Lie algebras

Shlomo Sternberg

April 23, 2004

Contents

Chapter 1

The Campbell Baker Hausdorff Formula

1.1 The problem.

Recall the power series:

$$\exp X = 1 + X + \frac{1}{2}X^2 + \frac{1}{3!}X^3 + \cdots, \quad \log(1+X) = X - \frac{1}{2}X^2 + \frac{1}{3}X^3 + \cdots.$$

We want to study these series in a ring where convergence makes sense; for example in the ring of $n \times n$ matrices. The exponential series converges everywhere, and the series for the logarithm converges in a small enough neighborhood of the origin. Of course,

$$\log(\exp X) = X; \quad \exp(\log(1+X)) = 1 + X$$

where these series converge, or as formal power series.

In particular, if A and B are two elements which are close enough to 0 we can study the convergent series

$$\log[(\exp A)(\exp B)]$$

which will yield an element C such that $\exp C = (\exp A)(\exp B)$. The problem is that A and B need not commute. For example, if we retain only the linear and constant terms in the series we find

$$\log[(1 + A + \cdots)(1 + B + \cdots)] = \log(1 + A + B + \cdots) = A + B + \cdots.$$

On the other hand, if we go out to terms second order, the non-commutativity begins to enter:

$$\log[(1 + A + \frac{1}{2}A^2 + \cdots)(1 + B + \frac{1}{2}B^2 + \cdots)] =$$

$$A + B + \frac{1}{2}A^2 + AB + \frac{1}{2}B^2 - \frac{1}{2}(A + B + \cdots)^2$$
$$= A + B + \frac{1}{2}[A, B] + \cdots$$

where
$$[A, B] := AB - BA \tag{1.1}$$
is the **commutator** of A and B, also known as the **Lie bracket** of A and B.

Collecting the terms of degree three we get, after some computation,

$$\frac{1}{12}\left(A^2B + AB^2 + B^2A + BA^2 - 2ABA - 2BAB\right) = \frac{1}{12}[A, [A, B]] + \frac{1}{12}[B, [B, A]].$$

This suggests that the series for $\log[(\exp A)(\exp B)]$ can be expressed entirely in terms of successive Lie brackets of A and B. This is so, and is the content of the Campbell-Baker-Hausdorff formula.

One of the important consequences of the mere existence of this formula is the following. Suppose that $\mathbf{g}$ is the Lie algebra of a Lie group G. Then the *local* structure of G near the identity, i.e. the rule for the product of two elements of G sufficiently closed to the identity is determined by its Lie algebra $\mathbf{g}$. Indeed, the exponential map is locally a diffeomorphism from a neighborhood of the origin in $\mathbf{g}$ onto a neighborhood W of the identity, and if $U \subset W$ is a (possibly smaller) neighborhood of the identity such that $U \cdot U \subset W$, the the product of $a = \exp \xi$ and $b = \exp \eta$, with $a \in U$ and $b \in U$ is then completely expressed in terms of successive Lie brackets of ξ and η.

We will give two proofs of this important theorem. One will be geometric - the explicit formula for the series for $\log[(\exp A)(\exp B)]$ will involve integration, and so makes sense over the real or complex numbers. We will derive the formula from the "Maurer-Cartan equations" which we will explain in the course of our discussion. Our second version will be more algebraic. It will involve such ideas as the universal enveloping algebra, comultiplication and the Poincaré-Birkhoff-Witt theorem. In both proofs, many of the key ideas are at least as important as the theorem itself.

1.2 The geometric version of the CBH formula.

To state this formula we introduce some notation. Let ad A denote the operation of bracketing on the left by A, so

$$\mathrm{ad}A(B) := [A, B].$$

Define the function ψ by
$$\psi(z) = \frac{z \log z}{z - 1}$$
which is defined as a convergent power series around the point $z = 1$ so

$$\psi(1 + u) = (1 + u)\frac{\log(1 + u)}{u} = (1 + u)(1 - \frac{u}{2} + \frac{u^2}{3} + \cdots) = 1 + \frac{u}{2} - \frac{u^2}{6} + \cdots .$$

In fact, we will also take this as a *definition* of the formal power series for ψ in terms of u. The Campbell-Baker-Hausdorff formula says that

$$\log((\exp A)(\exp B)) = A + \int_0^1 \psi\left((\exp \mathrm{ad}\ A)(\exp t\mathrm{ad}\ B)\right) B dt. \tag{1.2}$$

Remarks.

1. The formula says that we are to substitute

$$u = (\exp \mathrm{ad}\ A)(\exp t\mathrm{ad}\ B) - 1$$

into the definition of ψ, apply this operator to the element B and then integrate. In carrying out this computation we can ignore all terms in the expansion of ψ in terms of ad A and ad B where a factor of ad B occurs on the right, since $(\mathrm{ad}\ B)B = 0$. For example, to obtain the expansion through terms of degree three in the Campbell-Baker-Hausdorff formula, we need only retain quadratic and lower order terms in u, and so

$$\begin{aligned} u &= \mathrm{ad}\ A + \frac{1}{2}(\mathrm{ad}\ A)^2 + t\mathrm{ad}\ B + \frac{t^2}{2}(\mathrm{ad}\ B)^2 + \cdots \\ u^2 &= (\mathrm{ad}\ A)^2 + t(\mathrm{ad}\ B)(\mathrm{ad}\ A) + \cdots \\ \int_0^1 \left(1 + \frac{u}{2} - \frac{u^2}{6}\right) dt &= 1 + \frac{1}{2}\mathrm{ad}\ A + \frac{1}{12}(\mathrm{ad}\ A)^2 - \frac{1}{12}(\mathrm{ad}\ B)(\mathrm{ad}\ A) + \cdots, \end{aligned}$$

where the dots indicate either higher order terms or terms with ad B occurring on the right. So up through degree three (1.2) gives

$$\log(\exp A)(\exp B) = A + B + \frac{1}{2}[A,B] + \frac{1}{12}[A,[A,B]] - \frac{1}{12}[B,[A,B]] + \cdots$$

agreeing with our preceding computation.

2. The meaning of the exponential function on the left hand side of the Campbell-Baker-Hausdorff formula differs from its meaning on the right. On the right hand side, exponentiation takes place in the algebra of endomorphisms of the ring in question. In fact, we will want to make a fundamental reinterpretation of the formula. We want to think of A, B, etc. as elements of a Lie algebra, $\mathbf{g}$. Then the exponentiations on the right hand side of (1.2) are still taking place in $\mathrm{End}(\mathbf{g})$. On the other hand, if $\mathbf{g}$ is the Lie algebra of a Lie group G, then there is an exponential map: exp: $\mathbf{g} \to G$, and this is what is meant by the exponentials on the left of (1.2). This exponential map is a diffeomorphism on some neighborhood of the origin in $\mathbf{g}$, and its inverse, log, is defined in some neighborhood of the identity in G. This is the meaning we will attach to the logarithm occurring on the left in (1.2).

3. The most crucial consequence of the Campbell-Baker-Hausdorff formula is that it shows that the local structure of the Lie group G (the multiplication law for elements near the identity) is completely determined by its Lie algebra.

4. For example, we see from the right hand side of (1.2) that group multiplication and group inverse are analytic if we use exponential coordinates.

5. Consider the function τ defined by

$$\tau(w) := \frac{w}{1-e^{-w}}. \tag{1.3}$$

This is a familiar function from analysis, as it enters into the Euler-Maclaurin formula, see below. (It is the exponential generating function of $(-1)^k b_k$ where the b_k are the Bernoulli numbers.) Then

$$\psi(z) = \tau(\log z).$$

6. The formula is named after three mathematicians, Campbell, Baker, and Hausdorff. But this is a misnomer. Substantially earlier than the works of any of these three, there appeared a paper by Friedrich Schur, "Neue Begruendung der Theorie der endlichen Transformationsgruppen," *Mathematische Annalen* **35** (1890), 161-197. Schur writes down, as convergent power series, the composition law for a Lie group in terms of "canonical coordinates", i.e., in terms of linear coordinates on the Lie algebra. He writes down recursive relations for the coefficients, obtaining a version of the formulas we will give below. I am indebted to Prof. Schmid for this reference.

Our strategy for the proof of (1.2) will be to prove a differential version of it:

$$\frac{d}{dt}\log\left((\exp A)(\exp tB)\right) = \psi\left((\exp \mathrm{ad}\ A)(\exp t\ \mathrm{ad}\ B)\right)B. \tag{1.4}$$

Since $\log(\exp A(\exp tB)) = A$ when $t = 0$, integrating (1.4) from 0 to 1 will prove (1.2). Let us define $\Gamma = \Gamma(t) = \Gamma(t, A, B)$ by

$$\Gamma = \log\left((\exp A)(\exp tB)\right). \tag{1.5}$$

Then

$$\exp\Gamma = \exp A \exp tB$$

and so

$$\begin{aligned}
\frac{d}{dt}\exp\Gamma(t) &= \exp A\frac{d}{dt}\exp tB \\
&= \exp A(\exp tB)B \\
&= (\exp\Gamma(t))B \quad \text{so} \\
(\exp -\Gamma(t))\frac{d}{dt}\exp\Gamma(t) &= B.
\end{aligned}$$

We will prove (1.4) by finding a general expression for

$$\exp(-C(t))\frac{d}{dt}\exp(C(t))$$

where $C = C(t)$ is a curve in the Lie algebra, $\mathbf{g}$, see (1.11) below.

In our derivation of (1.4) from (1.11) we will make use of an important property of the adjoint representation which we might as well state now: For any $g \in G$, define the linear transformation

$$\mathrm{Ad}\ g : \mathbf{g} \to \mathbf{g} : X \mapsto gXg^{-1}.$$

(In geometrical terms, this can be thought of as follows: (The differential of) Left multiplication by g carries $\mathbf{g} = T_I(G)$ into the tangent space, $T_g(G)$ to G at the point g. Right multiplication by g^{-1} carries this tangent space back to $\mathbf{g}$ and so the combined operation is a linear map of $\mathbf{g}$ into itself which we call Ad g. Notice that Ad is a representation in the sense that

$$\mathrm{Ad}\ (gh) = (\mathrm{Ad}\ g)(\mathrm{Ad}\ h)\ \ \forall g, h \in G.$$

In particular, for any $A \in \mathbf{g}$, we have the one parameter family of linear transformations $\mathrm{Ad}(\exp tA)$ and

$$\begin{aligned} \frac{d}{dt}\mathrm{Ad}\ (\exp tA)X &= (\exp tA)AX(\exp -tA) + (\exp tA)X(-A)(\exp -tA) \\ &= (\exp tA)[A, X](\exp -tA) \quad \text{so} \\ \frac{d}{dt}\mathrm{Ad}\ \exp tA &= \mathrm{Ad}(\exp tA) \circ \mathrm{ad}\ A. \end{aligned}$$

But ad A is a linear transformation acting on $\mathbf{g}$ and the solution to the differential equation

$$\frac{d}{dt}M(t) = M(t)\mathrm{ad}\ A, \quad M(0) = I$$

(in the space of linear transformations of $\mathbf{g}$) is $\exp t\,\mathrm{ad}\,A$. Thus $\mathrm{Ad}(\exp tA) = \exp(t\,\mathrm{ad}\,A)$. Setting $t = 1$ gives the important formula

$$\mathrm{Ad}\ (\exp A) = \exp(\mathrm{ad}\ A). \tag{1.6}$$

As an application, consider the Γ introduced above. We have

$$\begin{aligned} \exp(\mathrm{ad}\ \Gamma) &= \mathrm{Ad}\ (\exp \Gamma) \\ &= \mathrm{Ad}\ ((\exp A)(\exp tB)) \\ &= (\mathrm{Ad}\ \exp A)(\mathrm{Ad}\ \exp tB) \\ &= (\exp \mathrm{ad}\ A)(\exp \mathrm{ad}\ tB) \end{aligned}$$

hence

$$\mathrm{ad}\ \Gamma = \log((\exp \mathrm{ad}\ A)(\exp \mathrm{ad}\ tB)). \tag{1.7}$$

1.3 The Maurer-Cartan equations.

If G is a Lie group and $\gamma = \gamma(t)$ is a curve on G with $\gamma(0) = A \in G$, then $A^{-1}\gamma$ is a curve which passes through the identity at $t = 0$. Hence $A^{-1}\gamma'(0)$ is a tangent vector at the identity, i.e. an element of $\mathbf{g}$, the Lie algebra of G.

In this way, we have defined a linear differential form θ on G with values in $\mathbf{g}$. In case G is a subgroup of the group of all invertible $n \times n$ matrices (say over the real numbers), we can write this form as

$$\theta = A^{-1}dA.$$

We can then think of the A occurring above as a collection of n^2 real valued functions on G (the matrix entries considered as functions on the group) and dA as the matrix of differentials of these functions. The above equation giving θ is then just matrix multiplication. For simplicity, we will work in this case, although the main theorem, equation (1.8) below, works for any Lie group and is quite standard.

The definitions of the groups we are considering amount to constraints on A, and then differentiating these constraints show that $A^{-1}dA$ takes values in $\mathbf{g}$, and gives a description of $\mathbf{g}$. It is best to explain this by examples:

- O(n): $AA^\dagger = I, \quad dAA^\dagger + AdA^\dagger = 0$ or

  $$A^{-1}dA + \left(A^{-1}dA\right)^\dagger = 0.$$

  $o(n)$ consists of antisymmetric matrices.

- Sp(n): Let

  $$J := \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

  and let Sp(n) consist of all matrices satisfying

  $$AJA^\dagger = J.$$

  Then

  $$dAJa^\dagger + AJdA^\dagger = 0$$

  or

  $$(A^{-1}dA)J + J(A^{-1}dA)^\dagger = 0.$$

  The equation $BJ + JB^\dagger = 0$ defines the Lie algebra sp(n).

- Let J be as above and define Gl(n,C) to consist of all invertible matrices satisfying

  $$AJ = JA.$$

  Then

  $$dAJ = JdA = 0.$$

  and so

  $$A^{-1}dAJ = A^{-1}JdA = JA^{-1}dA.$$

We return to general considerations: Let us take the exterior derivative of the defining equation $\theta = A^{-1}dA$. For this we need to compute $d(A^{-1})$: Since

$$d(AA^{-1}) = 0$$

we have

$$dA \cdot A^{-1} + Ad(A^{-1}) = 0$$

or

$$d(A^{-1}) = -A^{-1}dA \cdot A^{-1}.$$

This is the generalization to matrices of the formula in elementary calculus for the derivative of $1/x$. Using this formula we get

$$d\theta = d(A^{-1}dA) = -(A^{-1}dA \cdot A^{-1}) \wedge dA = -A^{-1}dA \wedge A^{-1}dA$$

or the **Maurer-Cartan equation**

$$d\theta + \theta \wedge \theta = 0. \tag{1.8}$$

If we use commutator instead of multiplication we would write this as

$$d\theta + \frac{1}{2}[\theta, \theta] = 0. \tag{1.9}$$

The Maurer-Cartan equation is of central importance in geometry and physics, far more important than the Campbell-Baker-Hausdorff formula itself.

Suppose we have a map $g : \mathbf{R}^2 \to G$, with s, t coordinates on the plane. Pull θ back to the plane, so

$$g^*\theta = g^{-1}\frac{\partial g}{\partial s}ds + g^{-1}\frac{\partial g}{\partial t}dt$$

Define

$$\alpha = \alpha(s,t) := g^{-1}\frac{\partial g}{\partial s}$$

and

$$\beta := \beta(s,t) = g^{-1}\frac{\partial g}{\partial t}$$

so that

$$g^*\theta = \alpha ds + \beta dt.$$

Then collecting the coefficient of $ds \wedge dt$ in the Maurer Cartan equation gives

$$\frac{\partial \beta}{\partial s} - \frac{\partial \alpha}{\partial t} + [\alpha, \beta] = 0. \tag{1.10}$$

This is the version of the Maurer Cartan equation we shall use in our proof of the Campbell Baker Hausdorff formula. Of course this version is completely equivalent to the general version, since a two form is determined by its restriction to all two dimensional surfaces.

1.4 Proof of CBH from Maurer-Cartan.

Let $C(t)$ be a curve in the Lie algebra $\mathbf{g}$ and let us apply (1.10) to

$$g(s,t) := \exp[sC(t)]$$

so that

$$\begin{aligned}
\alpha(s,t) &= g^{-1}\frac{\partial g}{\partial s}\\
&= \exp[-sC(t)]\exp[sC(t)]C(t)\\
&= C(t)\\
\beta(s,t) &= g^{-1}\frac{\partial g}{\partial t}\\
&= \exp[-sC(t)]\frac{\partial}{\partial t}\exp[sC(t)] \text{ so by (1.10)}\\
\frac{\partial\beta}{\partial s} - C'(t) + [C(t),\beta] &= 0.
\end{aligned}$$

For fixed t consider the last equation as the differential equation (in s)

$$\frac{d\beta}{ds} = -(\text{ad } C)\beta + C', \quad \beta(0) = 0$$

where $C := C(t),\ C' := C'(t)$.

If we expand $\beta(s,t)$ as a formal power series in s (for fixed t):

$$\beta(s,t) = a_1 s + a_2 s^2 + a_3 s^3 + \cdots$$

and compare coefficients in the differential equation we obtain $a_1 = C'$, and

$$na_n = -(\text{ad } C)a_{n-1}$$

or

$$\beta(s,t) = sC'(t) + \frac{1}{2}s(-\text{ad } C(t))C'(t) + \cdots + \frac{1}{n!}s^n(-\text{ad } C(t))^{n-1}C'(t) + \cdots .$$

If we define

$$\phi(z) := \frac{e^z - 1}{z} = 1 + \frac{1}{2!}z + \frac{1}{3!}z^2 + \cdots$$

and set $s = 1$ in the expression we derived above for $\beta(s,t)$ we get

$$\exp(-C(t))\frac{d}{dt}\exp(C(t)) = \phi(-\text{ad } C(t))C'(t). \tag{1.11}$$

Now to the proof of the Campbell-Baker-Hausdorff formula. Suppose that A and B are chosen sufficiently near the origin so that

$$\Gamma = \Gamma(t) = \Gamma(t,A,B) := \log((\exp A)(\exp tB))$$

is defined for all $|t| \leq 1$. Then, as we remarked,

$$\exp \Gamma = \exp A \exp tB$$

so $\exp \mathrm{ad}\ \Gamma = (\exp \mathrm{ad}\ A)(\exp t\ \mathrm{ad}\ B)$ and hence

$$\mathrm{ad}\ \Gamma = \log\left((\exp \mathrm{ad}\ A)(\exp t\ \mathrm{ad}\ B)\right).$$

We have

$$\begin{aligned}
\frac{d}{dt}\exp\Gamma(t) &= \exp A \frac{d}{dt}\exp tB \\
&= \exp A(\exp tB)B \\
&= (\exp\Gamma(t)B \quad \text{so} \\
(\exp -\Gamma(t))\frac{d}{dt}\exp\Gamma(t) &= B \text{ and therefore} \\
\phi(-\mathrm{ad}\ \Gamma(t))\Gamma'(t) &= B \quad \text{by (1.11) so} \\
\phi(-\log\left((\exp \mathrm{ad}\ A)(\exp t\ \mathrm{ad}\ B)\right))\Gamma'(t) &= B.
\end{aligned}$$

Now for $|z - 1| < 1$

$$\begin{aligned}
\phi(-\log z) &= \frac{e^{-\log z} - 1}{-\log z} \\
&= \frac{z^{-1} - 1}{-\log z} \\
&= \frac{z - 1}{z \log z} \quad \text{so} \\
\psi(z)\phi(-\log z) &\equiv 1 \quad \text{where } \psi(z) := \frac{z \log z}{z - 1} \text{ so} \\
\Gamma'(t) &= \psi\left((\exp \mathrm{ad}\ A)(\exp t \mathrm{ad}\ B)\right) B.
\end{aligned}$$

This proves (1.4) and integrating from 0 to 1 proves (1.2).

