cartan subalgebras of a semi-simple algebra, then we will know that the Weyl group is determined, up to isomorphism, by the semi-simple algebra, and does not depend on the choice of Cartan subalgebra.

We define

$$\langle\beta,\alpha\rangle := \frac{2(\beta,\alpha)}{(\alpha,\alpha)}.$$

So

$$\begin{aligned} \langle\beta,\alpha\rangle &= \beta(h_\alpha) && (5.7)\\ &= r - q \in \mathbf{Z} && (5.8) \end{aligned}$$

and

$$s_\alpha(\beta) = \beta - \langle\beta,\alpha\rangle. \tag{5.9}$$

So far, we have defined the reflection s_α purely in terms of the root struction on E, which is the real subspace of $\mathbf{h}^*$ generated by the roots. But in fact, s_α, and hence the entire Weyl group arises as (an) automorphism(s) of $\mathbf{g}$ which preserve $\mathbf{h}$. Indeed, we know that $e_\alpha, f_\alpha, h_\alpha$ span a subalgebra $sl(2)_\alpha$ isomorphic to $sl(2)$. Now $\exp \operatorname{ad} e_\alpha$ and $\exp \operatorname{ad}(-f_\alpha)$ are elements of $\mathcal{E}(\mathbf{g})$. Consider

$$\tau_\alpha := (\exp \operatorname{ad} e_\alpha)(\exp \operatorname{ad}(-f_\alpha))(\exp \operatorname{ad} e_\alpha) \in \mathcal{E}(\mathbf{g}). \tag{5.10}$$

We claim that

Proposition 4 *The automorphism τ_α preserves $\mathbf{h}$ and on $\mathbf{h}$ it is given by*

$$\tau_\alpha(h) = h - \alpha(h)h_\alpha. \tag{5.11}$$

In particular, the transformation induced by τ_α on E is s_α.

Proof. It suffices to prove (5.11). If $\alpha(h) = 0$, then both $\operatorname{ad} e_\alpha$ and $\operatorname{ad} f_\alpha$ vanish on h so $\tau_\alpha(h) = h$ and (5.11) is true. Now h_α and $\ker \alpha$ span $\mathbf{h}$. So we need only check (5.11) for h_α where it says that $\tau(h_\alpha) = -h_\alpha$. But we have already verified this for the algebra $sl(2)$. QED

We can also verify (5.11) directly. We have

$$\exp(\operatorname{ad} e_\alpha)(h) = h - \alpha(h)e_\alpha$$

for any $h \in \mathbf{h}$. Now $[f_\alpha, e_\alpha] = -h_\alpha$ so

$$(\operatorname{ad} f_\alpha)^2(e_\alpha) = [f_\alpha, -h_\alpha] = [h_\alpha, f_\alpha] = -2f_\alpha.$$

So

$$\begin{aligned} \exp(-\operatorname{ad} f_\alpha)(\exp \operatorname{ad} e_\alpha)h &= (\mathrm{id} - \operatorname{ad} f_\alpha + \frac{1}{2}(\operatorname{ad} f_\alpha)^2)(h - \alpha(h)e_\alpha) \\ &= h - \alpha(h)e_\alpha - \alpha(h)f_\alpha - \alpha(h)h_\alpha + \alpha(h)f_\alpha \\ &= h - \alpha(h)h_\alpha - \alpha(h)e_\alpha. \end{aligned}$$

If we now apply $\exp \operatorname{ad} e_\alpha$ to this last expression and use the fact that $\alpha(h_\alpha) = 2$, we get the right hand side of (5.11).