1.5 The differential of the exponential and its inverse.

Once again, equation (1.11), which we derived from the Maurer-Cartan equation, is of significant importance in its own right, perhaps more than the use we made of it - to prove the Campbell-Baker-Hausdorff theorem. We will rewrite this equation in terms of more familiar geometric operations, but first some preliminaries:

The exponential map exp sends the Lie algebra $\mathbf{g}$ into the corresponding Lie group, and is a differentiable map. If $\xi \in \mathbf{g}$ we can consider the differential of exp at the point ξ:

$$d(\exp)_\xi : \mathbf{g} = T\mathbf{g}_\xi \to TG_{\exp \xi}$$

where we have identified **g** with its tangent space at ξ which is possible since **g** is a vector space. In other words, $d(\exp)_\xi$ maps the tangent space to **g** at the point ξ into the tangent space to G at the point $\exp(\xi)$. At $\xi = 0$ we have

$$d(\exp)_0 = \mathrm{id}$$

and hence, by the implicit function theorem, $d(\exp)_\xi$ is invertible for sufficiently small ξ. Now the Maurer-Cartan form, evaluated at the point $\exp\xi$ sends $TG_{\exp\xi}$ back to **g**:

$$\theta_{\exp\xi} : TG_{\exp\xi} \to \mathbf{g}.$$

Hence

$$\theta_{\exp\xi} \circ d(\exp)_\xi : \mathbf{g} \to \mathbf{g}$$

and is invertible for sufficiently small ξ. We claim that

$$\tau(\operatorname{ad}\xi) \circ \left(\theta_{\exp\xi} \circ d(\exp_\xi)\right) = \mathrm{id} \tag{1.12}$$

where τ is as defined above in (1.3). Indeed, we claim that (1.12) is an immediate consequence of (1.11).

Recall the definition (1.3) of the function τ as $\tau(z) = 1/\phi(-z)$. Multiply both sides of (1.11) by $\tau(\operatorname{ad} C(t))$ to obtain

$$\tau(\operatorname{ad} C(t)) \exp(-C(t)) \frac{d}{dt} \exp(C(t)) = C'(t). \tag{1.13}$$

Choose the curve C so that $\xi = C(0)$ and $\eta = C'(0)$. Then the chain rule says that

$$\frac{d}{dt} \exp(C(t))_{|t=0} = d(\exp)_\xi(\eta).$$

Thus

$$\left(\exp(-C(t)) \frac{d}{dt} \exp(C(t))\right)_{|t=0} = \theta_{\exp\xi} d(\exp)_\xi \eta,$$

the result of applying the Maurer-Cartan form θ (at the point $\exp(\xi)$) to the image of η under the differential of exponential map at $\xi \in \mathbf{g}$. Then (1.13) at $t = 0$ translates into (1.12). QED

1.6 The averaging method.

In this section we will give another important application of (1.10): For fixed $\xi \in \mathbf{g}$, the differential of the exponential map is a linear map from $\mathbf{g} = T_\xi(\mathbf{g})$ to $T_{\exp\xi}G$. The (differential of) left translation by $\exp\xi$ carries $T_{\exp\xi}(G)$ back to $T_eG = \mathbf{g}$. Let us denote this composite by $\exp_\xi^{-1} d(\exp)_\xi$. So

$$\theta_{\exp\xi} \circ d(\exp)_\xi = d\exp_\xi^{-1} d(\exp)_\xi : \quad \mathbf{g} \to \mathbf{g}$$

is a linear map. We claim that for any $\eta \in \mathbf{g}$

$$\exp_\xi^{-1} d(\exp)_\xi(\eta) = \int_0^1 \mathrm{Ad}_{\exp(-s\xi)} \eta \, ds. \tag{1.14}$$

We will prove this by applying(1.10) to

$$g(s,t) = \exp\left(t(\xi + s\eta)\right).$$

Indeed,

$$\beta(s,t) := g(s,t)^{-1}\frac{\partial g}{\partial t} = \xi + s\eta$$

so

$$\frac{\partial \beta}{\partial s} \equiv \eta$$

and

$$\beta(0,t) \equiv \xi.$$

The left hand side of (1.14) is $\alpha(0,1)$ where

$$\alpha(s,t) := g(s,t)^{-1}\frac{\partial g}{\partial s}$$

so we may use (1.10) to get an ordinary differential equation for $\alpha(0,t)$. Defining

$$\gamma(t) := \alpha(0,t),$$

(1.10) becomes

$$\frac{d\gamma}{dt} = \eta + [\gamma, \xi]. \tag{1.15}$$

For any $\zeta \in \mathbf{g}$,

$$\begin{aligned} \frac{d}{dt}\mathrm{Ad}_{\exp -t\xi}\zeta &= \mathrm{Ad}_{\exp -t\xi}[\zeta, \xi] \\ &= [\mathrm{Ad}_{\exp -t\xi}\zeta, \xi]. \end{aligned}$$

So for constant $\zeta \in \mathbf{g}$,

$$\mathrm{Ad}_{\exp -t\xi}\zeta$$

is a solution of the homogeneous equation corresponding to (1.15). So, by Lagrange's method of variation of constants, we look for a solution of (1.15) of the form

$$\gamma(t) = \mathrm{Ad}_{\exp -t\xi}\zeta(t)$$

and (1.15) becomes

$$\zeta'(t) = \mathrm{Ad}_{\exp t\xi}\eta$$

or

$$\gamma(t) = \mathrm{Ad}_{\exp -t\xi}\int_0^t \mathrm{Ad}_{\exp s\xi}\eta ds$$

is the solution of (1.15) with $\gamma(0) = 0$. Setting $s = 1$ gives

$$\gamma(1) = \mathrm{Ad}_{\exp -\xi}\int_0^1 \mathrm{Ad}_{\exp s\xi} ds$$

and replacing s by $1 - s$ in the integral gives (1.14).

1.7 The Euler MacLaurin Formula.

We pause to remind the reader of a different role that the τ function plays in mathematics. We have seen in (1.12) that τ enters into the inverse of the exponential map. In a sense, this formula is taking into account the non-commutativity of the group multiplication, so τ is helping to relate the non-commutative to the commutative.

But much earlier in mathematical history, τ was introduced to relate the discrete to the continuous: Let D denote the differentiation operator in one variable. Then if we think of D as the one dimensional vector field $\partial/\partial h$ it generates the one parameter group $\exp hD$ which consists of translation by h. In particular, taking $h = 1$ we have

$$\left(e^D f\right)(x) = f(x+1).$$

This equation is equally valid in a purely algebraic sense, taking f to be a polynomial and

$$e^D = 1 + D + \frac{1}{2}D^2 + \frac{1}{3!}D^3 + \cdots .$$

This series is infinite. But if p is a polynomial of degree d, then $D^k p = 0$ for $k > D$ so when applied to any polynomial, the above sum is really finite. Since

$$D^k e^{ah} = a^k e^{ah}$$

it follows that if F is any formal power series in one variable, we have

$$F(D)e^{ah} = F(a)e^{ah} \tag{1.16}$$

in the ring of power series in two variables. Of course, under suitable convergence conditions this is an equality of functions of h.

For example, the function $\tau(z) = z/(1 - e^{-z})$ converges for $|z| < 2\pi$ since $\pm 2\pi i$ are the closest zeros of the denominator (other than 0) to the origin. Hence

$$\tau\left(\frac{d}{dh}\right)\frac{e^{zh}}{z} = e^{zh}\frac{1}{1-e^{-z}} \tag{1.17}$$

holds for $0 < |z| < 2\pi$. Here the infinite order differential operator on the left is regarded as the limit of the finite order differential operators obtained by truncating the power series for τ at higher and higher orders.

Let $a < b$ be integers. Then for any non-negative values of h_1 and h_2 we have

$$\int_{a-h_1}^{b+h_2} e^{zx}dx = e^{h_2 z}\frac{e^{bz}}{z} - e^{-h_2 z}\frac{e^{az}}{z}$$

for $z \neq 0$. So if we set

$$D_1 := \frac{d}{dh_1}, \quad D_2 := \frac{d}{dh_2},$$

the for $0 < |z| < 2\pi$ we have

$$\tau(D_1)\tau(D_2)\int_{a-h_1}^{b+h_2} e^{zx}dx = \tau(z)e^{h_2 z}\frac{e^{bz}}{z} - \tau(-z)e^{-h_1 z}\frac{e^{az}}{z}$$

because $\tau(D_1)f(h_2) = f(h_2)$ when applied to any function of h_2 since the constant term in τ is one and all of the differentiations with respect to h_1 give zero.

Setting $h_1 = h_2 = 0$ gives

$$\tau(D_1)\tau(D_2)\int_{a-h_1}^{b+h_2} e^{zx}dx\Bigg|_{h_1=h_2=0} = \frac{e^{az}}{1-e^z} + \frac{e^{bz}}{1-e^{-z}}, 0 < |z| < 2\pi.$$

On then other hand, the geometric sum gives

$$\sum_{k=a}^{b} e^{kz} = e^{az}\left(1 + e^z + e^{2z} + \cdots + e^{(b-a)z}\right) = e^{az}\frac{1-e^{(b-a+1)z}}{1-e^z}$$

$$= \frac{e^{az}}{1-e^z} + \frac{e^{bz}}{1-e^{-z}}.$$

We have thus proved the following exact Euler-MacLaurin formula:

$$\tau(D_1)\tau(D_2)\int_{a-h_1}^{b+h_2} f(x)dx\Bigg|_{h_1=h_2=0} = \sum_{k=a}^{b} f(k), \tag{1.18}$$

where the sum on the right is over integer values of k and we have proved this formula for functions f of the form $f(x) = e^{zx}$, $0 < |z| < 2\pi$. It is also true when $z = 0$ by passing to the limit or by direct evaluation.

Repeatedly differentiating (1.18) (with $f(x) = e^{zx}$) with respect to z gives the corresponding formula with $f(x) = x^n e^{zx}$ and hence for all functions of the form $x \mapsto p(x)e^{zx}$ where p is a polynomial and $|z| < 2\pi$.

There is a corresponding formula with remainder for C^k functions.

1.8 The universal enveloping algebra.

We will now give an alternative (algebraic) version of the Campbell-Baker-Hausdorff theorem. It depends on several notions which are extremely important in their own right, so we pause to develop them.

A **universal algebra** of a Lie algebra L is a map $\epsilon : L \to UL$ where UL is an associative algebra with unit such that

1. ϵ is a Lie algebra homomorphism, i.e. it is linear and

$$\epsilon[x,y] = \epsilon(x)\epsilon(y) - \epsilon(y)\epsilon(x)$$

2. If A is any associative algebra with unit and $\alpha : L \to A$ is any Lie algebra homomorphism then there exists a unique homomorphism ϕ of associative algebras such that
$$\alpha = \phi \circ \epsilon.$$

It is clear that if UL exists, it is unique up to a unique isomorphism. So we may then talk of *the* universal algebra of L. We will call it the universal enveloping algebra and sometimes put in parenthesis, i.e. write $U(L)$.

In case $L = \mathbf{g}$ is the Lie algebra of left invariant vector fields on a group G, we may think of L as consisting of left invariant first order homogeneous differential operators on G. Then we may take UL to consist of all left invariant differential operators on G. In this case the construction of UL is intuitive and obvious. The ring of differential operators $\mathcal{D}$ on any manifold is filtered by degree: $\mathcal{D}^n$ consisting of those differential operators with total degree at most n. The quotient, $\mathcal{D}^n/\mathcal{D}^{n-1}$ consists of those homogeneous differential operators of degree n, i.e. homogeneous polynomials in the vector fields with function coefficients. For the case of left invariant differential operators on a group, these vector fields may be taken to be left invariant, and the function coefficients to be constant. In other words, $(UL)^n/(UL)^{n-1}$ consists of all symmetric polynomial expressions, homogeneous of degree n in L. This is the content of the Poincaré-Birkhoff-Witt theorem. In the algebraic case we have to do some work to get all of this. We first must construct $U(L)$.

1.8.1 Tensor product of vector spaces.

Let $E_1, \ldots, E_m$ be vector spaces and (f, F) a multilinear map $f : E_1 \times \cdots \times E_m \to F$. Similarly (g, G). If ℓ is a linear map $\ell : F \to G$, and $g = \ell \circ f$ then we say that ℓ is a morphism of (f, F) to (g, G). In this way we make the set of all (f, F) into a *category*. Want a universal object in this category; that is, an object with a unique morphism into every other object. So want a pair $(t, \mathcal{T})$ where $\mathcal{T}$ is a vector space, $t : E_1 \times \cdots \times E_m \to \mathcal{T}$ is a multilinear map, and for every (f, F) there is a unique linear map $\ell_f : \mathcal{T} \to F$ with
$$f = \ell_f \circ t$$
.

Uniqueness. By the universal property $t = \ell'_t \circ t'$, $t' = \ell'_t \circ t$ so $t = (\ell'_t \circ \ell_{t'}) \circ t$, but also $t = t \circ$id. So $\ell'_t \circ \ell_{t'} =$id. Similarly the other way. Thus $(t, \mathcal{T})$, if it exists, is unique up to a unique morphism. This is a standard argument valid in any category proving the uniqueness of "initial elements".

Existence. Let M be the free vector space on the symbols $x_1, \ldots, x_m$, $x_i \in E_i$. Let N be the subspace generated by all the
$$(x_1, \ldots, x_i + x_i', \ldots, x_m) - (x_1, \ldots, x_i, \ldots, x_m) - (x_1, \ldots, x_i', \ldots, x_m)$$
and all the
$$(x_1, \ldots, , ax_i, \ldots, x_m) - a(x_1, \ldots, x_i, \ldots, x_m)$$

for all $i = 1, \dots, m, x_i, x_i' \in E_i,\ a \in k$. Let $\mathcal{T} = M/N$ and

$$t((x_1, \dots, x_m)) = (x_1, \dots, x_m)/N.$$

This is universal by its very construction. QED

We introduce the notation

$$\mathcal{T} = \mathcal{T}(E_1 \times \cdots \times E_m) =: E_1 \otimes \cdots \otimes E_m.$$

The universality implies an isomorphism

$$(E_1 \otimes \cdots \otimes E_m) \otimes (E_{m+1} \otimes \cdots \otimes E_{m+n}) \cong E_1 \otimes \cdots \otimes E_{m+n}.$$

1.8.2 The tensor product of two algebras.

If A and B are algebras, they are they are vector spaces, so we can form their tensor product as vector spaces. We define a product structure on $A \otimes B$ by defining

$$(a_1 \otimes b_1) \cdot (a_2 \otimes b_2) := a_1 a_2 \otimes b_1 b_2.$$

It is easy to check that this extends to give an algebra structure on $A \otimes B$. In case A and B are associative algebras so is $A \otimes B$, and if in addition both A and B have unit elements, then $1_A \otimes 1_B$ is a unit element for $A \otimes B$. We will frequently drop the subscripts on the unit elements, for it is easy to see from the position relative to the tensor product sign the algebra to which the unit belongs. In other words, we will write the unit for $A \otimes B$ as $1 \otimes 1$. We have an isomorphism of A into $A \otimes B$ given by

$$a \mapsto a \otimes 1$$

when both A and B are associative algebras with units. Similarly for B. Notice that

$$(a \otimes 1) \cdot (1 \otimes b) = a \otimes b = (1 \otimes b) \cdot (a \otimes 1).$$

In particular, an element of the form $a \otimes 1$ commutes with an element of the form $1 \otimes b$.

1.8.3 The tensor algebra of a vector space.

Let V be a vector space. The **tensor algebra** of a vector space is the solution of the universal problem for maps α of V into an associative algebra: it consists of an algebra TV and a map $\iota : V \to TV$ such that ι is linear, and for any linear map $\alpha : V \to A$ where A is an associative algebra there exists a unique algebra homomorphism $\psi : TV \to A$ such that $\alpha = \psi \circ \iota$. We set

$$T^n V := V \otimes \cdots \otimes V \quad n-\text{factors}.$$

We define the multiplication to be the isomorphism

$$T^n V \otimes T^m V \to T^{n+m} V$$

obtained by "dropping the parentheses," i.e. the isomorphism given at the end of the last subsection. Then

$$TV := \bigoplus T^n V$$

(with T^0V the ground field) is a solution to this universal problem, and hence the unique solution.

1.8.4 Construction of the universal enveloping algebra.

If we take $V = L$ to be a Lie algebra, and let I be the two sided ideal in TL generated the elements $[x, y] - x \otimes y + y \otimes x$ then

$$UL := TL/I$$

is a universal algebra for L. Indeed, any homomorphism α of L into an associative algebra A extends to a unique algebra homomorphism $\psi : TL \to A$ which must vanish on I if it is to be a Lie algebra homomorphism.

1.8.5 Extension of a Lie algebra homomorphism to its universal enveloping algebra.

If $h : L \to M$ is a Lie algebra homomorphism, then the composition

$$\epsilon_M \circ h : L \to UM$$

induces a homomorphism

$$UL \to UM$$

and this assignment sending Lie algebra homomorphisms into associative algebra homomorphisms is functorial.