5.6 Bases.

$\Delta \subset \Phi$ is a called a **Base** if it is a basis of E (so $\#\Delta = \ell = \dim_{\mathbf{R}} E = \dim_{\mathbf{C}} \mathbf{h}$) and every $\beta \in \Phi$ can be written as $\sum_{\alpha \in \Delta} k_\alpha \alpha,\ k_\alpha \in \mathbf{Z}$ with either all the coefficients $k_\alpha \geq 0$ or all ≤ 0. Roots are accordingly called positive or negative and we define the height of a root by

$$\operatorname{ht} \beta := \sum_\alpha k_\alpha.$$

Given a base, we get partial order on E by defining $\lambda \succ \mu$ iff $\lambda - \mu$ is a sum of positive roots or zero. We have

$$(\alpha, \beta) \leq 0, \quad \alpha, \beta \in \Delta \tag{5.12}$$

since otherwise $(\alpha, \beta) > 0$ and

$$s_\alpha(\beta) = \beta - \frac{2(\beta, \alpha)}{(\alpha, \alpha)}\alpha$$

is a root with the coefficient of $\beta = 1 > 0$ and the coefficient of $\alpha < 0$, contradicting the definition which says that roots must have all coefficients non-negative or non-positive.

To construct a base, choose a $\gamma \in E, \mid (\gamma, \beta) \neq 0\ \forall\ \beta \in \Phi$. Such an element is called **regular**. Then every root has positive or negative scalar product with γ, dividing the set of roots into two subsets:

$$\Phi = \Phi^+ \cup \Phi^-, \quad \Phi^- = -\Phi^+.$$

A root $\beta \in \Phi^+$ is called **decomposable** if $\beta = \beta_1 + \beta_2, \beta_1, \beta_2 \in \Phi^+$, indecomposable otherwise. Let $\Delta(\gamma)$ consist of the indecomposable elements of $\Phi^+(\gamma)$.

Theorem 3 *$\Delta(\gamma)$ is a base, and every base is of the form $\Delta(\gamma)$ for some γ.*

Proof. Every $\beta \in \Phi^+$ can be written as a non-negative integer combination of $\Delta(\gamma)$ for otherwise choose one that can not be so written with (γ, β) as small as possible. In particular, β is not indecomposable. Write $\beta = \beta_1 + \beta_2, \beta_i \in \Phi^+$. Then $\beta \not\in \Delta(\gamma), (\gamma, \beta) = (\gamma, \beta_1) + (\gamma, \beta_2)$ and hence $(\gamma, \beta_1) < (\gamma, \beta)$ and $(\gamma, \beta_2) < (\gamma, \beta)$. By our choice of β this means β_1 and β_2 *are* non-negative integer combinations of elements of $\Delta(\gamma)$ and hence so is β, contradiction.

Now (5.12) holds for $\Delta = \Delta(\gamma)$ for if not, $\alpha - \beta$ is a root, so either $\alpha - \beta \in \Phi^+$ so $\alpha = \alpha - \beta + \beta$ is decomposable or $\beta - \alpha \in \Phi^+$ and β is decomposable.

This implies that $\Delta(\gamma)$ is linearly independent: for suppose $\sum_\alpha c_\alpha \alpha = 0$ and let p_α be the positive coefficients and $-q_\beta$ the negative ones, so

$$\sum_\alpha p_\alpha \alpha = \sum_\beta q_\beta \beta$$

all coefficients positive. Let ϵ be this common vector. Then $(\epsilon, \epsilon) = \sum p_\alpha q_\beta (\alpha, \beta) \leq 0$ so $\epsilon = 0$ which is impossible unless all the coefficients vanish, since all scalar products with γ are strictly positive. Since the elements of Φ span E this shows that $\Delta(\gamma)$ is a basis of E and hence a base.

Now let us show that every base is of the desired form: For any base Δ, let $\Phi^+ = \Phi^+(\Delta)$ denote the set of those roots which are non-negative integral combinations of the elements of Δ and let $\Phi^- = \Phi^-(\Delta)$ denote the ones which are non-positive integral combinations of elements of Δ. Define $\delta_\alpha,\ \alpha \in \Delta$ to be the projection of α onto the orthogonal complement of the space spanned by the other elements of the base. Then

$$(\delta_\alpha, \alpha') = 0, \quad \alpha \neq \alpha',\ (\delta_\alpha, \alpha) = (\delta_\alpha, \delta_\alpha) > 0$$

so $\gamma = \sum r_\alpha \delta_\alpha,\ r_\alpha > 0$ satisfies

$$(\gamma, \alpha) > 0\ \forall\ \alpha \in \Delta$$

hence

$$\Phi^+(\Delta) \subset \Phi^+(\gamma)$$

and

$$\Phi^-(\Delta) \subset \Phi^-(\gamma)$$

hence

$$\Phi^+(\Delta) = \Phi^+(\gamma) \ \text{ and } \ \ \Phi^-(\Delta) = \Phi^-(\gamma).$$

Since every element of Φ^+ can be written as a sum of elements of Δ with non-negative integer coefficients, the only indecomposable elements can be the Δ, so $\Delta(\gamma) \subset \Delta$ but then they must be equal since they have the same cardinality $\ell = \dim E$. QED

5.7 Weyl chambers.

Define $P_\beta := \beta^\perp$. Then $E - \bigcup P_\beta$ is the union of **Weyl chambers** each consisting of regular γ's with the same Φ^+. So the Weyl chambers are in one to one correspondence with the bases, and the Weyl group permutes them.

Fix a base, Δ. Our goal in this section is to prove that the reflections $s_\alpha,\ \alpha \in \Delta$ generate the Weyl group, W, and that W acts simply transitively on the Weyl chambers.

Each $s_\alpha,\ \alpha \in \Delta$ sends $\alpha \mapsto -\alpha$. But acting on $\lambda = \sum c_\beta \beta$, the reflection s_α does not change the coefficient of any other element of the base. If $\lambda \in \Phi^+$ and $\lambda \neq \alpha$, we must have $c_\beta > 0$ for some $\beta \neq \alpha$ in the base Δ. Then the coefficient of β in the expansion of $s_\alpha(\lambda)$ is positive, and hence all its coefficients must be non-negative. So $s_\alpha(\lambda) \in \Phi^+$. In short, the only element of Φ^+ sent into Φ^- is α. So if

$$\delta := \frac{1}{2}\sum_{\beta\in\Phi^+} \beta \text{ then } s_\alpha\delta = \delta - \alpha.$$

If $\beta \in \Phi^+,\ \beta \notin \Delta$, then we can not have $(\beta, \alpha) \leq 0\ \forall \alpha \in \Delta$ for then $\beta \cup \Delta$ would be linearly independent. So $\beta - \alpha$ is a root for some $\alpha \in \Delta$, and since we have changed only one coefficient, it must be a positive root. Hence any $\beta \in \Phi$ can be written as

$$\beta = \alpha_1 + \cdots + \alpha_p \quad \alpha_i \in \Delta$$

where all the partial sums are positive roots.

Let γ be any vector in a Euclidean space, and let s_γ denote reflection in the hyperplane orthogonal to γ. Let R be any orthogonal transformation. Then

$$s_{R\gamma} = R s_\gamma R^{-1} \tag{5.13}$$

as follows immediately from the definition.

Let $\alpha_1, \ldots, \alpha_i \in \Delta$, and, for short, let us write $s_i := s_{\alpha_i}$.

Lemma 10 *If* $s_1 \cdots s_{i-1}\alpha_i < 0$ *then* $\exists j < i, j \geq 1$ *so that*

$$s_1 \cdots s_i = s_1 \cdots s_{j-1}s_{j+1} \cdots s_{i-1}.$$

Proof. Set $\beta_{i-1} := \alpha_i,\ \beta_j := s_{j+1}\cdots s_{i-1}\alpha_i,\ \ j < i-1$. Since $\beta_{i-1} \in \Phi^+$ and $\beta_0 \in \Phi^-$ there must be some j for which $\beta_j \in \Phi^+$ and $s_j\beta_j = \beta_{j-1} \in \Phi^-$ implying that that $\beta_j = \alpha_j$ so by (5.13) with $R = s_{j+1}\cdots s_{i-1}$ we conclude that

$$s_j = (s_{j+1}\cdots s_{i-1})s_i(s_{j+1}\cdots s_{i-1})^{-1}$$

or

$$s_j s_{j+1} \cdots s_i = s_{j+1}\cdots s_{i-1}$$

implying the lemma. QED

As a consequence, if $s = s_1 \cdots s_t$ is a shortest expression for s, then, since $s_t\alpha_t \in \Phi^-$, we must have $s\alpha_t \in \Phi^-$.