1.8.6 Universal enveloping algebra of a direct sum.

Suppose that: $L = L_1 \oplus L_2$, with $\epsilon_i : L_i \to U(L_i)$, and $\epsilon : L \to U(L)$ the canonical homomorphisms. Define

$$f : L \to U(L_1) \otimes U(L_2), \quad f(x_1 + x_2) = \epsilon_1(x_1) \otimes 1 + 1 \otimes \epsilon_2(x_2).$$

This is a homomorphism because x_1 and x_2 commute. It thus extends to a homomorphism

$$\psi : U(L) \to U(L_1) \otimes U(L_2).$$

Also,

$$x_1 \mapsto \epsilon(x_1)$$

is a Lie algebra homomorphism of $L_1 \to U(L)$ which thus extends to a unique algebra homomorphism

$$\phi_1 : U(L_1) \to U(L)$$

and similarly $\phi_2 : U(L_2) \to U(L)$. We have

$$\phi_1(x_1)\phi_2(x_2) = \phi_2(x_2)\phi_1(x_1), \quad x_1 \in L_1, x_2 \in L_2$$

since $[x_1, x_2] = 0$. As the $\epsilon_i(x_i)$ generate $U(L_i)$, the above equation holds with x_i replaced by arbitrary elements $u_i \in U(L_i), i = 1, 2$. So we have a homomorphism

$$\phi : U(L_1) \otimes U(L_2) \to U(L), \quad \phi(u_1 \otimes u_2) := \phi_1(u_1)\phi_2(u_2).$$

We have

$$\phi \circ \psi(x_1 + x_2) = \phi(x_1 \otimes 1) + \phi(1 \otimes x_2) = x_1 + x_2$$

so $\phi \circ \psi = \mathrm{id}$, on L and hence on $U(L)$ and

$$\psi \circ \phi(x_1 \otimes 1 + 1 \otimes x_2) = x_1 \otimes 1 + 1 \otimes x_2$$

so $\psi \circ \phi = \mathrm{id}$ on $L_1 \otimes 1 + 1 \otimes L_2$ and hence on $U(L_1) \otimes U(L_2)$. Thus

$$U(L_1 \oplus L_2) \cong U(L_1) \otimes U(L_2).$$

1.8.7 Bialgebra structure.

Consider the map $L \to U(L) \otimes U(L)$:

$$x \mapsto x \otimes 1 + 1 \otimes x.$$

Then

$$(x \otimes 1 + 1 \otimes x)(y \otimes 1 + 1 \otimes y) =$$

$$xy \otimes 1 + x \otimes y + y \otimes x + +1 \otimes xy,$$

and multiplying in the reverse order and subtracting gives

$$[x \otimes 1 + 1 \otimes x, y \otimes 1 + 1 \otimes y] = [x, y] \otimes 1 + 1 \otimes [x, y].$$

Thus the map $x \mapsto x \otimes 1 + 1 \otimes x$ determines an algebra homomorphism

$$\Delta : U(L) \to U(L) \otimes U(L).$$

Define

$$\varepsilon : U(L) \to k, \;\; \varepsilon(1) = 1, \;\; \varepsilon(x) = 0, x \in L$$

and extend as an algebra homomorphism. Then

$$(\varepsilon \otimes \mathrm{id})(x \otimes 1 + 1 \otimes x) = 1 \otimes x, \quad x \in L.$$

We identify $k \otimes L$ with L and so can write the above equation as

$$(\varepsilon \otimes \mathrm{id})(x \otimes 1 + 1 \otimes x) = x, \quad x \in L.$$

The algebra homomorphism

$$(\varepsilon \otimes \mathrm{id}) \circ \Delta : U(L) \to U(L)$$

is the identity (on 1 and on) L and hence is the identity. Similarly

$$(\mathrm{id} \otimes \varepsilon) \circ \Delta = \mathrm{id}.$$

A vector space C with a map $\Delta : C \to C \otimes C$, (called a **comultiplication**) and a map $\varepsilon : D \to k$ (called a **co-unit**) satisfying

$$(\varepsilon \otimes \mathrm{id}) \circ \Delta = \mathrm{id}$$

and

$$(\mathrm{id} \otimes \varepsilon) \circ \Delta = \mathrm{id}$$

is called a **co-algebra**. If C is an algebra and both Δ and ε are algebra homomorphisms, we say that C is a **bi-algebra** (sometimes shortened to "bigebra"). So we have proved that $(U(L), \Delta, \varepsilon)$ is a bialgebra.

Also

$$[(\Delta \otimes \mathrm{id}) \circ \Delta](x) = x \otimes 1 \otimes 1 + 1 \otimes x \otimes 1 + 1 \otimes 1 \otimes x = [(\mathrm{id} \otimes \Delta) \circ \ \Delta](x)$$

for $x \in L$ and hence for all elements of $U(L)$. Hence the comultiplication is is coassociative. (It is also co-commutative.)

1.9 The Poincaré-Birkhoff-Witt Theorem.

Suppose that V is a vector space made into a Lie algebra by declaring that all brackets are zero. Then the ideal I in TV defining $U(V)$ is generated by $x \otimes y - y \otimes x$, and the quotient TV/I is just the symmetric algebra, SV. So the universal enveloping algebra of the trivial Lie algebra is the symmetric algebra.

For any Lie algebra L define U_nL to be the subspace of UL generated by products of at most n elements of L, i.e. by all products

$$\epsilon(x_1) \cdots \epsilon(x_m), \qquad m \leq n.$$

For example,,

$$U_0L = k, \text{ the ground field}$$

and

$$U_1L = k \oplus \epsilon(L).$$

We have

$$U_0L \subset U_1L \subset \cdots \subset U_nL \subset U_{n+1}L \subset \cdots$$

and

$$U_mL \cdot U_nL \subset U_{m+n}L.$$

We define

$$\mathrm{gr}_n\, UL := U_nL/U_{n-1}L$$

and

$$\mathrm{gr}\, UL := \bigoplus \mathrm{gr}_n\, UL$$

with the multiplication

$$\mathrm{gr}_m\, UL \times \mathrm{gr}_n\, UL \to \mathrm{gr}_{m+n}\, UL$$

induced by the multiplication on UL.

If $a \in U_nL$ we let $\overline{a} \in \mathrm{gr}_n\, UL$ denote its image by the projection $U_nL \to U_nL/U_{n-1}L = \mathrm{gr}_n\, UL$. We may write a as a sum of products of at most n elements of L:

$$a = \sum_{m_\mu \le n} c_\mu \epsilon(x_{\mu,1}) \cdots \epsilon(x_{\mu,m_\mu}).$$

Then $\overline{a}$ can be written as the corresponding homogeneous sum

$$\overline{a} = \sum_{m_\mu = n} c_\mu \overline{\epsilon(x_{\mu,1})} \cdots \overline{\epsilon(x_{\mu,m_\mu})}.$$

In other words, as an algebra, $\mathrm{gr}\, UL$ is generated by the elements $\overline{\epsilon(x)}, \quad x \in L$. But all such elements commute. Indeed, for $x, y \in L$,

$$\epsilon(x)\epsilon(y) - \epsilon(y)\epsilon(x) = \epsilon([x,y]).$$

by the defining property of the universal enveloping algebra. The right hand side of this equation belongs to U_1L. Hence

$$\overline{\epsilon(x)\epsilon(y)} - \overline{\epsilon(y)\epsilon(x)} = 0$$

in $\mathrm{gr}_2\, UL$. This proves that $\mathrm{gr}\, UL$ is *commutative*. Hence, by the universal property of the symmetric algebra, there exists a unique algebra homomorphism

$$\mathbf{w} : SL \to \mathrm{gr}\, UL$$

extending the linear map

$$L \to \mathrm{gr}\, UL, \quad x \mapsto \overline{\epsilon(x)}.$$

Since the $\overline{\epsilon(x)}$ generate $\mathrm{gr}\, UL$ as an algebra, we know that this map is surjective. The **Poincaré-Birkhoff-Witt theorem** asserts that

$$\mathbf{w} : SL \to \mathrm{gr}\, UL \quad \text{is an isomorphism.} \tag{1.19}$$

Suppose that we choose a basis $x_i,\ i \in I$ of L where I is a totally ordered set. Since

$$\overline{\epsilon(x_i)\epsilon(x_j)} = \overline{\epsilon(x_j)\epsilon(x_i)}$$

we can rearrange any product of $\overline{\epsilon(x_i)}$ so as to be in increasing order. This shows that the elements

$$x_M := \epsilon(x_{i_1}) \cdots \epsilon(x_{i_m}), \quad M := (i_1, \ldots, i_m)\ i_1 \le \cdots i_m$$

span UL as a vector space. We claim that (1.19) is equivalent to

Theorem 1 Poincaré-Birkhoff-Witt. *The elements x_M form a basis of UL.*

Proof that (1.19) is equivalent to the statement of the theorem. For any expression x_M as above, we denote its length by $\ell(M) = m$. The elements $\overline{x_M}$ are the images under $\mathbf{w}$ of the monomial basis in $S_m(L)$. As we know that $\mathbf{w}$ is surjective, equation (1.19) is equivalent to the assertion that $\mathbf{w}$ is injective. This amounts to the non-existence of a relation of the form

$$\sum_{\ell(M)=n} c_M x_M = \sum_{\ell(M)<n} c_M x_M$$

with some non-zero coefficients on the left hand side. But any non-trivial relation between the x_M can be rewritten in the above form by moving the terms of highest length to one side. QED

We now turn to the proof of the theorem:

Let V be the vector space with basis z_M where M runs over all ordered sequences $i_1 \leq i_2 \leq \cdots \leq i_n$. (Recall that we have chosen a well ordering on I and that the $x_i\ {}_{i\in I}$ form a basis of L.)

Furthermore, the empty sequence, $z_\emptyset$ is allowed, and we will identify the symbol $z_\emptyset$ with the number $1 \in k$. If $i \in I$ and $M = (i_1, \ldots, i_n)$ we write $i \leq M$ if $i \leq i_1$ and then let (i, M) denote the ordered sequence $(i, i_1, \ldots, i_n)$. In particular, we adopt the convention that if $M = \emptyset$ is the empty sequence then $i \leq M$ for all i in which case $(i, M) = (i)$. Recall that if $M = (i_1, \ldots, i_n))$ we set $\ell(M) = n$ and call it the length of M. So, for example, $\ell(i, M) = \ell(M) + 1$ if $i \leq M$.

Lemma 1 *We can make V into an L module in such a way that*

$$x_i z_M = z_{iM} \quad \textit{whenever } i \leq M. \tag{1.20}$$

Proof of lemma. We will inductively define a map

$$L \times V \to V, \quad (x, v) \mapsto xv$$

and then show that it satisfies the equation

$$xyv - yxv = [x, y]v, \quad x, y \in L, \quad v \in V, \tag{1.21}$$

which is the condition that makes V into an L module. Our definition will be such that (1.20) holds. In fact, we will define $x_i z_M$ inductively on $\ell(M)$ and on i. So we start by defining

$$x_i z_\emptyset = z_{(i)}$$

which is in accordance with (1.20). This defines $x_i z_M$ for $\ell(M) = 0$. For $\ell(M) = 1$ we define

$$x_i z_{(j)} = z_{(i,j)} \quad \text{if } i \leq j$$

while if $i > j$ we set

$$x_i z_{(j)} = x_j z_{(i)} + [x_i, x_j] z_\emptyset = z_{(j,i)} + \sum c_{ij}^k z_{(k)}$$

where

$$[x_i, x_j] = \sum c_{ij}^k x_k$$

is the expression for the Lie bracket of x_i with x_j in terms of our basis. These c_{ij}^k are known as the **structure constants** of the Lie algebra, L in terms of the given basis. Notice that the first of these two cases is consistent with (and forced on us) by (1.20) while the second is forced on us by (1.21). We now have defined $x_i z_M$ for all i and all M with $\ell(M) \leq 1$, and we have done so in such a way that (1.20) holds, and (1.21) holds where it makes sense (i.e. for $\ell(M) = 0$).

So suppose that we have defined $x_j z_N$ for all j if $\ell(N) < \ell(M)$ and for all $j < i$ if $\ell(N) = \ell(M)$ in such a way that

$$x_j z_N \text{ is a linear combination of } z_L\text{'s with } \ell(L) \leq \ell(N) + 1 \quad (*).$$

We then define

$$x_i z_M = \begin{cases} z_{iM} \text{ if } i \leq M \\ x_j(x_i z_N) + [x_i, x_j] z_N \text{ if } M = (jN) \text{ with } i > j. \end{cases} \tag{1.22}$$

This makes sense since $x_i z_N$ is already defined as a linear combination of z_L's with $\ell(L) \leq \ell(N) + 1 = \ell(M)$ and because $[x_i, x_j]$ can be written as a linear combination of the x_k as above. Furthermore $(*)$ holds with j and N replaced by M. Furthermore, (1.20) holds by construction. We must check (1.21). By linearity, this means that we must show that

$$x_i x_j z_N - x_j x_i z_N = [x_i, x_j] z_N.$$

If $i = j$ both sides are zero. Also, since both sides are anti-symmetric in i and j, we may assume that $i > j$. If $j \leq N$ and $i > j$ then this equation holds by definition. So we need only deal with the case where $j \not\leq N$ which means that $N = (kP)$ with $k \leq P$ and $i > j > k$. So we have, by definition,

$$\begin{aligned} x_j z_N &= x_j z_{(kP)} \\ &= x_j x_k z_P \\ &= x_k x_j z_P + [x_j, x_k] z_P. \end{aligned}$$

Now if $j \leq P$ then $x_j z_P = z_{(jP)}$ and $k < (jP)$. If $j \not\leq P$ then $x_j z_P = z_Q + w$ where still $k \leq Q$ and w is a linear combination of elements of length $< \ell(N)$. So we know that (1.21) holds for $x = x_i, y = x_k$ and $v = z_{(jP)}$ (if $j \leq P$) or $v = z_Q$ (otherwise). Also, by induction, we may assume that we have verified (1.21) for all N' of length $< \ell(N)$. So we may apply (1.21) to $x = x_i, \ \ y = x_k$ and $v = x_j z_P$ and also to $x = x_i, \ y = [x_j, x_k], \ v = z_P$. So

$$x_i x_j z_N = x_k x_i x_j z_P + [x_i, x_k] x_j z_P + [x_j, x_k] x_i z_P + [x_i, [x_j, x_k]] z_P.$$

Similarly, the same result holds with i and j interchanged. Subtracting this interchanged version from the preceding equation the two middle terms from each equation cancel and we get

$$\begin{aligned}
(x_i x_j - x_j x_i) z_N &= x_k(x_i x_j - x_j x_i) z_P + ([x_i,[x_j,x_k]] - [x_j,[x_i,x_k])z_P \\
&= x_k[x_i,x_j]z_P + ([x_i,[x_j,x_k]] - [x_j,[x_i,x_k])z_P \\
&= [x_i,x_j]x_k z_P + ([x_k,[x_i,x_j]] + [x_i,[x_j,x_k]] - [x_j,[x_i,x_k])z_P \\
&= [x_i,x_j]z_N.
\end{aligned}$$

(In passing from the second line to the third we used (1.21) applied to z_P (by induction) and from the third to the last we used the antisymmetry of the bracket and Jacobi's equation.)QED

Proof of the PBW theorem. We have made V into an L and hence into a $U(L)$ module. By construction, we have, inductively,

$$x_M z_\emptyset = z_M.$$

But if

$$\sum c_M x_M = 0$$

then

$$0 = \sum c_M z_M = \left(\sum c_M x_M\right) z_\emptyset$$

contradicting the fact the the z_M are independent. QED

In particular, the map $\epsilon : L \to U(L)$ is an injection, and so we may identify L as a subspace of $U(L)$.

1.10 Primitives.

An element x of a bialgebra is called **primitive** if

$$\Delta(x) = x \otimes 1 + 1 \otimes x.$$

So the elements of L are primitives in $U(L)$.

We claim that *these are the only primitives.*

First prove this for the case L is abelian so $U(L) = S(L)$. Then we may think of $S(L) \otimes S(L)$ as polynomials in twice the number of variables as those of $S(L)$ and

$$\Delta(f)(u, v) = f(u + v).$$

The condition of being primitive says that

$$f(u + v) = f(u) + f(v).$$

Taking homogeneous components, the same equality holds for each homogeneous component. But if f is homogeneous of degree n, taking $u = v$ gives

$$2^n f(u) = 2f(u)$$

so $f = 0$ unless $n = 1$.

Taking gr, this shows that for any Lie algebra the primitives are contained in $U_1(L)$. But

$$\Delta(c + x) = c(1 \otimes 1) + x \otimes 1 + 1 \otimes x$$

so the condition on primitivity requires $c = 2c$ or $c = 0$. QED

1.11 Free Lie algebras

1.11.1 Magmas and free magmas on a set

A set M with a map:

$$M \times M \to M, \quad (x, y) \mapsto xy$$

is called a **magma**. Thus a magma is a set with a binary operation with no axioms at all imposed.

Let X be any set. Define X_n inductively by $X_1 := X$ and

$$X_n = \coprod_{p+q=n} X_p \times X_q$$

for $n \geq 2$. Thus X_2 consists of all expressions ab where a and b are elements of X. (We write ab instead of (a, b).) An element of X_3 is either an expression of the form $(ab)c$ or an expression of the form $a(bc)$. An element of X_4 has one out of five forms: $a((bc)d), a(b(cd)), ((ab)(cd)), ((ab)c)d$ or $(a(bc))d$.

Set

$$M_X := \coprod_{n=1}^{\infty} X_n.$$

An element $w \in M_X$ is called a non-associative word, and its length $\ell(w)$ is the unique n such that $w \in X_n$. We have a "multiplication" map $M_X \times M_X$ given by the inclusion

$$X_p \times X_q \hookrightarrow X_{p+q}.$$

Thus the multiplication on M_X is concatenation of non-associative words.

If N is any magma, and $f : X \to N$ is any map, we define $F : M_X \to N$ by $F = f$ on X_1, by

$$F : X_2 \to N, \quad F(ab) = f(a)f(b)$$

and inductively

$$F : X_p \times X_q \to N, \quad F(uv) = F(u)F(v).$$

Any element of X_n has a unique expression as uv where $u \in X_p$ and $v \in X_q$ for a unique (p, q) with $p + q = n$, so this inductive definition is valid.

It is clear that F is a magna homomorphism and is uniquely determined by the original map f. Thus M_X is the "free magma on X" or the "universal magma on X" in the sense that it is the solution to the universal problem associated to a map from X to any magma.

Let A_X be the vector space of finite formal linear combinations of elements of M_X. So an element of A_X is a finite sum $\sum c_m m$ with $m \in M_X$ and c_m in the ground field. The multiplication in M_X extends by bi-linearity to make A_X into an algebra. If we are given a map $X \to B$ where B is any algebra, we get a unique magna homomorphism $M_X \to B$ extending this map (where we think of B as a magma) and then a unique algebra map $A_X \to B$ extending this map by linearity.

Notice that the algebra A_X is graded since every element of M_X has a length and the multiplication on M_X is graded. Hence A_X is the free algebra on X in the sense that it solves the universal problem associated with maps of X to algebras.