Keeping Δ fixed in the ensuing discussion, we will call the elments of Δ **simple** roots, and the corresponding reflections **simple** reflections. Let W' denote the subgroup of W generated by the simple reflections, $s_\alpha, \alpha \in \Delta$. (Eventually we will prove that this is all of W.) It now follows that if $s \in W'$ and $s\Delta = \Delta$ then $s = id$. Indeed, if $s \neq id$, write s in a minimal fashion as a product of simple reflections. By what we have just proved, it must send some simple root into a negative root. So W' permutes the Weyl chambers without fixed points. We now show that W' acts transitively on the Weyl chambers:

Let $\gamma \in E$ be a regular element. We claim

$$\exists \;\; s \in W' \text{ with } (s(\gamma), \alpha) > 0 \;\forall\; \alpha \in \Delta.$$

Indeed, choose $s \in W'$ so that $(s(\gamma), \delta)$ is as large as possible. Then

$$\begin{aligned} (s(\gamma), \delta) &\geq (s_\alpha s(\gamma), \delta) \\ &= (s(\gamma), s_\alpha \delta) \\ &= (s(\gamma), \delta) - (s(\gamma), \alpha) \text{ so} \\ (s(\gamma), \alpha) &\geq 0 \;\;\forall\; \alpha \in \Delta. \end{aligned}$$

We can't have equality in this last inequality since $s(\gamma)$ is not orthogonal to any root. This proves that W' acts transitively on all Weyl chambers and hence on all bases.

We next claim that every root belongs to at least one base. Choose a (non-regular) $\gamma' \perp \alpha$, but $\gamma' \notin P_\beta,\; \beta \neq \alpha$. Then choose γ close enough to γ' so that $(\gamma, \alpha) > 0$ and $(\gamma, \alpha) < |(\gamma, \beta)| \;\forall\; \beta \neq \alpha$. Then in $\Phi^+(\gamma)$ the element α must be indecomposable. If β is any root, we have shown that there is an $s' \in W'$ with $s'\beta = \alpha_i \in \Delta$. By (5.13) this implies that every reflection $s_\beta$in W is conjugate by an element of W' to a simple reflection: $s_\beta = s' s_i {s'}^{-1} \in W'$. Since W is generated by the s_β, this shows that $W' = W$.

5.8 Length.

Define the length of an element of W as the minimal word length in its expression as a product of simple roots. Define $n(s)$ to be the number of positive roots made negative by s. We know that $n(s) = \ell(s)$ if $\ell(s) = 0$ or 1. We claim that

$$\ell(s) = n(s)$$

in general.

Proof by induction on $\ell(s)$. Write $s = s_1 \cdots s_i$ in reduced form and let $\alpha = \alpha_i$. We have $s\alpha \in \Phi^-$. Then $n(ss_i) = n(s) - 1$ since s_i leaves all positive roots positive except α. Also $\ell(ss_i) = \ell(s) - 1$. So apply induction. QED

Let $C = C(\Delta)$ be the Weyl chamber associated to the base Δ. Let $\overline{C}$ denote its closure.

Lemma 11 *If $\lambda, \mu \in \overline{C}$ and $s \in W$ satisfies $s\lambda = \mu$ then s is a product of simple reflections which fix λ. In particular, $\lambda = \mu$. So $\overline{C}$ is a fundamental domain for the action of W on E.*

Proof. By induction on $\ell(s)$. If $\ell(s) = 0$ then $s = id$ and the assertion is clear with the empty product. So we may assume that $n(s) > 0$, so s sends some positive root to a negative root, and hence must send some simple root to a negative root. So let $\alpha \in \Delta$ be such that $s\alpha \in \Phi^-$. Since $\mu \in \overline{C}$, we have $(\mu, \beta) \geq 0$, $\forall \beta \in \Phi^+$ and hence $(\mu, s\alpha) \leq 0$. So

$$\begin{aligned} 0 &\geq (\mu, s\alpha) \\ &= (s^{-1}\mu, \alpha) \\ &= (\lambda, \alpha) \\ &\geq 0. \end{aligned}$$

So $(\lambda, \alpha) = 0$ so $s_\alpha \lambda = \lambda$ and hence $ss_\alpha \lambda = \mu$. But $n(ss_\alpha) = n(s) - 1$ since $s_\alpha = -\alpha$ and s_α permutes all the other positive roots. So $\ell(ss_\alpha) = \ell(s) - 1$ and we can apply induction to conclude that $s = (ss_\alpha)s_\alpha$ is a product of simple reflections which fix λ.