1.11.2 The Free Lie Algebra L_X.

In A_X let I be the two-sided ideal generated by all elements of the form $aa,\ a \in A_X$ and $(ab)c + (bc)a + (ca)b, \quad a, b, c \in A_X$. We set

$$L_X := A_X/I$$

and call L_X the free Lie algebra on X. Any map from X to a Lie algebra L extends to a unique algebra homomorphism from L_X to L.

We claim that the ideal I defining L_X is graded. This means that if $a = \sum a_n$ is a decomposition of an element of I into its homogeneous components, then each of the a_n also belong to I. To prove this, let $J \subset I$ denote the set of all $a = \sum a_n$ with the property that all the homogeneous components a_n belong to I. Clearly J is a two sided ideal. We must show that $I \subset J$. For this it is enough to prove the corresponding fact for the generating elements. Clearly if

$$a = \sum a_p,\ \ b = \sum b_q, c = \sum c_r$$

then

$$(ab)c + (bc)a + (ca)b = \sum_{p,q,r} \left((a_p b_q)c_r + (b_q c_r)a_p + (c_r a_p)b_q\right).$$

But also if $x = \sum x_m$ then

$$x^2 = \sum x_n^2 + \sum_{m<n} (x_m x_n + x_n x_m)$$

and

$$x_m x_n + x_n x_m = (x_m + x_n)^2 - x_m^2 - x_n^2 \in I$$

so $I \subset J$.

The fact that I is graded means that L_X inherits the structure of a graded algebra.

1.11.3 The free associative algebra Ass(X).

Let V_X be the vector space of all finite formal linear combinations of elements of X. Define

$$\mathrm{Ass}_X = T(V_X),$$

the tensor algebra of V_X. Any map of X into an associative algebra A extends to a unique linear map from V_X to A and hence to a unique algebra homomorphism from Ass_X to A. So Ass_X is the free associative algebra on X.

We have the maps $X \to L_X$ and $\epsilon : L_X \to U(L_X)$ and hence their composition maps X to the associative algebra $U(L_X)$ and so extends to a unique homomorphism

$$\Psi : \mathrm{Ass}_X \to U(L_X).$$

On the other hand, the commutator bracket gives a Lie algebra structure to Ass_X and the map $X \to \mathrm{Ass}_X$ thus give rise to a Lie algebra homomorphism

$$L_X \to \mathrm{Ass}_X$$

which determines an associative algebra homomorphism

$$\Phi : U(L_X) \to \mathrm{Ass}_X.$$

both compositions $\Phi \circ \Psi$ and $\Psi \circ \Phi$ are the identity on X and hence, by uniqueness, the identity everywhere. We obtain the important result that $U(L_X)$ and Ass_X are canonically isomorphic:

$$U(L_X) \cong \mathrm{Ass}_X. \tag{1.23}$$

Now the Poincaré-Birkhoff -Witt theorem guarantees that the map $\epsilon : L_X \to U(L_X)$ is injective. So under the above isomorphism, the map $L_X \to \mathrm{Ass}_X$ is injective. On the other hand, by construction, the map $X \to V_X$ induces a surjective Lie algebra homomorphism from L_X into the Lie subalgebra of Ass_X generated by X. So we see that the under the isomorphism (1.23) $L_X \subset U(L_X)$ is mapped isomorphically onto the Lie subalgebra of Ass_X generated by X.

Now the map

$$X \to \mathrm{Ass}_X \otimes \mathrm{Ass}_X, \quad x \mapsto x \otimes 1 + 1 \otimes x$$

extends to a unique algebra homomorphism

$$\Delta : \mathrm{Ass}_X \to \mathrm{Ass}_X \otimes \mathrm{Ass}_X.$$

Under the identification (1.23) this is none other than the map

$$\Delta : U(L_X) \to U(L_X) \otimes U(L_X)$$

and hence we conclude that L_X is the set of primitive elements of Ass_X:

$$L_X = \{w \in \mathrm{Ass}_X | \Delta(w) = w \otimes 1 + 1 \otimes w.\} \tag{1.24}$$

under the identification (1.23).

1.12 Algebraic proof of CBH and explicit formulas.

We recall our constructs of the past few sections: X denotes a set, L_X the free Lie algebra on X and Ass_X the free associative algebra on X so that Ass_X may be identified with the universal enveloping algebra of L_X. Since Ass_X may be identified with the non-commutative polynomials indexed by X, we may consider its completion, F_X, the algebra of formal power series indexed by X. Since the free Lie algebra L_X is graded we may also consider its completion which we shall denote by $\mathbf{L}_X$. Finally let m denote the ideal in F_X generated by X. The maps

$$\exp : m \to 1+m, \quad \log : 1+m \to m$$

are well defined by their formal power series and are mutual inverses. (There is no convergence issue since everything is within the realm of formal power series.) Furthermore exp is a bijection of the set of $\alpha \in m$ satisfying $\Delta\alpha = \alpha \otimes 1 + 1 \otimes \alpha$ to the set of all $\beta \in 1+m$ satisfying $\Delta\beta = \beta \otimes \beta$.

1.12.1 Abstract version of CBH and its algebraic proof.

In particular, since the set $\{\beta \in 1+m | \Delta\beta = \beta \otimes \beta\}$ forms a group, we conclude that for any $A, B \in \mathbf{L}_X$ there exists a $C \in \mathbf{L}_X$ such that

$$\exp C = (\exp A)(\exp B).$$

This is the abstract version of the Campbell-Baker-Hausdorff formula. It depends basically on two algebraic facts: That the universal enveloping algebra of the free Lie algebra is the free associative algebra, and that the set of primitive elements in the universal enveloping algebra (those satisfying $\Delta\alpha = \alpha \otimes 1 + 1 \otimes \alpha$) is precisely the original Lie algebra.

1.12.2 Explicit formula for CBH.

Define the map

$$\Phi : m \cap \mathrm{Ass}_X \to L_X,$$

$$\Phi(x_1 \dots x_n) := [x_1, [x_2, \dots, [x_{n-1}, x_n] \cdots] = \mathrm{ad}(x_1) \cdots \mathrm{ad}(x_{n-1})(x_n),$$

and let $\Theta : \mathrm{Ass}_X \to \mathrm{End}(L_X)$ be the algebra homomorphism extending the Lie algebra homomorphism $\mathrm{ad} : L_X \to \mathrm{End}(L_X)$. We claim that

$$\Phi(uv) = \Theta(u)\Phi(v), \quad \forall\ u \in \mathrm{Ass}_X,\ v \in m \cap \mathrm{Ass}_X. \tag{1.25}$$

Proof. It is enough to prove this formula when u is a monomial, $u = x_1 \cdots x_n$. We do this by induction on n. For $n = 0$ the assertion is obvious and for $n = 1$

it follows from the definition of Φ. Suppose $n > 1$. Then

$$\begin{aligned}\Phi(x_1 \cdots x_n v) &= \Theta(x_1)\Phi(x_2 \cdots x_n v)\\ &= \Theta(x_1)\Theta(x_2 \dots x_n)\Phi(v)\\ &= \Theta(x_1 \cdots x_n)\Phi(v).\ \text{QED}\end{aligned}$$

Let L_X^n denote the $n-$th graded component of L_X. So L_X^1 consists of linear combinations of elements of X, L_X^2 is spanned by all brackets of pairs of elements of X, and in generalL_X^n is spanned by elements of the form

$$[u, v],\ u \in L_X^p,\ v \in L_X^q,\ p + q = n.$$

We claim that

$$\Phi(u) = nu \ \ \forall\ u \in L_X^n. \tag{1.26}$$

For $n = 1$ this is immediate from the definition of Φ. So by induction it is enough to verify this on elements of the form $[u, v]$ as above. We have

$$\begin{aligned}\Phi([u, v]) &= \Phi(uv - vu)\\ &= \Theta(u)\Phi(v) - \Theta(v)\Phi(u)\\ &= q\Theta(u)v - p\Theta(v)u \ \ \text{by induction}\\ &= q[u, v] - p[v, u]\\ &\qquad \text{since } \Theta(w) = \mathrm{ad}(w) \text{ for } w \in L_X\\ &= (p + q)[u, v] \quad \text{QED.}\end{aligned}$$

We can now write down an explicit formula for the $n-$th term in the Campbell-Baker-Hausdorff expansion. Consider the case where X consists of two elements $X = \{x, y\},\ x \neq y$. Let us write

$$z = \log\left((\exp x)(\exp y)\right) \quad z \in \mathbf{L}_X, \quad z = \sum_1^\infty z_n(x, y).$$

We want an explicit expression for $z_n(x, y)$. We know that

$$z_n = \frac{1}{n}\Phi(z_n)$$

and z_n is a sum of non-commutative monomials of degree n in x and y. Now

$$\begin{aligned}
(\exp x)(\exp y) &= \left(\sum_{p=0}^{\infty}\frac{x^p}{p!}\right)\left(\sum_{q=0}^{\infty}\frac{y^q}{q!}\right)\\
&= 1+\sum_{p+q\geq 1}\frac{x^py^q}{p!q!} \quad \text{so}\\
z &= \log((\exp x)(\exp y))\\
&= \sum_{m=1}^{\infty}\frac{(-1)^{m+1}}{m}\left(\sum_{p+q\geq 1}\frac{x^py^q}{p!q!}\right)^m\\
&= \sum_{p_i+q_i\geq 1}\frac{(-1)^{m+1}}{m}\frac{x^{p_1}y^{q_1}x^{p_2}y^{q_2}\cdots x^{p_m}y^{q_m}}{p_1!q_1!\cdots p_m!q_m!}.
\end{aligned}$$

We want to apply $\frac{1}{n}\Phi$ to the terms in this last expression which are of total degree n so as to obtain z_n. So let us examine what happens when we apply Φ to an expression occurring in the numerator: If $q_m \geq 2$ we get 0 since we will have $\operatorname{ad}(y)(y) = 0$. Similarly we will get 0 if $q_m = 0, p_m \geq 2$. Hence the only terms which survive are those with $q_m = 1$ or $q_m = 0, p_m = 1$. Accordingly we decompose z_n into these two types:

$$z_n = \frac{1}{n}\sum_{p+q=n}(z'_{p,q} + z''_{p,q}), \tag{1.27}$$

where

$$z'_{p,q} = \sum \frac{(-1)^{m+1}}{m}\frac{\operatorname{ad}(x)^{p_1}\operatorname{ad}(y)^{q_1}\cdots\operatorname{ad}(x)^{p_m}y}{p_1!q_1!\cdots p_m!} \quad \text{summed over all}$$

$$p_1+\cdots+p_m = p,\ q_1+\cdots+q_{m-1} = q-1,\quad q_i+p_i\geq 1,\ p_m\geq 1$$

and

$$z''_{p,q} = \sum \frac{(-1)^{m+1}}{m}\frac{\operatorname{ad}(x)^{p_1}\operatorname{ad}(y)^{q_1}\cdots\operatorname{ad}(y)^{q_{m-1}}(x)}{p_1!q_1!\cdots q_{m-1}!} \quad \text{summed over}$$

$$p_1+\cdots+p_{m-1} = p-1,\quad q_1+\cdots+q_{m-1} = q,$$
$$p_i+q_i\geq 1\ (i=1,\ldots,m-1)\ q_{m-1}\geq 1.$$

The first four terms are:

$$\begin{aligned}
z_1(x,y) &= x+y\\
z_2(x,y) &= \frac{1}{2}[x,y]\\
z_3(x,y) &= \frac{1}{12}[x,[x,y]] + \frac{1}{12}[y,[y,x]]\\
z_4(x,y) &= \frac{1}{24}[x,[y,[x,y]]].
\end{aligned}$$

Chapter 2

$sl(2)$ and its Representations.

In this chapter (and in most of the succeeding chapters) all Lie algebras and vector spaces are over the complex numbers.

2.1 Low dimensional Lie algebras.

Any one dimensional Lie algebra must be commutative, since $[X, X] = 0$ in any Lie algebra.

If $\mathbf{g}$ is a two dimensional Lie algebra, say with basis X, Y then $[aX+bY, cX+dY] = (ad - bc)[X, Y]$, so that there are two possibilities: $[X, Y] = 0$ in which case $\mathbf{g}$ is commutative, or $[X, Y] \neq 0$, call it B, and the Lie bracket of any two elements of $\mathbf{g}$ is a multiple of B. So if C is not a multiple of B, we have $[C, B] = cB$ for some $c \neq 0$, and setting $A = c^{-1}C$ we get a basis A, B of $\mathbf{g}$ with the bracket relations

$$[A, B] = B.$$

This is an interesting Lie algebra; it is the Lie algebra of the group of all affine transformations of the line, i.e. all transformations of the form

$$x \mapsto ax + b, \quad a \neq 0.$$

For this reason it is sometimes called the "$ax + b$ group". Since

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} ax + b \\ 1 \end{pmatrix}$$

we can realize the group of affine transformations of the line as a group of two by two matrices. Writing

$$a = \exp tA, \quad b = tB$$

so that
$$\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} = \exp t \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \exp t \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix}$$
we see that our algebra $\mathbf{g}$ with basis A, B and $[A, B] = B$ is indeed the Lie algebra of the $ax + b$ group.

In a similar way, we could list all possible three dimensional Lie algebras, by first classifying them according to $\dim[\mathbf{g}, \mathbf{g}]$ and then analyzing the possibilities for each value of this dimension. Rather than going through all the details, we list the most important examples of each type. If $\dim[\mathbf{g}, \mathbf{g}] = 0$ the algebra is commutative so there is only one possibility.

A very important example arises when $\dim[\mathbf{g}, \mathbf{g}] = 1$ and that is the Heisenberg algebra, with basis P, Q, Z and bracket relations

$$[P, Q] = Z, \quad [Z, P] = [Z, Q] = 0.$$

Up to constants (such as Planck's constant and i) these are the famous Heisenberg commutation relations. Indeed, we can realize this algebra as an algebra of operators on functions of one variable x: Let $P = D =$ differentiation, let Q consist of multiplication by x. Since, for any function $f = f(x)$ we have

$$D(xf) = f + xf'$$

we see that $[P, Q] = \mathrm{id}$, so setting $Z = \mathrm{id}$, we obtain the Heisenberg algebra.

As an example with $\dim[\mathbf{g}, \mathbf{g}] = 2$ we have (the complexification of) the Lie algebra of the group of Euclidean motions in the plane. Here we can find a basis h, x, y of $\mathbf{g}$ with brackets given by

$$[h, x] = y, \quad [h, y] = -x, \quad [x, y] = 0.$$

More generally we could start with a commutative two dimensional algebra and adjoin an element h with $\operatorname{ad} h$ acting as an arbitrary linear transformation, A of our two dimensional space.

The item of study of this chapter is the algebra $sl(2)$ of **all two by two matrices of trace zero**, where $[\mathbf{g}, \mathbf{g}] = \mathbf{g}$.

2.2 $sl(2)$ and its irreducible representations.

Indeed $sl(2)$ is spanned by the matrices:

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

They satisfy
$$[h, e] = 2e, \quad [h, f] = -2f, \quad [e, f] = h.$$

Thus every element of $sl(2)$ can be expressed as a sum of brackets of elements of $sl(2)$, in other words
$$[sl(2), sl(2)] = sl(2).$$

The bracket relations above are also satisfied by the matrices

$$\rho_2(h) := \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \ \rho_2(e) := \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \ \rho_2(f) := \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix},$$

the matrices

$$\rho_3(h) := \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}, \ \rho_3(e) := \begin{pmatrix} 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \ \rho_3(f) := \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix},$$

and, more generally, the $(n+1)\times(n+1)$ matrices given by

$$\rho_n(h) := \begin{pmatrix} n & 0 & \cdots & \cdots & 0 \\ 0 & n-2 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -n+2 & 0 \\ 0 & 0 & \cdots & \cdots & -n \end{pmatrix}, \ \rho_n(e) = \begin{pmatrix} 0 & n & \cdots & \cdots & 0 \\ 0 & 0 & n-1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix},$$

$$\rho_n(f) := \begin{pmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 1 & & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & n & 0 \end{pmatrix}.$$

These representations of $sl(2)$ are all irreducible, as is seen by successively applying $\rho_n(e)$ to any non-zero vector until a vector with non-zero element in the first position and all other entries zero is obtained. Then keep applying $\rho_n(f)$ to fill up the entire space.

These are all the finite dimensional irreducible representations of $sl(2)$ as can be seen as follows: In $U(sl(2))$ we have

$$[h, f^k] = -2kf^k, \quad [h, e^k] = 2ke^k \tag{2.1}$$

$$[e, f^k] = -k(k-1)f^{k-1} + kf^{k-1}h. \tag{2.2}$$

Equation (2.1) follows from the fact that bracketing by any element is a derivation and the fundamental relations in $sl(2)$. Equation (2.2) is proved by induction: For $k = 1$ it is true from the defining relations of $sl(2)$. Assuming it for k, we have

$$\begin{aligned} [e, f^{k+1}] &= [e, f]f^k + f[e, f^k] \\ &= hf^k - k(k-1)f^k + kf^k h \\ &= [h, f^k] + f^k h - k(k-1)f^k + kf^k h \\ &= -2kf^k - k(k-1)f^k + (k+1)f^k h \\ &= -(k+1)kf^k + (k+1)f^k h. \end{aligned}$$

We may rewrite (2.2) as

$$\left[e, \frac{1}{k!}f^k\right] = (-k+1)\frac{1}{(k-1)!}f^{k-1} + \frac{1}{(k-1)!}f^{k-1}h. \tag{2.3}$$

In any finite dimensional module V, the element h has at least one eigenvector. This follows from the fundamental theorem of algebra which assert that any polynomial has at least one root; in particular the characteristic polynomial of any linear transformation on a finite dimensional space has a root. So there is a vector w such that $hw = \mu w$ for some complex number μ. Then

$$h(ew) = [h, e]w + ehw = 2ew + \mu ew = (\mu + 2)(ew).$$

Thus ew is again an eigenvector of h, this time with eigenvalue $\mu+2$. Successively applying e yields a vector v_λ such that

$$hv_\lambda = \lambda v_\lambda, \quad ev_\lambda = 0. \tag{2.4}$$

Then $U(sl(2))v_\lambda$ is an invariant subspace, hence all of V. We say that v is a **cyclic vector** for the action of $\mathbf{g}$ on V if $U(\mathbf{g})v = V$,

We are thus led to study all modules for $sl(2)$ with a cyclic vector v_λ satisfying (2.4). In any such space the elements

$$\frac{1}{k!}f^k v_\lambda$$

span, and are eigenspaces of h of weight $\lambda - 2k$. For any $\lambda \in \mathbf{C}$ we can construct such a module as follows: Let $\mathbf{b}_+$ denote the subalgebra of $sl(2)$ generated by h and e. Then $U(\mathbf{b}_+)$, the universal enveloping algebra of $\mathbf{b}_+$ can be regarded as a subalgebra of $U(sl(2))$. We can make $\mathbf{C}$ into a $\mathbf{b}_+$ module, and hence a $U(\mathbf{b}_+)$ module by

$$h \cdot 1 := \lambda, \quad e \cdot 1 := 0.$$

Then the space

$$U(sl(2)) \otimes_{U(\mathbf{b}^+)} \mathbf{C}$$

with e acting on $\mathbf{C}$ as 0 and h acting via multiplication by λ is a cyclic module with cyclic vector $v_\lambda = 1 \otimes 1$ which satisfies (2.4). It is a "universal" such module in the sense that any other cyclic module with cyclic vector satisfying (2.4) is a homomorphic image of the one we just constructed.

This space $U(sl(2)) \otimes_{U(\mathbf{b}^+)} \mathbf{C}$ is infinite dimensional. It is irreducible unless there is some $\frac{1}{k!}f^k v_\lambda$ with

$$e\left(\frac{1}{k!}f^k v_\lambda\right) = 0$$

where k is an integer ≥ 1. Indeed, any non-zero vector w in the space is a finite linear combination of the basis elements $\frac{1}{k!}f^k v_\lambda$; choose k to be the largest integer so that the coefficient of the corrresponding element does not vanish. Then successive application of the element e (k-times) will yield a multiple of

v_λ, and if this multiple is non-zero, then $U(sl(2))w = U(sl(2))v_\lambda$ is the whole space.

But

$$e\left(\frac{1}{k!}f^k v_\lambda\right) = \left[e, \left(\frac{1}{k!}f^k\right)\right] v_\lambda = (1-k+\lambda)\frac{1}{(k-1)!}f^{k-1}v_\lambda.$$

This vanishes only if λ is an integer and $k = \lambda + 1$, in which case there is a unique finite dimensional quotient of dimension $k+1$. QED

The finite dimensional irreducible representations having zero as a weight are all odd dimensional and have only even weights. We will call them "even". They are called "integer spin" representations by the physicists. The others are "odd" or "half spin" representations.

2.3 The Casimir element.

In $U(sl(2))$ consider the element

$$C := \frac{1}{2}h^2 + ef + fe \tag{2.5}$$

called the **Casimir element** or simply the "Casimir" of $sl(2)$.

Since $ef = fe + [e,f] = fe + h$ in $U(sl(2))$ we also can write

$$C = \frac{1}{2}h^2 + h + 2fe. \tag{2.6}$$

This implies that if v is a "highest weight vector" in a $sl(2)$ module satisfying $ev = 0,\ hv = \lambda v$ then

$$Cv = \frac{1}{2}\lambda(\lambda+2)v. \tag{2.7}$$

Now in $U(sl(2))$ we have

$$\begin{aligned}
[h, C] &= 2([h,f]e + f[h,e]) \\
&= 2(-2fe + 2fe) \\
&= 0
\end{aligned}$$

and

$$\begin{aligned}
[C, e] &= \frac{1}{2}\cdot 2(eh + he) + 2e - 2he \\
&= eh - he + 2e \\
&= -[h,e] + 2e \\
&= 0.
\end{aligned}$$

Similarly

$$[C, f] = 0.$$

In other words, C lies in the **center** of the universal enveloping algebra of $sl(2)$, i.e. it commutes with all elements. If V is a module which possesses a "highest weight vector" v_λ as above, and if V has the property that v_λ is a cyclic vector, meaning that $V = U(L)v_\lambda$ then C takes on the constant value

$$C = \frac{\lambda(\lambda+2)}{2}\mathrm{Id}$$

since C is central and v_λ is cyclic.

2.4 $sl(2)$ is simple.

An ideal I in a Lie algebra $\mathbf{g}$ is a subspace of $\mathbf{g}$ which is invariant under the adjoint representation. In other words, I is an ideal if $[\mathbf{g}, I] \subset I$. If a Lie algebra $\mathbf{g}$ has the property that its only ideals are 0 and $\mathbf{g}$ itself, and if $\mathbf{g}$ is not commutative, we say that $\mathbf{g}$ is **simple**. Let us prove that $sl(2)$ is simple. Since $sl(2)$ is not commutative, we must prove that the only ideals are 0 and $sl(2)$ itself. We do this by introducing some notation which will allow us to generalize the proof in the next chapter. Let

$$\mathbf{g} = sl(2)$$

and set

$$\mathbf{g}_{-1} := \mathbf{C}f, \quad \mathbf{g}_0 := \mathbf{C}h, \quad \mathbf{g}_1 := \mathbf{C}e$$

so that $\mathbf{g}$, as a vector space, is the direct sum of the three one dimensional spaces

$$\mathbf{g} = \mathbf{g}_{-1} \oplus \mathbf{g}_0 \oplus \mathbf{g}_1.$$

Correspondingly, write any $x \in \mathbf{g}$ as

$$x = x_{-1} + x_0 + x_1.$$

If we let

$$d := \frac{1}{2}h$$

then we have

$$\begin{aligned} x &= x_{-1} + x_0 + x_1, \\ [d, x] &= -x_{-1} + 0 + x_1, \text{ and} \\ [d, [d, x]] &= x_{-1} + 0 + x_1. \end{aligned}$$

Since the matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

is invertible, we see that we can solve for the "components" x_{-1}, x_0 and x_1 in terms of x, $[d, x]$, $[d, [d, x]]$. This means that if I is an ideal, then

$$I = I_1 \oplus I_0 \oplus I_1$$

where

$$I_{-1} := I \cap \mathbf{g}_{-1},\ I_0 := I \cap \mathbf{g}_0,\ I_1 := I \cap \mathbf{g}_1.$$

Now if $I_0 \neq 0$ then $d = \frac{1}{2}h \in I$, and hence $e = [d, e]$ and $f = -[d, f]$ also belong to I so $I = sl(2)$. If $I_{-1} \neq 0$ so that $f \in I$, then $h = [e, f] \in I$ so $I = sl(2)$. Similarly, if $I_1 \neq 0$ so that $e \in I$ then $h = [e, f] \in I$ so $I = sl(2)$.

Thus if $I \neq 0$ then $I = sl(2)$ and we have proved that $sl(2)$ is simple.

2.5 Complete reducibility.

We will use the Casimir element C to prove that every finite dimensional representation W of $sl(2)$ is **completely reducible**, which means that if W' is an invariant subspace there exists a complementary invariant subspace W'' so that $W = W' \oplus W''$. Indeed we will prove:

Theorem 2 *1. Every finite dimensional representation of $sl(2)$ is completely reducible.*

2. Each irreducible subspace is a cyclic highest weight module with highest weight n where n is a non-negative integer.

3. When the representation is decomposed into a direct sum of irreducible components, the number of components with even highest weight is the multiplicity of 0 as an an eigenvector of h and

4. the number of components with odd highest weight is the multiplicity of 1 as an eigenvalue of h.

Proof. We know that every irreducible finite dimensional representation is a cyclic module with integer highest weight, that those with even highest weight contain 0 as an eigenvalue of h with multiplicity one and do not contain 1 as an eigenvalue of h, and that those with odd highest weight contain 1 as an eigenvalue of h with multiplicity one, and do not contain 0 as an eigenvalue. So 2), 3) and 4) follow from 1). We must prove 1).

We first prove

Proposition 1 *Let $0 \to V \to W \to k \to 0$ be an exact sequence of $sl(2)$ modules and such that the action of $sl(2)$ on k is trivial (as it must be, since $sl(2)$ has no non-trivial one dimensional modules). Then this sequence splits, i.e. there is a line in W supplementary to V on which $sl(2)$ acts trivially.*

This proposition is, of course, a special case of the theorem we want to prove. But we shall see that it is sufficient to prove the theorem.

Proof of proposition. It is enough to prove the proposition for the case that V is an irreducible module. Indeed, if V_1 is a submodule, then by induction on $\dim V$ we may assume the theorem is known for $0 \to V/V_1 \to W/V_1 \to k \to 0$ so that there is a one dimensional invariant subspace M in W/V_1 supplementary to V/V_1 on which the action is trivial. Let N be the inverse image of M in W. By another application of the proposition, this time to the sequence

$$0 \to V_1 \to N \to M \to 0$$

we find an invariant line, P, in N complementary to V_1. So $N = V_1 \oplus P$. Since $(W/V_1) = (V/V_1) \oplus M$ we must have $P \cap V = \{0\}$. But since dim $W =$ dim $V + 1$, we must have $W = V \oplus P$. In other words P is a one dimensional subspace of W which is complementary to V.

Next we are reduced to proving the proposition for the case that $sl(2)$ acts faithfully on V. Indeed, let $I =$ the kernel of the action on V. Since $sl(2)$ is simple, either $I = sl(2)$ or $I = 0$. Suppose that $I = sl(2)$. For all $x \in sl(2)$ we have, by hypothesis, $xW \subset V$, and for $x \in I = sl(2)$ we have $xV = 0$. Hence

$$[sl(2), sl(2)] = sl(2)$$

acts trivially on all of W and the proposition is obvious. So we are reduced to the case that V is irreducible and the action, ρ, of $sl(2)$ on V is injective. We have our Casimir element C whose image in $\operatorname{End} W$ must map $W \to V$ since every element of $sl(2)$ does. On the other hand, $C = \frac{1}{2}n(n+2)\ \mathrm{Id} \neq 0$ since we are assuming that the action of $sl(2)$ on the irreducible module V is not trivial. In particular, the restriction of C to V is an isomorphism. Hence $\ker C_\rho : W \to V$ is an invariant line supplementary to V. We have proved the proposition.

Proof of theorem from proposition. Let $0 \to E' \to E$ be an exact sequence of $sl(2)$ modules, and we may assume that $E' \neq 0$. We want to find an invariant complement to E' in E. Define W to be the subspace of $\operatorname{Hom}_k(E, E')$ whose restriction to E' is a scalar times the identity, and let $V \subset W$ be the subspace consisting of those linear transformations whose restrictions to E' is zero. Each of these is a submodule of $\operatorname{End}(E)$. We get a sequence

$$0 \to V \to W \to k \to 0$$

and hence a complementary line of invariant elements in W. In particular, we can find an element, T which is invariant, maps $E \to E'$, and whose restriction to E' is non-zero. Then $\ker T$ is an invariant complementary subspace. QED

2.6 The Weyl group.

We have

$$\exp e = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \exp -f = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

so

$$(\exp e)(\exp -f)(\exp e) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Since

$$\exp \operatorname{ad} x = Ad(\exp x)$$

we see that

$$\tau := (\exp \operatorname{ad} e)(\exp \operatorname{ad}(-f))(\exp \operatorname{ad} e)$$

consists of conjugation by the matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Thus

$$\tau(h) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = -h,$$

$$\tau(e) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} = -f$$

and similarly $\tau(f) = -e$. In short

$$\tau : e \mapsto -f,\ f \mapsto -e,\ h \mapsto -h.$$

In particular, τ induces the "reflection" $h \mapsto -h$ on $\mathbf{C}h$ and hence the reflection $\mu \mapsto -\mu$ (which we shall also denote by s) on the (one dimensional) dual space. In any finite dimensional module V of $sl(2)$ the action of the element $\tau = (\exp e)(\exp -f)(\exp e)$ is defined, and

$$(\tau)^{-1} h(\tau) = \operatorname{Ad}\left(\tau^{-1}\right)(h) = s^{-1}h = sh$$

so if

$$hu = \mu u$$

then

$$h(\tau u) = \tau(\tau)^{-1}h(\tau)u = \tau s(h)u = -\mu\tau u = (s\mu)\tau u.$$

So if

$$V_\mu : \{u \in V | hu = \mu u\}$$

then

$$\tau(V_\mu) = V_{s\mu}. \tag{2.8}$$

The two element group consisting of the identity and the element s (acting as a reflection as above) is called the Weyl group of $sl(2)$. Its generalization to an arbitrary simple Lie algebra, together with the generalization of formula (2.8) will play a key role in what follows.

Chapter 3

The classical simple algebras.

In this chapter we introduce the “classical” finite dimensional simple Lie algebras, which come in four families: the algebras $sl(n+1)$ consisting of all traceless $(n+1)\times(n+1)$ matrices, the orthogonal algebras, on even and odd dimensional spaces (the structure for the even and odd cases are different) and the symplectic algebras (whose definition we will give below). We will prove that they are indeed simple by a uniform method - the method that we used in the preceding chapter to prove that $sl(2)$ is simple. So we axiomatize this method.

3.1 Graded simplicity.

We introduce the following conditions on the Lie algebra $\mathbf{g}$:

$$\mathbf{g} = \bigoplus_{i=-1}^{\infty} \mathbf{g}_i \tag{3.1}$$

$$[\mathbf{g}_i, \mathbf{g}_j] \subset \mathbf{g}_{i+j} \tag{3.2}$$

$$[\mathbf{g}_1, \mathbf{g}_{-1}] = \mathbf{g}_0 \tag{3.3}$$

$$[\mathbf{g}_{-1}, x] = 0 \Rightarrow x = 0,\ \forall\, x \in \mathbf{g}_i,\ \forall i \geq 0 \tag{3.4}$$

$$\text{There exists a } d \in \mathbf{g}_0 \text{ satisfying } [d, x] = kx,\ x \in \mathbf{g}_k,\ \forall k, \tag{3.5}$$

and

$$\mathbf{g}_{-1} \text{ is irreducible under the (adjoint) action of } \mathbf{g}_0. \tag{3.6}$$

Condition (3.4) means that if $x \in \mathbf{g}_i, i \geq 0$ is such that $[y, x] = 0$ for all $y \in \mathbf{g}_{-1}$ then $x = 0$.

We wish to show that any non-zero $\mathbf{g}$ satisfying these six conditions is simple. We know that $\mathbf{g}_{-1}, \mathbf{g}_0$ and $\mathbf{g}_1$ are all non-zero, since $0 \neq d \in \mathbf{g}_0$ by (3.5) and

$[\mathbf{g}_{-1}, \mathbf{g}_1] = \mathbf{g}_0$ by (3.3). So $\mathbf{g}$ can not be the one dimensional commutative algebra, and hence what we must show is that any non-zero ideal I of $\mathbf{g}$ must be all of $\mathbf{g}$.

We first show that any ideal I must be a **graded ideal**, i.e. that

$$I = I_{-1} \oplus I_0 \oplus I_1 \oplus \cdots, \quad \text{where } I_j := I \cap \mathbf{g}_j.$$

Indeed, write any $x \in \mathbf{g}$ as $x = x_{-1} + x_0 + x_1 + \cdots + x_k$ and successively bracket by d to obtain

$$\begin{array}{rcl} x & = & x_{-1} + x_0 + x_1 + \cdots + x_k \\ [d, x] & = & -x_{-1} + 0 + x_1 + \cdots + kx_k \\ [d, [d, x]] & = & x_{-1} + 0 + x_1 + \cdots + k^2 x_k \\ \vdots & \vdots & \vdots \\ (\operatorname{ad} d)^k x & = & (-1)^k x_{-1} + 0 + x_1 + \cdots + k^k x_k \\ (\operatorname{ad} d)^{k+1} x & = & (-1)^{k+1} x_{-1} + 0 + x_1 + \cdots + k^{k+1} x_k. \end{array}$$

The matrix

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ -1 & 0 & 1 & \cdots & k \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ (-1)^k & 0 & 1 & \cdots & k^k \\ (-1)^{k+1} & 0 & 1 & \cdots & k^{k+1} \end{pmatrix}$$

is non singular. Indeed, it is a van der Monde matrix, that is a matrix of the form

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ t_1 & t_2 & 1 & \cdots & t_{k+2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ t_1^k & t_2^k & 1 & \cdots & t_{k+2}^k \\ t_1^{k+1} & t_2^{k+1} & 1 & \cdots & t_{k+2}^{k+1} \end{pmatrix}$$

whose determinant is

$$\prod_{i<j} (t_i - t_j)$$

and hence non-zero if all the t_j are distinct. Since $t_1 = -1, t_2 = 0, t_3 = 1$ etc. in our case, our matrix is invertible, and so we can solve for each of the components of x in terms of the $(\operatorname{ad} d)^j x$. In particular, if $x \in I$ then all the $(\operatorname{ad} d)^j x \in I$ since I is an ideal, and hence all the component x_j of x belong to I as claimed.

The subspace $I_{-1} \subset \mathbf{g}_{-1}$ is invariant under the adjoint action of $\mathbf{g}_0$ on $\mathbf{g}_{-1}$, and since we are assuming that this action is irreducible, there are two possibilities: $I_{-1} = 0$ or $I_{-1} = \mathbf{g}_{-1}$. We will show that in the first case $I = 0$ and in the second case that $I = \mathbf{g}$.

Indeed, if $I_{-1} = 0$ we will show inductively that $I_j = 0$ for all $j \geq 0$. Suppose $0 \neq y \in \mathbf{g}_0$. Since every element of $[I_{-1}, y]$ belongs to I and to $\mathbf{g}_{-1}$ we conclude that $[\mathbf{g}_{-1}, y] = 0$ and hence that $y = 0$ by (3.4). Thus $I_0 = 0$. Suppose that we know that $I_{j-1} = 0$. Then the same argument shows that any $y \in I_j$ satisfies $[\mathbf{g}_{-1}, y] = 0$ and hence $y = 0$. So $I_j = 0$ for all j, and since I is the sum of all the I_j we conclude that $I = 0$.

Now suppose that $I_{-1} = \mathbf{g}_{-1}$. Then $\mathbf{g}_0 = [\mathbf{g}_{-1}, \mathbf{g}_1] = [I_{-1}, \mathbf{g}_1] \subset I$. Furthermore, since $d \in \mathbf{g}_0 \subset I$ we conclude that $\mathbf{g}_k \subset I$ for all $k \neq 0$ since every element y of such a $\mathbf{g}_k$ can be written as $y = \frac{1}{k}[d, y] \in I$. Hence $I = \mathbf{g}$. QED

For example, the Lie algebra of all polynomial vector fields, where

$$g_k = \{\sum X^i \frac{\partial}{\partial x_i} \quad X^i \text{ homogenous polynomials of degree } k+1\}$$

is a simple Lie algebra. Here d is the Euler vector field

$$d = x_1 \frac{\partial}{\partial x_1} + \cdots + x_n \frac{\partial}{\partial x_n}.$$

This algebra is infinite dimensional. We are primarily interested in the finite dimensional Lie algebras.