5.9 Conjugacy of Borel subalgebras

We need to prove this for semi-simple algebras since the radical is contained in every maximal solvable subalgebra.

Define a **standard** Borel subalgebra (relative to a choice of **CSA** $\mathbf{h}$ and a system of simple roots, Δ) to be

$$\mathbf{b}(\Delta) := \mathbf{h} \oplus \bigoplus_{\beta \in \Phi^+(\Delta)} \mathbf{g}_\beta.$$

Define the corresponding nilpotent Lie algebra by

$$\mathbf{n}_+(\Delta) := \bigoplus_{\beta \in \Phi^+} \mathbf{g}_\beta.$$

Since each s_α can be realized as $(\exp e_\alpha)(\exp -f_\alpha)(\exp e_\alpha)$ every element of W can be realized as an element of $\mathcal{E}(\mathbf{g})$. Hence all standard Borel subalgebras relative to a given Cartan subalgebra are conjugate.

Notice that if x normalizes a Borel subalgebra, $\mathbf{b}$, then

$$[\mathbf{b} + \mathbf{C}x, \mathbf{b} + \mathbf{C}x] \subset \mathbf{b}$$

and so $\mathbf{b} + \mathbf{C}x$ is a solvable subalgebra containing $\mathbf{b}$ and hence must coincide with $\mathbf{b}$:

$$N_{\mathbf{g}}(\mathbf{b}) = \mathbf{b}.$$

In particular, if $x \in \mathbf{b}$ then its semi-simple and nilpotent parts lie in $\mathbf{b}$.

From now on, fix a standard **BSA**, $\mathbf{b}$. We want to prove that any other **BSA**, $\mathbf{b}'$ is conjugate to $\mathbf{b}$. We may assume that the theorem is known for Lie algebras of smaller dimension, or for $\mathbf{b}'$ with $\mathbf{b} \cap \mathbf{b}'$ of greater dimension, since if dim

$\mathbf{b}\cap\mathbf{b}' = \dim\mathbf{b}$, so that $\mathbf{b}'\supset\mathbf{b}$, we must have $\mathbf{b}'=\mathbf{b}$ by maximality. Therefore we can proceed by downward induction on the dimension of the intersection $\mathbf{b}\cap\mathbf{b}'$.

Suppose $\mathbf{b}\cap\mathbf{b}'\neq 0$. Let $\mathbf{n}'$ be the set of nilpotent elements in $\mathbf{b}\cap\mathbf{b}'$. So $\mathbf{n}'=\mathbf{n}^+\cap\mathbf{b}'$.

Also $[\mathbf{b}\cap\mathbf{b}',\mathbf{b}\cap\mathbf{b}']\subset\mathbf{n}^+\cap\mathbf{b}'=\mathbf{n}'$ so $\mathbf{n}'$ is a nilpotent ideal in $\mathbf{b}\cap\mathbf{b}'$. Suppose that $\mathbf{n}'\neq 0$. Then since $\mathbf{g}$ contains no solvable ideals,

$$\mathbf{k}:=N_{\mathbf{g}}(\mathbf{n}')\neq\mathbf{g}.$$

Consider the action of $\mathbf{n}'$ on $\mathbf{b}/(\mathbf{b}\cap\mathbf{b}')$. By Engel, there exists a $y\notin\mathbf{b}\cap\mathbf{b}'$ with $[x,y]\in\mathbf{b}\cap\mathbf{b}'\ \forall x\in\mathbf{n}'$. But $[x,y]\in[\mathbf{b},\mathbf{b}]\subset\mathbf{n}^+$ and so $[x,y]\in\mathbf{n}'$. So $y\in\mathbf{k}$. Thus $y\in\mathbf{k}\cap\mathbf{b},\ y\notin\mathbf{b}\cap\mathbf{b}'$. Similarly, we can interchange the roles of $\mathbf{b}$ and $\mathbf{b}'$ in the above argument, replacing $\mathbf{n}^+$ by the nilpotent subalgebra $[\mathbf{b}',\mathbf{b}']$ of $\mathbf{b}'$, to conclude that there exists a $y'\in\mathbf{k}\cap\mathbf{b}',\ y'\notin\mathbf{b}\cap\mathbf{b}'$. In other words, the inclusions

$$\mathbf{k}\cap\mathbf{b}\supset\mathbf{b}\cap\mathbf{b}',\quad \mathbf{k}\cap\mathbf{b}'\supset\mathbf{b}\cap\mathbf{b}'$$

are strict.

Both $\mathbf{b}\cap\mathbf{k}$ and $\mathbf{b}'\cap\mathbf{k}$ are solvable subalgebras of $\mathbf{k}$. Let $\mathbf{c},\mathbf{c}'$ be **BSA**'s containing them. By induction, there is a $\sigma\in\mathcal{E}(\mathbf{k})\subset\mathcal{E}(\mathbf{g})$ with $\sigma(\mathbf{c}')=\mathbf{c}$. Now let $\mathbf{b}''$ be a **BSA** containing $\mathbf{c}$. We have

$$\mathbf{b}''\cap\mathbf{b}\supset\mathbf{c}\cap\mathbf{b}\supset\mathbf{k}\cap\mathbf{b}\supset\mathbf{b}'\cap\mathbf{b}$$

with the last inclusion strict. So by induction there is a $\tau\in\mathcal{E}(\mathbf{g})$ with $\tau(\mathbf{b}'')=\mathbf{b}$. Hence

$$\tau\sigma(\mathbf{c}')\subset\mathbf{b}.$$

Then

$$\mathbf{b}\cap\tau\sigma(\mathbf{b}')\supset\tau\sigma(\mathbf{c}')\cap\tau\sigma(\mathbf{b}')\supset\tau\sigma(\mathbf{b}'\cap\mathbf{k})\supset\tau\sigma(\mathbf{b}\cap\mathbf{b}')$$

with the last inclusion strict. So by induction we can further conjugate $\tau\sigma\mathbf{b}'$ into $\mathbf{b}$.

So we must now deal with the case that $\mathbf{n}'=0$, but we will still assume that $\mathbf{b}\cap\mathbf{b}'\neq 0$. Since any Borel subalgebra contains both the semi-simple and nilpotent parts of any of its elements, we conclude that $\mathbf{b}\cap\mathbf{b}'$ consists entirely of semi-simple elements, and so is a toral subalgebra, call it $\mathbf{t}$. If $x\in\mathbf{b},t\in\mathbf{t}=\mathbf{b}\cap\mathbf{b}'$ and $[x,t]\in\mathbf{t}$, then we must have $[x,t]=0$, since all elements of $[\mathbf{b},\mathbf{b}]$ are nilpotent. So

$$N_{\mathbf{b}}(\mathbf{t})=C_{\mathbf{b}}(\mathbf{t}).$$

Let $\mathbf{c}$ be a **CSA** of $C_{\mathbf{b}}(\mathbf{t})$. Since a Cartan subalgebra is its own normalizer, we have $\mathbf{t}\subset\mathbf{c}$. So we have

$$\mathbf{t}\subset\mathbf{c}\subset C_{\mathbf{b}}(\mathbf{t})=N_{\mathbf{b}}(\mathbf{t})\subset N_{\mathbf{b}}(\mathbf{c}).$$

Let $t\in\mathbf{t},\ n\in N_{\mathbf{b}}(\mathbf{c})$. Then $[t,n]\in\mathbf{c}$ and successive brackets by t will eventually yield 0, since $\mathbf{c}$ is nilpotent. Thus $(\mathrm{ad}\ t)^k n=0$ for some k, and since t is semi-simple, $[t,n]=0$. Thus $n\in C_{\mathbf{b}}(\mathbf{t})$ and hence $n\in\mathbf{c}$ since $\mathbf{c}$ is its own normalizer in $C_{\mathbf{b}}(\mathbf{t})$. Thus $\mathbf{c}$ is a **CSA** of $\mathbf{b}$. We can now apply the conjugacy theorem for **CSA**'s of solvable algebras to conjugate $\mathbf{c}$ into $\mathbf{h}$.