3.2 $sl(n+1)$

Write the most general matrix in $sl(n+1)$ as

$$\begin{pmatrix} -\operatorname{tr} A & w^* \\ v & A \end{pmatrix}$$

where A is an arbitrary $n \times n$ matrix, v is a column vector and $w^* = (w_1, \ldots, w_n)$ is a row vector. Let $\mathbf{g}_{-1}$ consist of matrices with just the top row, i.e. with $v = A = 0$. Let $\mathbf{g}_1$ consist of matrices with just the left column, i.e. with $A = w^* = 0$. Let $\mathbf{g}_0$ consist of matrices with just the central block, i.e. with $v = w^* = 0$. Let

$$d = \frac{1}{n+1} \begin{pmatrix} -n & 0 \\ 0 & I \end{pmatrix}$$

where I is the $n \times n$ identity matrix. Thus g_0 acts on g_{-1} as the algebra of all endomorphisms, and so g_{-1} is irreducible. We have

$$[\begin{pmatrix} 0 & 0 \\ v & 0 \end{pmatrix}, \begin{pmatrix} 0 & w^* \\ 0 & 0 \end{pmatrix}] = \begin{pmatrix} -\langle w^*, v\rangle & 0 \\ 0 & v \otimes w^* \end{pmatrix},$$

where $\langle w^*, v\rangle$ denotes the value of the linear function w^* on the vector v, and this is precisely the trace of the rank one linear transformation $v \otimes w^*$. Thus all our axioms are satisfied. The algebra $sl(n+1)$ is simple.

3.3 The orthogonal algebras.

The algebra $o(2)$ is one dimensional and (hence) commutative. In our (real) Euclidean three dimensional space, the algebra $o(3)$ has a basis X, Y, Z (infinitesimal rotations about each of the axes) with bracket relations

$$[X,Y] = Z,\ [Y,Z] = X,\ [Z,X] = Y,$$

(the usual formulae for "vector product" in three dimensions". But we are over the complex numbers, so can consider the basis $X + iY, -X + iY, iZ$ and find that

$$[iZ, X+iY] = X+iY,\ \ [iZ, -X+iY] = -(-X+iY),\ \ [X+iY, -X+iY] = 2iZ.$$

These are the bracket relations for $sl(2)$ with $e = X + iY,\ f = -X + iY,\ h = iZ$. In other words, the complexification of our three dimensional world is the irreducible three dimensional representation of $sl(2)$ so $o(3) = sl(2)$ which is simple.

To study the higher dimensional orthogonal algebras it is useful to make two remarks:

If V is a vector space with a non-degenerate symmetric bilinear form $(\ ,\ )$, we get an isomorphism of V with its dual space V^* sending every $u \in V$ to the linear function ℓ_u where $\ell_u(v) = (v, u)$. This gives an identification of

$$\mathrm{End}(V) = V \otimes V^* \ \text{ with } \ V \otimes V.$$

Under this identification, the elements of $o(V)$ become identified with the anti-symmetric two tensors, that is with elements of $\wedge^2(V)$. (In terms of an orthonormal basis, a matrix A belongs to $o(V)$ if and only if it is anti-symmetric.)

Explicitly, an element $u \wedge v$ becomes identified with the linear transformation $A_{u\wedge v}$ where

$$A_{u\wedge v}x = (x, v)u - (u, x)v.$$

This has the following consequence. Suppose that $z \in V$ with $(z, z) \neq 0$, and let w be any element of V. Then

$$A_{w\wedge z}z = (z, z)w - (z, w)z$$

and so $U(o(V))z = V$. On the other hand, suppose that $u \in V$ with $(u, u) = 0$. We can find $v \in V$ with $(v, v) = 0$ and $(v, u) = 1$. Now suppose in addition that $\dim V \geq 3$. We can then find a $z \in V$ orthogonal to the plane spanned by u and v and with $(z, z) = 1$. Then

$$A_{z\wedge v}u = z,$$

so $z \in U(o(V))e$ and hence $U(o(V))u = V$. We have proved:

1 *If* $\dim V \geq 3$, *then every non-zero vector in* V *is cyclic, i.e the representation of* $o(V)$ *on* V *is irreducible.*

(In two dimensions this is false - the line spanned by a vector e with $(e,e)=0$ is a one dimensional invariant subspace.)

We now show that

2 *$o(V)$ is simple for dim $V \geq 5$.*

For this, begin by writing down the bracket relations for elements of $o(V)$ in terms of their parametrization by elements of $\wedge^2 V$. Direct computation shows that

$$[A_{u\wedge v}, A_{x\wedge y}] = (v,x)A_{u\wedge y} - (u,x)A_{v\wedge y} - (v,y)A_{u\wedge x} + (u,y)A_{v\wedge x}. \qquad (3.7)$$

Now let $n = \dim V - 2$ and choose a basis

$$u, v, x_1, \ldots, x_n$$

of V where

$$(u,u) = (u,x_i) = (v,v) = (v,x_i) = 0 \ \forall i, \ (u,v) = 1, \ (x_i,x_j) = \delta_{ij}.$$

Let $\mathbf{g} := o(V)$ and write W for the subspace spanned by the x_i. Set

$$d := A_{u\wedge v}$$

and

$$\mathbf{g}_{-1} := \{A_{v\wedge x}, \ x \in W\}, \quad \mathbf{g}_0 := o(W) \oplus \mathbf{C}d, \ \mathbf{g}_1 := \{A_{u\wedge x}, \ x \in W\}.$$

It then follows from (3.7) that d satisfies (3.5). The spaces $\mathbf{g}_{-1}$ and $\mathbf{g}_1$ look like copies of W with the $o(W)$ part of $\mathbf{g}_0$ acting as $o(W)$, hence irreducibly since $\dim W \geq 3$. All our remaining axioms are easily verified. Hence $o(V)$ is simple for $\dim V \geq 5$.

We have seen that $o(3) = sl(2)$ is simple.

However $o(4)$ is not simple, being isomorphic to $sl(2) \oplus sl(2)$: Indeed, if Z_1 and Z_2 are vector spaces equipped with non-degenerate anti-symmetric bilinear forms $\langle\ ,\ \rangle_1$ and $\langle\ ,\ \rangle_2$ then $Z_1 \otimes Z_2$ has a non-degenerate symmetric bilinear form $(\ ,\ )$ determined by

$$(u_1 \otimes u_2, v_1 \otimes v_2) = \langle u_1, v_1\rangle_1 \langle u_2, v_2\rangle_2.$$

The algebra $sl(2)$ acting on its basic two dimensional representation infinitesimally preserves the antisymmetric form given by

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = x_1 y_2 - x_2 y_1.$$

Hence, if we take $Z = Z_1 = Z_2$ to be this two dimensional space, we see that $sl(2) \oplus sl(2)$ acts as infinitesimal orthogonal transformations on $Z \otimes Z$ which is four dimensional. But $o(4)$ is six dimensional so the embedding of $sl(2) \oplus sl(2)$ in $o(4)$ is in fact an isomorphism since $3 + 3 = 6$.

3.4 The symplectic algebras.

We consider an even dimensional space with coordinates $q_1, q_2, \ldots, p_1, p_2, \ldots$. The polynomials have a Poisson bracket

$$\{f, g\} := \sum \left(\frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} \right). \tag{3.8}$$

This is clearly anti-symmetric, and direct computation will show that the Jacobi identity is satisfied. Here is a more interesting proof of Jacobi's identity: Notice that if f is a constant, then $\{f, g\} = 0$ for all g. So in doing bracket computations we can ignore constants. On the other hand, if we take g to be successively $q_1, \ldots, q_n, p_1, \ldots, p_n$ in (3.8) we see that the partial derivatives of f are completely determined by how it brackets with all g, in fact with all linear g. If we fix f, the map

$$h \mapsto \{f, h\}$$

is a **derivation**, i.e. it is linear and satisfies

$$\{f, h_1 h_2\} = \{f, h_1\} h_2 + h_1 \{f, h_2\}.$$

This follows immediately from from the definition (3.8). Now Jacobi's identity amounts to the assertion that

$$\{\{f, g\}, h\} = \{f, \{g, h\}\} - \{g, \{f, h\}\},$$

i.e. that the derivation

$$h \mapsto \{\{f, g\}, h\}$$

is the commutator of the of the derivations

$$h \mapsto \{f, h\} \quad \text{and} \qquad h \mapsto \{g, h\}.$$

It is enough to check this on linear polynomials h, and hence on the polynomials q_j and p_k. If we take $h = q_j$ then

$$\{f, q_j\} = \frac{\partial f}{\partial p_j}, \quad \{g, q_j\} = \frac{\partial g}{\partial p_j}$$

so

$$\begin{aligned}
\{f, \{g, q_j\}\} &= \sum \left(\frac{\partial f}{\partial p_i} \frac{\partial^2 g}{\partial q_i \partial p_j} - \frac{\partial f}{\partial q_i} \frac{\partial^2 g}{\partial p_i \partial p_j} \right) \\
\{f, \{f, q_j\}\} &= \sum \left(\frac{\partial g}{\partial p_i} \frac{\partial^2 f}{\partial q_i \partial p_j} - \frac{\partial g}{\partial q_i} \frac{\partial^2 f}{\partial p_i \partial p_j} \right) \text{ so} \\
\{f, \{g, q_j\}\} - \{g, \{f, q_j\}\} &= \frac{\partial}{\partial p_j} \{f, g\} \\
&= \{\{f, g\}, q_j\}
\end{aligned}$$

as desired, with a similar computation for p_k.

The symplectic algebra $sp(2n)$ is defined to be the subalgebra consisting of all homogeneous quadratic polynomials. We divide these polynomials into three groups as follows: Let $\mathbf{g}_1$ consist of homogeneous polynomials in the q's alone, so $\mathbf{g}_1$ is spanned by the q_iq_j. Let $\mathbf{g}_{-1}$ be the quadratic polynomials in the p's alone, and let $\mathbf{g}_0$ be the mixed terms, so spanned by the q_ip_j. It is easy to see that $\mathbf{g}_0 \sim gl(n)$ and that $[\mathbf{g}_{-1}, \mathbf{g}_1] = \mathbf{g}_0$. To check that $\mathbf{g}_{-1}$ is irreducible under $\mathbf{g}_0$, observe that $[p_1q_j, p_kp_\ell] = 0$ if $j \neq k$ or ℓ, and $[p_1q_j, p_jp_\ell]$ is a multiple of p_1p_ℓ. So we can by one or two brackets carry any non-zero element of $\mathbf{g}_{-1}$ into a non-zero multiply of p_1^2, and then get any monomial from p_1^2 by bracketing with p_iq_1 appropriately. The element d is given by $\frac{1}{2}(p_1q_1 + \cdots + p_nq_n)$.

We have shown that the symplectic algebra is simple, but we haven't really explained what it is. Consider the space of V of homogenous linear polynomials, i.e all polynomials of the form

$$\ell = a_1q_1 + \cdots + a_nq_n + b_1p_q + \cdots + b_np_n.$$

Define an anti-symmetric bilinear form ω on V by setting

$$\omega(\ell, \ell') := \{\ell, \ell'\}.$$

From the formula (3.8) it follows that the Poisson bracket of two linear functions is a constant, so ω does indeed define an antisymmetric bilinear form on V, and we know that this bilinear form is non-degenerate. Furthermore, if f is a homogenous quadratic polynomial, and ℓ is linear, then $\{f, \ell\}$ is again linear, and if we denote the map

$$\ell \mapsto \{f, \ell\}$$

by $A = A_f$, then Jacobi's identity translates into

$$\omega(A\ell, \ell') + \omega(\ell A\ell') = 0 \tag{3.9}$$

since $\{\ell, \ell'\}$ is a constant. Condition (3.9) can be interpreted as saying that A belongs to the Lie algebra of the group of all linear transformations R on V which preserve ω, i.e. which satisfy

$$\omega(R\ell, R\ell') = \omega(\ell, \ell').$$

This group is known as the symplectic group. The form ω induces an isomorphism of V with V^* and hence of $\mathrm{Hom}(V, V) = V \otimes V^*$ with $V \otimes V$, and this time the image of the set of A satisfying (3.9) consists of all symmetric tensors of degree two, i.e. of $S^2(V)$. (Just as in the orthogonal case we got the anti-symmetric tensors). But the space $S^2(V)$ is the same as the space of homogenous polynomials of degree two. In other words, the symplectic algebra as defined above is the same as the Lie algebra of the symplectic group.

It is an easy theorem in linear algebra, that if V is a vector space which carries a non-degenerate anti-symmetric bilinear form, then V must be even dimensional, and if dim $V = 2n$ then it is isomorphic to the space constructed above. We will not pause to prove this theorem.

3.5 The root structures.

We are going to choose a basis for each of the classical simple algebras which generalizes the basis e, f, h that we chose for $sl(2)$. Indeed, for each classical simple algebra $\mathbf{g}$ we will first choose a maximal commutative subalgebra $\mathbf{h}$ all of whose elements are semi-simple = diagonizable in the adjoint representation. Since the adjoint action of all the elements of $\mathbf{h}$ commute, this means that they can be simultaneously diagonalized. Thus we can decompose $\mathbf{g}$ into a direct sum of simultaneous eigenspaces

$$\mathbf{g} = \mathbf{h} \oplus \bigoplus_{\alpha} \mathbf{g}_\alpha \tag{3.10}$$

where $0 \neq \alpha \in \mathbf{h}^*$ and

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

The linear functions α are called **roots** (originally because the $\alpha(h)$ are roots of the characteristic polynomial of $\mathrm{ad}(h)$). The simultaneous eigenspace $\mathbf{g}_\alpha$ is called the root space corresponding to α. The collection of all roots will usually be denoted by Φ.

Let us see how this works for each of the classical simple algebras.

3.5.1 $A_n = sl(n+1)$.

We choose $\mathbf{h}$ to consist of the diagonal matrices in the algebra $sl(n+1)$ of all $(n+1) \times (n+1)$ matrices with trace zero. As a basis of $\mathbf{h}$ we take

$$\begin{aligned}
h_1 &:= \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \\
h_2 &:= \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \\
\vdots &:= \vdots \\
h_n &:= \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & -1 \end{pmatrix}.
\end{aligned}$$

Let L_i denote the linear function which assigns to each diagonal matrix its i-th (diagonal) entry, Let E_{ij} denote the matrix with one in the i,j position and zero's elsewhere. Then

$$[h, E_{ij}] = (L_i(h) - L_j(h))E_{ij} \quad \forall \ \ h \in \mathbf{h}$$

so the linear functions of the form

$$L_i - L_j, \quad i \neq j$$

are the roots.

We may subdivide the set of roots into two classes: the **positive roots**

$$\Phi^+ := \{L_i - L_j; \ i < j\}$$

and the **negative roots**

$$\Phi^- := -\Phi^+ = \{L_j - L_i, \ i < j\}.$$

Every root is either positive or negative. If we define

$$\alpha_i := L_i - L_{i+1}$$

then every positive root can be written as a sum of the α_i:

$$L_i - L_j = \alpha_i + \cdots + \alpha_{j-1}.$$

We have

$$\alpha_i(h_i) = 2,$$

and for $i \neq j$

$$\alpha_i(h_{i\pm 1}) = -1, \quad \alpha_i(h_j) = 0, \ j \neq i \pm 1. \tag{3.11}$$

The elements

$$E_{i,i+1}, \ h_i, \ E_{i+1,i}$$

form a subalgebra of $sl(n+1)$ isomorphic to $sl(2)$. We may call it $sl(2)_i$.

3.5.2 $C_n = sp(2n), n \geq 2$.

Let $\mathbf{h}$ consist of all linear combinations of $p_1q_1, \ldots, p_nq_n$ and let L_i be defined by

$$L_i\left(a_1p_1q_1 + \cdots + a_np_nq_n\right) = a_i$$

so $L_1, \ldots, L_n$ is the basis of $\mathbf{h}^*$ dual to the basis $p_1q_1, \ldots, p_nq_n$ of $\mathbf{h}$.

If $h = a_1p_1q_1 + \cdots + a_np_nq_n$ then

$$\begin{aligned} [h, q^iq^j] &= (a_i + a_j)q^iq^j \\ [h, q^ip^j] &= (a_i - a_j)q^ip^j \\ [h, p^ip^j] &= -(a_i + a_j)p^ip^j \end{aligned}$$

so the roots are

$$\pm(L_i + L_j) \text{ all } i, j \quad \text{and} \quad L_i - L_j \ i \neq j.$$

We can divide the roots Φ into positive and negative roots by setting

$$\Phi^+ = \{L_i + L_j\}_{\text{all } ij} \cup \{L_i - L_j\}_{i<j}.$$

If we set

$$\alpha_1 := L_1 - L_2, \dots, \alpha_{n-1} := L_{n-1} - L_n, \alpha_n := 2L_n$$

then every positive root is a sum of the α_i. Indeed, $L_{n-1} + L_n = \alpha_{n-1} + \alpha_n$ and $2L_{n-1} = 2\alpha_{n-1} + \alpha_n$ and so on. In particular $2\alpha_{n-1} + \alpha_n$ is a root.

If we set

$$h_1 := p_1 q_1 - p_2 q_2, \dots, h_{n-1} := p_{n-1} q_{n-1} - p_n q_n, \ h_n := p_n q_n$$

then

$$\alpha_i(h_i) = 2$$

while for $i \neq j$

$$\begin{aligned} \alpha_i(h_{i\pm1}) &= -1, \quad i = 1, \dots, n-1 \\ \alpha_i(h_j) &= 0, \ j \neq i \pm 1, i = 1, \dots, n \\ \alpha_n(h_{n-1}) &= -2. \end{aligned} \tag{3.12}$$

In particular, the elements $h_i, q_i p_{i+1}, q_{i+1} p_i$ for $i = 1, \dots, n-1$ form a subalgebra isomorphic to $sl(2)$ as do the elements $h_n, \frac{1}{2}q_n^2, -\frac{1}{2}p_n^2$. We call these subalgebras $sl(2)_i, \ i = 1, \dots, n$.

3.5.3 $D_n = o(2n), \ n \geq 3$.

We choose a basis $u_1, \dots, u_n, v_1, \dots, v_n$ of our orthogonal vector space V such that

$$(u_i, u_j) = (v_i, v_j) = 0, \forall \ i, j, \quad (u_i, v_j) = \delta_{ij}.$$

We let $\mathbf{h}$ be the subalgebra of $o(V)$ spanned by the $A_{u_i v_i}$, $i = 1, \dots, n$. Here we have written A_{xy} instead of $A_{x \wedge y}$ in order to save space. We take

$$A_{u_1 v_1}, \dots, A_{u_n v_n}$$

as a basis of $\mathbf{h}$ and let $L_1, \dots, L_n$ be the dual basis. Then

$$\pm L_k \pm L_\ell \ k \neq \ell$$

are the roots since from (3.7) we have

$$\begin{aligned} [A_{u_i v_i}, A_{u_k u_\ell}] &= (\delta_{ik} + \delta_{i\ell}) A_{u_k u_\ell} \\ [A_{u_i v_i}, A_{u_k v_\ell}] &= (\delta_{ik} - \delta_{i\ell}) A_{u_k v_\ell} \\ [A_{u_i v_i}, A_{v_k v_\ell}] &= -(\delta_{ik} + \delta_{i\ell}) A_{v_k v_\ell}. \end{aligned}$$

We can choose as positive roots the

$$L_k + L_\ell, L_k - L_\ell, \quad k < \ell$$

and set

$$\alpha_i := L_i - L_{i+1},\ i = 1, \dots, n-1, \quad \alpha_n := L_{n-1} + L_n.$$

Every positive root is a sum of these simple roots. If we set

$$h_i := A_{u_i v_i} - A_{u_{i+1} v_{i+1}}, i = 1, \dots n-1,$$

and

$$h_n = A_{u_{n-1} v_{n-1}} + A_{u_n v_n}$$

then

$$\alpha_i(h_i) = 2$$

and for $i \neq j$

$$\begin{aligned} \alpha_i(h_j) &= 0 \ \ j \neq i \pm 1,\ i = 1, \dots n-2 \\ \alpha_i(h_{i\pm 1}) &= -1 \ \ i = 1, \dots, n-2 \\ \alpha_{n-1}(h_{n-2}) &= -1 \\ \alpha_n(h_{n-2}) &= -1 \\ \alpha_n(h_{n-1}) &= 0. \end{aligned} \tag{3.13}$$

For $i = 1, \dots, n-1$ the elements $h_i, A_{u_i v_{i+1}}, A_{u_{i+1} v_i}$ form a subalgebra isomorphic to $sl(2)$ as do $h_n, A_{u_{n-1} u_n}, A_{v_{n-1} v_n}$.

3.5.4 $B_n = o(2n+1)$ $n \geq 2$.

We choose a basis $u_1, \dots, u_n, v_1, \dots, v_n, x$ of our orthogonal vector space V such that

$$(u_i, u_j) = (v_i, v_j) = 0, \forall\ i, j, \quad (u_i, v_j) = \delta_{ij},$$

and

$$(x, u_i) = (x, v_i) = 0\ \forall\ i, (x, x) = 1.$$

As in the even dimensional case we let $\mathbf{h}$ be the subalgebra of $o(V)$ spanned by the $A_{u_i v_i}$, $i = 1, \dots, n$ and take

$$A_{u_1 v_1}, \dots, A_{u_n v_n}$$

as a basis of $\mathbf{h}$ and let $L_1, \dots, L_n$ be the dual basis. Then

$$\pm L_i \pm L_j \ \ i \neq j, \pm L_i$$

are roots. We take

$$L_i \pm L_j,\ 1 \leq i < j \leq n,\ \text{together with } L_i,\ i = 1, \dots, n$$

to be the positive roots, and

$$\alpha_i := L_i - L_{i+1},\ i = 1, \ldots, n-1,\ \alpha_n := L_n$$

to be the simple roots. We let

$$h_i := A_{u_i v_i} - A_{u_{i+1} v_{i+1}}, i = 1, \ldots n-1,$$

as in the even case, but set

$$h_n := 2A_{u_n v_n}.$$

Then every positive root can be written as a sum of the simple roots,

$$\alpha_i(h_i) = 2,\ i = 1, \ldots n,$$

and for $i \neq j$

$$\begin{array}{rcl} \alpha_i(h_j) & = & 0 \quad j \neq i \pm 1,\ i = 1, \ldots n \\ \alpha_i(h_{i\pm1}) & = & -1 \quad i = 1, \ldots, n-2, n \\ \alpha_{n-1}(h_n) & = & -2 \end{array} \tag{3.14}$$

Notice that in this case $\alpha_{n-1} + 2\alpha_n = L_{n-1} + L_n$ is a root. Finally we can construct subalgebras isomorphic to $sl(2)$, with the first $n-1$ as in the even orthogonal case and the last $sl(2)$ spanned by $h_n, A_{u_n x}, -A_{v_n x}$.

3.5.5 Diagrammatic presentation.

The information of the last four subsections can be summarized in each of the following four diagrams:

The way to read this diagram is as follows: each node in the diagram stands for a simple root, reading from left to right, starting with α_1 at the left. (In the diagram D_ℓ the two rightmost nodes are $\alpha_{\ell-1}$ and α_ℓ, say the top $\alpha_{\ell-1}$ and the bottom α_ℓ.) Two nodes α_i and α_j are connected by (one or more) edges if and only if $\alpha_i(h_j) \neq 0$.

In all cases, the difference, $\alpha_i - \alpha_j$ is never a root, and, for $i \neq j,\ \alpha_i(h_j) \leq 0$ and is an integer. If, for $i \neq j$, $\alpha_i(h_j) < 0$ then $\alpha_i + \alpha_j$ is a root.

In two of the cases (B_ℓ and C_ℓ) it happens that $\alpha_i(h_j) = -2$. Then $\alpha_i + \alpha_j$ and $\alpha_i + 2\alpha_j$ are roots, and we draw a double bond with an arrow pointing towards α_j. In this case 2 is is the maximum integer such that $\alpha_i + k\alpha_j$ is a root. In all other cases, this maximum integer k is one if the nodes are connected (and zero it they are not).

3.6 Low dimensional coincidences.

We have already seen that $o(4) \sim sl(2) \oplus sl(2)$. We also have

$$o(6) \sim sl(4).$$

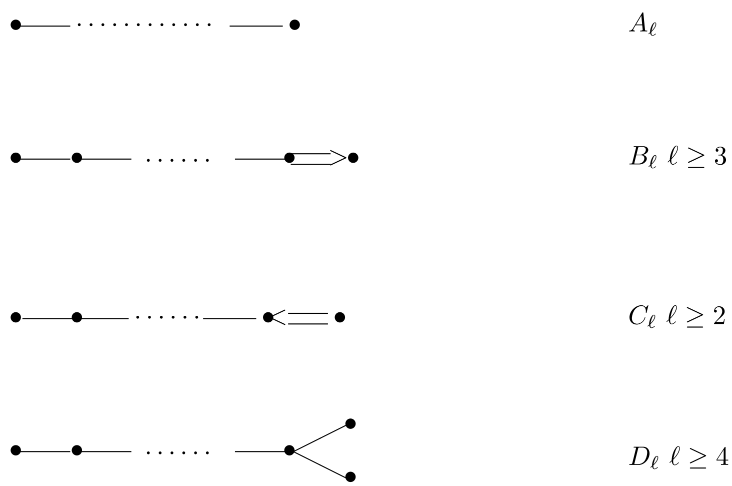


Figure 3.1: Dynkin diagrams of the classical simple algebras.

Both algebras are fifteen dimensional and both are simple. So to realize this isomorphism we need only find an orthogonal representation of $sl(4)$ on a six dimensional space. If we let $V = \mathbf{C}^4$ with the standard representation of $sl(4)$, we get a representation of $sl(4)$ on $\wedge^2(V)$ which is six dimensional. So we must describe a non-degenerate bilinear form on $\wedge^2 V$ which is invariant under the action of $sl(4)$. We have a map, wedge product, of

$$\wedge^2 V \times \wedge^2 V \to \wedge^4 V.$$

Furthermore this map is symmetric, and invariant under the action of $gl(4)$. However $sl(4)$ preserves a basis (a non-zero element) of $\wedge^4 V$ and so we may identify $\wedge^4 V$ with $\mathbf{C}$. It is easy to check that the bilinear form so obtained is non-degenerate

We also have the identification

$$sp(4) \sim o(5)$$

both algebras being ten dimensional. To see this let $V = \mathbf{C}^4$ with an antisymmetric form ω preserved by $Sp(4)$. Then $\omega \otimes \omega$ induces a symmetric bilinear form on $V \otimes V$ as we have seen. Sitting inside $V \otimes V$ as an invariant subspace is $\wedge^2 V$ as we have seen, which is six dimensional. But $\wedge^2 V$ is not irreducible as a representation of $sp(4)$. Indeed, $\omega \in \wedge^2 V^*$ is invariant, and hence its kernel is a five dimensional subspace of $\wedge^2 V$ which is invariant under $sp(4)$. We thus get a non-zero homomorphism $sp(4) \to o(5)$ which must be an isomorphism since $sp(4)$ is simple.

These coincidences can be seen in the diagrams. If we were to allow $\ell = 2$ in the diagram for B_ℓ it would be indistinguishable from C_2. If we were to allow $\ell = 3$ in the diagram for D_ℓ it would be indistinguishable from A_3.

3.7 Extended diagrams.

It follows from Jacobi's identity that in the decomposition (3.10), we have

$$[\mathbf{g}_\alpha, \mathbf{g}_{\alpha'}] \subset \mathbf{g}_{\alpha+\alpha'} \tag{3.15}$$

with the understanding that the right hand side is zero if $\alpha + \alpha'$ is not a root. In each of the cases examined above, every positive root is a linear combination of the simple roots with non-negative integer coefficients. Since the algebra is finite, there must be a **maximal** positive root β in the sense that $\beta + \alpha_i$ is not a root for any simple root. For example, in the case of $A_n = sl(n+1)$, the root $\beta := L_1 - L_{n+1}$ is maximal. The corresponding $\mathbf{g}_\beta$ consists of all $(n+1)\times(n+1)$ matrices with zeros everywhere except in the upper right hand corner. We can also consider the **minimal root** which is the negative of the maximal root, so

$$\alpha_0 := -\beta = L_{n+1} - L_1$$

in the case of A_n. Continuing to study this case, let

$$h_0 := h_{n+1} - h_1.$$

Then we have

$$\alpha_i(h_i) = 2, \;\; i = 0, \dots n$$

and

$$\alpha_0(h_1) = \alpha_0(h_n) = -1, \quad \alpha_0(h_i) = 0, \quad i \neq 0, 1, n.$$

This means that if we write out the $(n+1) \times (n+1)$ matrix whose entries are $\alpha_i(h_j), \; i, j = 0, \dots n$ we obtain a matrix of the form

$$2I - M$$

where $M_{ij} = 1$ if and only if $j = \pm 1$ with the understanding that $n + 1 = 0$ and $-1 = n$, i.e we do the subscript arithmetic mod n. In other words, M is the adjacency matrix of the cyclic graph with $n + 1$ vertices labeled $0, \dots n$. Also, we have

$$h_0 + h_1 + \cdots + h_n = 0.$$

If we apply α_i to this equation for $i = 0, \dots n$ we obtain

$$(2I - M)\mathbf{1} = 0,$$

where $\mathbf{1}$ is the column vector all of whose entries are 1. We can write this equation as

$$M\mathbf{1} = \mathbf{2}\mathbf{1}.$$

In other words, **1** is an eigenvector of M with eigenvalue 2.

In the chapters that follow we shall see that any finite dimensional simple Lie algebra has roots, simple roots, maximal roots etc. giving rise to a matrix M with integer entries which is irreducible (in the sense of non-negative matrices - definition later on) and which has an eigenvector with positive (integer) entries with eigenvalue 2. This will allow us to classify the simple (finite dimensional) Lie algebras.

Chapter 4

Engel-Lie-Cartan-Weyl

We return to the general theory of Lie algebras. Many of the results in this chapter are valid over arbitrary fields, indeed if we use the axioms to define a Lie algebra over a ring many of the results are valid in this generality. But some of the results depend heavily on the ring being an algebraically closed field of characteristic zero. As a compromise, throughout this chapter we deal with fields, and will assume that all vector spaces and all Lie algebras which appear are finite dimensional. We will indicate the necessary additional assumptions on the ground field as they occur. The treatment here follows Serre pretty closely.

4.1 Engel's theorem

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

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

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

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

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

Theorem 3 $\mathbf{g}$ *is nilpotent if and only if* $\operatorname{ad} x$ *is a nilpotent operator for each* $x \in \mathbf{g}$.

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

Theorem 4 Engel *Let* $\rho : \mathbf{g} \to \operatorname{End}(V)$ *be a representation such that* $\rho(x)$ *is nilpotent for each* $x \in \mathbf{g}$. *Then there exists a basis in terms of which* $\rho(\mathbf{g}) \subset \mathbf{n}^+(gl(d))$, *i.e. becomes strictly upper triangular. Here* $d =$dim V.

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

Theorem 5 *Under the hypotheses of Engel's theorem, if $V \neq 0$, there exists a non-zero vector $v \in V$ such that $\rho(x)v = 0\ \forall x \in \mathbf{g}$.*

Proof of Theorem 5 in seven easy steps.

- Replace $\mathbf{g}$ by its image, i.e. assume that $\mathbf{g} \subset \text{End } V$.
- Then $(\text{ad}\, x)y = L_x y - R_x y$ where L_x is the linear map of $\text{End}\, V$ into itself given by left multiplication by x, and R_x is given by right multiplication by x. Both L_x and R_x are nilpotent as operators since x is nilpotent. Also they commute. Hence by the binomial formula $(\text{ad}\, x)^n = (L_x - R_x)^n$ vanishes for sufficiently large n.
- We may assume (by induction) that for any Lie algebra, $\mathbf{m}$, of smaller dimension than that of $\mathbf{g}$ (and any representation) there exists a $v \in V$ such that $xv = 0\ \forall x \in \mathbf{m}$.
- Let $\mathbf{k} \subset \mathbf{g}$ be a subalgebra, $\mathbf{k} \neq \mathbf{g}$, and let

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

  be its normalizer. The claim is that

  3 $N(\mathbf{k})$ is strictly larger than $\mathbf{k}$.

  To see this, observe that each $x \in \mathbf{k}$ acts on $\mathbf{k}$ and on $\mathbf{g}/\mathbf{k}$ by nilpotent maps, and hence there is an $0 \neq \hat{y} \in \mathbf{g}/\mathbf{k}$ killed by all $x \in \mathbf{k}$. But then $y \notin \mathbf{k}$, and $[y, x] = -[x, y] \in \mathbf{k}$ for all $x \in \mathbf{k}$. So $y \in N(\mathbf{k}),\ y \notin \mathbf{k}$.
- If $\mathbf{g} \neq 0$, there is an ideal $\mathbf{i} \subset \mathbf{g}$ such that $\dim \mathbf{g}/\mathbf{i} = 1$. Indeed, let $\mathbf{i}$ be a maximal proper subalgebra of $\mathbf{g}$. Its normalizer is strictly larger, hence all of $\mathbf{g}$, so $\mathbf{i}$ is an ideal. The inverse image in $\mathbf{g}$ of a line in $\mathbf{g}/\mathbf{i}$ is a subalgebra, and is strictly larger than $\mathbf{i}$. Hence it must be all of $\mathbf{g}$.
- Choose such an ideal, $\mathbf{i}$. The subspace

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

  is invariant under $\mathbf{g}$. Indeed, if $y \in \mathbf{g},\ w \in W$ then $xyw = yxw + [x, y]w = 0$.
- $W \neq 0$ by induction. Take $y \in \mathbf{g},\ y \notin \mathbf{i}$. It preserves W and is nilpotent. Hence there is a non-zero $v \in W$ with $yv = 0$. Since y and $\mathbf{i}$ span $\mathbf{g}$, we have $xv = 0\ \ \forall x \in \mathbf{g}$. QED

No assumptions about the ground field went into this.

4.2 Solvable Lie algebras.

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

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

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

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

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

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

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

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

Theorem 6 [Lie.] *Let $\mathbf{g}$ be a solvable Lie algebra over an algebraically closed field k of characteristic zero, and (ρ, V) a finite dimensional representation of $\mathbf{g}$. Then we can find a basis of V so that $\rho(\mathbf{g})$ consists of upper triangular matrices.*

By induction on dim V this reduces to

Theorem 7 [Lie.] *Under the same hypotheses, there exists a (non-zero) common eigenvector v for all the $\rho(y)$, i.e. there is a vector $v \in V$ and a function $\chi : \mathbf{g} \to k$ such that*

$$\rho(y)v = \chi(y)v \quad \forall\ y \in \mathbf{g}. \tag{4.1}$$

Lemma 2 *Suppose that $\mathbf{i}$ is an ideal of $\mathbf{g}$ and (4.1) holds for all $y \in \mathbf{i}$. Then*

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

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

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

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

Thus V_n is invariant under $\mathbf{i}$ and for each $h \in \mathbf{i}$, $\mathrm{tr}_{|V_n} h = n\chi(h)$. In particular both x and h leave V_n invariant and $\mathrm{tr}_{|V_n}[x,h] = 0$ since the trace of any commutator is zero. This proves the lemma.
Proof of theorem by induction on dim $\mathbf{g}$, which we may assume to be positive. Let $\mathbf{m}$ be any subspace of $\mathbf{g}$ with $\mathbf{g} \supset \mathbf{m} \supset [\mathbf{g},\mathbf{g}]$. Then $[\mathbf{g},\mathbf{m}] \subset [\mathbf{g},\mathbf{g}] \subset \mathbf{m}$ so $\mathbf{m}$ is an ideal in $\mathbf{g}$. In particular, we may choose $\mathbf{m}$ to be a subspace of codimension 1 containing $[\mathbf{g},\mathbf{g}]$. By induction we can find a $v \in V$ and a $\chi : \mathbf{m} \to k$ such that (4.1) holds for all elements of $\mathbf{m}$. Let

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

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

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

since $\chi([x,h]) = 0$ by the lemma. Thus W is stable under all of $\mathbf{g}$. Pick $x \in \mathbf{g}$, $x \notin \mathbf{m}$, and let $v \in W$ be an eigenvector of x with eigenvalue λ, say. Then v is a simultaneous eigenvector for all of $\mathbf{g}$ with χ extended as

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

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

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

4.3 Linear algebra

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

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

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

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

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

$$u = s + n$$

where

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

is nilpotent. Also

$$ns = sn.$$

We claim that these two elements are uniquely determined by

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

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

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

Define

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

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

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

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

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

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

$$\ell(y)x.$$

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

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

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

Proposition 2 *If $u = s + n$ is the decomposition of u then $u_{pq} = s_{pq} + n_{pq}$ is the decomposition of u_{pq}.*

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

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

Proposition 3 *Suppose that ϕ is additive. Then*

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

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

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

where the middle equation is just the additivity. QED

As an immediate consequence we obtain

Proposition 4 *Notation as above. If $A \subset B \subset V_{p,q}$ with $u_{pq}B \subset A$ then for any additive map, $\phi(s)_{pq}B \subset A$*

Proposition 5 (over $\mathbf{C}$) *Let $u = s + n$ as above. If $\mathrm{tr}(u\phi(s)) = 0$ for $\phi(s) = \overline{s}$ then u is nilpotent.*

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

4.4 Cartan's criterion.

Let $\mathbf{g} \subset \mathrm{End}(V)$ be a Lie subalgebra where V is finite dimensional vector space over $\mathbf{C}$. Then

$$\mathbf{g} \text{ is solvable} \Leftrightarrow \quad \mathrm{tr}(xy) = 0 \ \forall x \in \mathbf{g}, \quad y \in [\mathbf{g}, \mathbf{g}].$$