So we may assume from now on that $\mathbf{t} \subset \mathbf{h}$. If $\mathbf{t} = \mathbf{h}$, then decomposing $\mathbf{b}'$ into root spaces under $\mathbf{h}$, we find that the non-zero root spaces must consist entirely of negative roots, and there must be at least one such, since $\mathbf{b}' \neq \mathbf{h}$. But then we can find a τ_α which conjugates this into a positive root, preserving $\mathbf{h}$, and then $\tau_\alpha(\mathbf{b}') \cap \mathbf{b}$ has larger dimension and we can further conjugate into $\mathbf{b}$.

So we may assume that

$$\mathbf{t} \subset \mathbf{h}$$

is strict.

If

$$\mathbf{b}' \subset C_{\mathbf{g}}(\mathbf{t})$$

then since we also have $\mathbf{h} \subset C_{\mathbf{g}}(\mathbf{t})$, we can find a **BSA**, $\mathbf{b}''$ of $C_{\mathbf{g}}(\mathbf{t})$ containing $\mathbf{h}$, and conjugate $\mathbf{b}'$ to $\mathbf{b}''$, since we are assuming that $\mathbf{t} \neq 0$ and hence $C_{\mathbf{g}}(\mathbf{t}) \neq \mathbf{g}$. Since $\mathbf{b}'' \cap \mathbf{b} \supset \mathbf{h}$ has bigger dimension than $\mathbf{b}' \cap \mathbf{b}$, we can further conjugate to $\mathbf{b}$ by the induction hypothesis.

If

$$\mathbf{b}' \not\subset C_{\mathbf{g}}(\mathbf{t})$$

then there is a common non-zero eigenvector for ad t in $\mathbf{b}'$, call it x. So there is a $t' \in \mathbf{t}$ such that $[t', x] = c'x,\ c' \neq 0$. Setting

$$t := \frac{1}{c'} t'$$

we have $[t, x] = x$. Let $\Phi_t \subset \Phi$ consist of those roots for which $\beta(t)$ is a positive rational number. Then

$$\mathbf{s} := \mathbf{h} \oplus \bigoplus_{\beta \in \Phi_t} \mathbf{g}_\beta$$

is a solvable subalgebra and so lies in a **BSA**, call it $\mathbf{b}''$. Since $\mathbf{t} \subset \mathbf{b}'', x \in \mathbf{b}''$ we see that $\mathbf{b}'' \cap \mathbf{b}'$ has strictly larger dimension than $\mathbf{b} \cap \mathbf{b}'$. Also $\mathbf{b}'' \cap \mathbf{b}$ has strictly larger dimension than $\mathbf{b} \cap \mathbf{b}'$ since $\mathbf{h} \subset \mathbf{b} \cap \mathbf{b}''$. So we can conjugate $\mathbf{b}'$ to $\mathbf{b}''$ and then $\mathbf{b}''$ to $\mathbf{b}$.

This leaves only the case $\mathbf{b} \cap \mathbf{b}' = 0$ which we will show is impossible. Let $\mathbf{t}$ be a maximal toral subalgebra of $\mathbf{b}'$. We can not have $\mathbf{t} = 0$, for then $\mathbf{b}'$ would consist entirely of nilpotent elements, hence nilpotent by Engel, and also self-normalizing as is every **BSA**. Hence it would be a **CSA** which is impossible since every **CSA** in a semi-simple Lie algebra is toral. So choose a **CSA**, $\mathbf{h}''$ containing $\mathbf{t}$, and then a standard **BSA** containing $\mathbf{h}''$. By the preceding, we know that $\mathbf{b}'$ is conjugate to $\mathbf{b}''$ and, in particular has the same dimension as $\mathbf{b}''$. But the dimension of each standard **BSA** (relative to any Cartan subalgebra) is strictly greater than half the dimension of $\mathbf{g}$, contradicting the hypothesis $\mathbf{g} \supset \mathbf{b} \oplus \mathbf{b}'$. QED