Proof. Suppose $\mathbf{g}$ is solvable. Choose a basis for which $\mathbf{g}$ is upper triangular. Then every $y \in [\mathbf{g}, \mathbf{g}]$ has zeros on the diagonal, Hence $\mathrm{tr}(xy) = 0$. For the reverse implication, it is enough to show that $[\mathbf{g}, \mathbf{g}]$ is nilpotent, and, by Engel, that each $u \in [\mathbf{g}, \mathbf{g}]$ is nilpotent. So it is enough to show that $\mathrm{tr}\, u\overline{s} = 0$, where s is the semisimple part of u, by Proposition 5 above. If it were true that $\overline{s} \in \mathbf{g}$ we would be done, but this need not be so. Write

$$u = \sum [x_i, y_i].$$

Now for $a, b, c \in \mathrm{End}(V)$

$$\begin{aligned}
\mathrm{tr}([a,b]c) &= \mathrm{tr}(abc - bac)\\
&= \mathrm{tr}(bca - bac)\\
&= \mathrm{tr}(b[c,a]) \quad \text{so}\\
\mathrm{tr}(u\overline{s}) &= \sum \mathrm{tr}([x_i, y_i]\overline{s})\\
&= \sum \mathrm{tr}(y_i[\overline{s}, x_i]).
\end{aligned}$$

So it is enough to show that $\mathrm{ad}\,\overline{s} : \mathbf{g} \to [\mathbf{g}, \mathbf{g}]$. We know that $\mathrm{ad}\,u : \mathbf{g} \to [\mathbf{g}, \mathbf{g}]$, and we can, by Lagrange interpolation, find a polynomial P such that $P(u) = \overline{s}$. The result now follows from Prop. 4:

Since $\mathrm{End}(V) \sim V_{1,1}$, take $A = [\mathbf{g}, \mathbf{g}]$ and $B = \mathbf{g}$. Then $\mathrm{ad}\,u = u_{1,1}$ so $u_{1,1}\mathbf{g} \subset [\mathbf{g}, \mathbf{g}[$ and hence $\overline{s}_{1,1}\mathbf{g} \subset [\mathbf{g}, \mathbf{g}]$ or $[\overline{s}, x] \in [\mathbf{g}, \mathbf{g}]\ \forall x \in \mathbf{g}$. QED

4.5 Radical.

If $\mathbf{i}$ is an ideal of $\mathbf{g}$ and $\mathbf{g}/\mathbf{i}$ is solvable, then $D^{(n)}(\mathbf{g}/\mathbf{i}) = 0$ implies that $D^{(n)}\mathbf{g} \subset \mathbf{i}$. If $\mathbf{i}$ itself is solvable with $D^{(m)}\mathbf{i} = 0$, then $D^{(m+n)}\mathbf{g} = 0$. So we have proved:

Proposition 6 *If* $\mathbf{i} \subset \mathbf{g}$ *is an ideal, and both* $\mathbf{i}$ *and* $\mathbf{g}/\mathbf{i}$ *are solvable, so is* $\mathbf{g}$.

If $\mathbf{i}$ and $\mathbf{j}$ are solvable ideals, then $(\mathbf{i}+\mathbf{j})/\mathbf{j} \sim \mathbf{i}/(\mathbf{i} \cap \mathbf{j})$ is solvable, being the homomorphic image of a solvable algebra. So, by the previous proposition:

Proposition 7 *If* $\mathbf{i}$ *and* $\mathbf{j}$ *are solvable ideals in* $\mathbf{g}$ *so is* $\mathbf{i}+\mathbf{j}$. *In particular, every Lie algebra* $\mathbf{g}$ *has a largest solvable ideal which contains all other solvable ideals. It is denoted by* rad $\mathbf{g}$ *or simply by* $\mathbf{r}$ *when* $\mathbf{g}$ *is fixed.*

An algebra $\mathbf{g}$ is called **semi-simple** if rad $\mathbf{g} = 0$. Since $D\mathbf{i}$ is an ideal whenever $\mathbf{i}$ is (by Jacobi's identity), if $\mathbf{r} \neq 0$ then the last non-zero $D^{(n)}\mathbf{r}$ is an abelian ideal. So an equivalent definition is: $\mathbf{g}$ is semi-simple if it has no non-zero abelian ideals.

We shall call a Lie algebra **simple** if it is not abelian and if it has no proper ideals. We shall show in the next section that every semi-simple Lie algebra is the direct sum of simple Lie algebras in a unique way.

4.6 The Killing form.

A bilinear form $(\ ,\ ) : \mathbf{g} \times \mathbf{g} \to k$ is called **invariant** if

$$([x, y], z) + (y, [x, z]) = 0 \quad \forall x, y, z \in \mathbf{g}. \tag{4.2}$$

Notice that if $(\ ,\ )$ is an invariant form, and $\mathbf{i}$ is an ideal, then $\mathbf{i}^{\perp}$ is again an ideal.

One way of producing invariant forms is from representations: if(ρ, V) is a representation of $\mathbf{g}$, then

$$(x, y)_\rho := \operatorname{tr} \rho(x)\rho(y)$$

is invariant. Indeed,

$$\begin{aligned} &([x,y],z)_\rho + (y,[x,z])_\rho \\ = \;& \operatorname{tr}\{(\rho(x)\rho(y) - \rho(y)\rho(x))\rho(z)) + \rho(y)(\rho(x)\rho(z) - \rho(z)\rho(x))\} \\ = \;& \operatorname{tr}\{\rho(x)\rho(y)\rho(z) - \rho(y)\rho(z)\rho(x)\} \\ = \;& 0. \end{aligned}$$

In particular, if we take $\rho = \operatorname{ad}$, $V = \mathbf{g}$ the corresponding bilinear form is called the **Killing form** and will be denoted by $(\ ,\ )_\kappa$. We will also sometimes write $\kappa(x, y)$ instead of $(x, y)_\kappa$.

Theorem 8 $\mathbf{g}$ *is semi-simple if and only if its Killing form is non-degenerate.*

Proof. Suppose $\mathbf{g}$ is not semi-simple and so has a non-zero abelian ideal, $\mathbf{a}$. We will show that $(x, y)_\kappa = 0\ \forall x \in \mathbf{a}, y \in \mathbf{g}$. Indeed, let $\sigma = \operatorname{ad} x \operatorname{ad} y$. Then σ maps $\mathbf{g} \to \mathbf{a}$ and $\mathbf{a} \to 0$. Hence in terms of a basis starting with elements of $\mathbf{a}$ and extending, it (is upper triangular and) has 0 along the diagonal. Hence $\operatorname{tr}\sigma = 0$. Hence if $\mathbf{g}$ is *not* semisimple then its Killing form is degenerate.

Conversely, suppose that $\mathbf{g}$ is semi-simple. We wish to show that the Killing form is non-degenerate. So let $\mathbf{u} := \mathbf{g}^\perp = \{x | \operatorname{tr} \operatorname{ad} x \operatorname{ad} y = 0\ \forall y \in \mathbf{g}\}$. If $x \in \mathbf{u}, z \in \mathbf{g}$ then

$$\begin{aligned} \operatorname{tr}\{\operatorname{ad}[x,z]\operatorname{ad} y\} &= \operatorname{tr}\{\operatorname{ad} x \operatorname{ad} z \operatorname{ad} y - \operatorname{ad} z \operatorname{ad} x \operatorname{ad} y)\} \\ &= \operatorname{tr}\{\operatorname{ad} x(\operatorname{ad} z \operatorname{ad} y - \operatorname{ad} y \operatorname{ad} z)\} \\ &= \operatorname{tr} \operatorname{ad} x \operatorname{ad}[z,y] \\ &= 0, \end{aligned}$$

so $\mathbf{u}$ is an ideal. In particular, $\operatorname{tr}_\mathbf{u}(\operatorname{ad} x_\mathbf{u} \operatorname{ad} y_\mathbf{u}) = \operatorname{tr}_\mathbf{g}(\operatorname{ad}_\mathbf{g} x \operatorname{ad}_\mathbf{g} y)$ for $x, y \in \mathbf{u}$, as can be seen from a block decomposition starting with a basis of $\mathbf{u}$ and extending to $\mathbf{g}$.

If we take $y \in D\mathbf{u}$, we see that $\operatorname{tr} \operatorname{ad} \mathbf{u} D \operatorname{ad} \mathbf{u} = 0$, so $\operatorname{ad} \mathbf{u}$ is solvable by Cartan's criterion. But the kernel of the map $\mathbf{u} \to \operatorname{ad} \mathbf{u}$ is the center of $\mathbf{u}$. So if $\operatorname{ad} \mathbf{u}$ is solvable, so is $\mathbf{u}$. QED

Proposition 8 *Let* $\mathbf{g}$ *be a semisimple algebra,* $\mathbf{i}$ *any ideal of* $\mathbf{g}$*, and* $\mathbf{i}^\perp$ *its orthocomplement with respect to its Killing form. Then* $\mathbf{i} \cap \mathbf{i}^\perp = 0$.

Indeed, $\mathbf{i} \cap \mathbf{i}^\perp$ is an ideal on which $tr \operatorname{ad} x \operatorname{ad} y \equiv 0$ hence is solvable by Cartan's criterion. Since $\mathbf{g}$ is semi-simple, there are no non-trivial solvable ideals. QED

Therefore

Proposition 9 *Every semi-simple Lie algebra is the direct sum of simple Lie algebras.*

Proposition 10 *$D\mathbf{g} = \mathbf{g}$ for a semi-simple Lie algebra.*

(Since this is true for each simple component.)

Proposition 11 *Let $\phi : \mathbf{g} \to \mathbf{s}$ be a surjective homomorphism of a semi-simple Lie algebra onto a simple Lie algebra. Then if $\mathbf{g} = \bigoplus \mathbf{g}_i$ is a decomposition of $\mathbf{g}$ into simple ideals, the restriction, ϕ_i of ϕ to each summand is zero, except for one summand where it is an isomorphism.*

Proof. Since $\mathbf{s}$ is simple, the image of every ϕ_i is 0 or all of $\mathbf{s}$. If ϕ_i is surjective for some i then it is an isomorphism since $\mathbf{g}_i$ is simple. There is at least one i for which it is surjective since ϕ is surjective. On the other hand, it can not be surjective for for two ideals, $\mathbf{g}_i, \mathbf{g}_j$ $i \neq j$ for then $\phi[\mathbf{g}_i, \mathbf{g}_j] = 0 \neq [\mathbf{s}, \mathbf{s}] = \mathbf{s}$. QED

4.7 Complete reducibility.

The basic theorem is

Theorem 9 [Weyl.] *Every finite dimensional representation of a semi-simple Lie algebra is completely reducible.*

Proof.

1. If $\rho : \mathbf{g} \to \operatorname{End} V$ is injective, then the form $(\ ,\ )_\rho$ is non-degenerate. Indeed, the ideal consisting of all x such that $(x, y)_\rho = 0\ \forall y \in \mathbf{g}$ is solvable by Cartan's criterion, hence 0.

2. The **Casimir operator.** Let (e_i) and (f_i) be bases of $\mathbf{g}$ which are dual with respect to some non-degenerate invariant bilinear form, $(,)$. So $(e_i, f_j) = \delta_{ij}$. As the form is non-degenerate and invariant, it defines a map of
$$\mathbf{g} \otimes \mathbf{g} \mapsto \operatorname{End} \mathbf{g}; \quad x \otimes y(w) = (y, w)x.$$
This map is an isomorphism and is a $\mathbf{g}$ morphism. Under this map,
$$\sum e_i \otimes f_i(w) = \sum (w, f_i)e_i = w$$
by the definition of dual bases. Hence under the inverse map
$$\operatorname{End} \mathbf{g} \mapsto \mathbf{g} \otimes \mathbf{g}$$
the identity element, id, corresponds to $\sum e_i \otimes f_i$ (and so this expression is independent of the choice of dual bases). Since id is annihilated by commutator by any element of End($\mathbf{g}$), we conclude that $\sum_i e_i \otimes f_i$ is annihilated by the action of all $(\operatorname{ad} x)_2 = \operatorname{ad} x \otimes 1 + 1 \otimes \operatorname{ad} x,\ x \in \mathbf{g}$. Indeed, for $x, e, f, y \in \mathbf{g}$ we have
$$\begin{aligned} ((\operatorname{ad} x)_2(e \otimes f))\, y &= (\operatorname{ad} xe \otimes f + e \otimes \operatorname{ad} xf)\, y \\ &= (f, y)[x, e] + ([x, f], y)e \\ &= (f, y)[x, e] - (f, [x, y])e \quad \text{by (4.2)} \\ &= ((\operatorname{ad} x)(e \otimes f) - (e \otimes f)(\operatorname{ad} x))\, y. \end{aligned}$$

Set

$$C := \sum_i e_i \cdot f_i \in U(L). \tag{4.3}$$

Thus C is the image of the element $\sum_i e_i \otimes f_i$ under the multiplication map $\mathbf{g} \otimes \mathbf{g} \mapsto U(\mathbf{g})$, and is independent of the choice of dual bases. Furthermore, C is annihilated by $\operatorname{ad} x$ acting on $U(\mathbf{g})$. In other words, it commutes with all elements of $\mathbf{g}$, and hence with all of $U(\mathbf{g})$; it is in the center of $U(\mathbf{g})$.

The C corresponding to the Killing form is called the **Casimir element**, its image in any representation is called the **Casimir operator**.

3. Suppose that $\rho : \mathbf{g} \to \operatorname{End} V$ is injective. The (image of the) central element corresponding to $(\ ,\ )_\rho$ defines an element of $\operatorname{End} V$ denoted by C_ρ and

$$\begin{aligned} \operatorname{tr} C_\rho &= \operatorname{tr} \rho(\sum e_i f_i) \\ &= \operatorname{tr} \sum \rho(e_i)\rho(f_i) \\ &= \sum_i (e_i, f_i) \\ &= \dim \mathbf{g} \end{aligned}$$

With these preliminaries, we can state the main proposition:

Proposition 12 *Let $0 \to V \to W \to k \to 0$ be an exact sequence of $\mathbf{g}$ modules, where $\mathbf{g}$ is semi-simple, and the action of $\mathbf{g}$ on k is trivial (as it must be). Then this sequence splits, i.e. there is a line in W supplementary to V on which $\mathbf{g}$ acts trivially.*

The proof of the proposition and of the theorem is almost identical to the proof we gave above for the special case of $sl(2)$. We will need only one or two additional arguments. As in the case of $sl(2)$, the proposition is a special case of the theorem we want to prove. But we shall see that it is sufficient to prove the theorem.

Proof of proposition. It is enough to prove the proposition for the case that V is an irreducible module. Indeed, if V_1 is a submodule, then by induction on dim V we may assume the theorem is known for $0 \to V/V_1 \to W/V_1 \to k \to 0$ so that there is a one dimensional invariant subspace M in W/V_1 supplementary to V/V_1 on which the action is trivial. Let N be the inverse image of M in W. By another application of the proposition, this time to the sequence

$$0 \to V_1 \to N \to M \to 0$$

we find an invariant line, P, in N complementary to V_1. So $N = V_1 \oplus P$. Since $(W/V_1) = (V/V_1) \oplus M$ we must have $P \cap V = \{0\}$. But since dim $W =$ dim $V + 1$, we must have $W = V \oplus P$. In other words P is a one dimensional subspace of W which is complementary to V.

Next we can reduce to proving the proposition for the case that $\mathbf{g}$ acts faithfully on V. Indeed, let $\mathbf{i}=$ the kernel of the action on V. For all $x \in \mathbf{g}$ we have, by hypothesis, $xW \subset V$, and for $x \in \mathbf{i}$ we have $xV = 0$. Hence $D\mathbf{i}$ acts trivially on W. But $\mathbf{i} = D\mathbf{i}$ since $\mathbf{i}$ is semi-simple. Hence $\mathbf{i}$ acts trivially on W and we may pass to $\mathbf{g}/\mathbf{i}$. This quotient is again semi-simple, since $\mathbf{i}$ is a sum of some of the simple ideals of $\mathbf{g}$.

So we are reduced to the case that V is irreducible and the action, ρ, of $\mathbf{g}$ on V is injective. Then we have an invariant element C_ρ whose image in $\operatorname{End} W$ must map $W \to V$ since every element of $\mathbf{g}$ does. (We may assume that $\mathbf{g} \neq 0$.) On the other hand, $C_\rho \neq 0$, indeed its trace is dim $\mathbf{g}$. The restriction of C_ρ to V can not have a non-trivial kernel, since this would be an invariant subspace. Hence the restriction of C_ρ to V is an isomorphism. Hence ker $C_\rho : W \to V$ is an invariant line supplementary to V. We have proved the proposition.

Proof of theorem from proposition. Let $0 \to E' \to E$ be an exact sequence of $\mathbf{g}$ modules, and we may assume that $E' \neq 0$. We want to find an invariant complement to E' in E. Define W to be the subspace of $\operatorname{Hom}_k(E, E')$ whose restriction to E' is a scalar times the identity, and let $V \subset W$ be the subspace consisting of those linear transformations whose restrictions to E' is zero. Each of these is a submodule of $\operatorname{End}(E)$. We get a sequence

$$0 \to V \to W \to k \to 0$$

and hence a complementary line of invariant elements in W. In particular, we can find an element, T which is invariant, maps $E \to E'$, and whose restriction to E' is non-zero. Then ker T is an invariant complementary subspace. QED

As an illustration of construction of the Casimir operator consider $\mathbf{g} = sl(2)$ with

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Then

$$\begin{aligned} \operatorname{tr}(\operatorname{ad} h)^2 &= 8 \\ \operatorname{tr}(\operatorname{ad} e)(\operatorname{ad} f) &= 4 \end{aligned}$$

so the dual basis to the basis h, e, f is $h/8, f/4, e/4$, or, if we divide the metric by 4, the dual basis is $h/2, f, e$ and so the Casimir operator C is

$$\frac{1}{2}h^2 + ef + fe = \frac{1}{2}h^2 + h + 2fe.$$

This coincides with the C that we used in Chapter II.

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, any 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

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

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

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

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 subspaces 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 polynomial 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 3 *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 $\operatorname{ad}(z + cx)$ 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 10 *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 ad 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 13 $\mathbf{h}$ *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 4 *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 5 $\phi : \mathbf{g} \to \mathbf{g}'$ *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 4 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 6 *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 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 $\mathbf{h}_2 = \mathbf{g}_0(\operatorname{ad} x)$ 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. Any **CSA** is nilpotent, hence solvable, hence contained in a **BSA**. This reduces the proof of the conjugacy theorem for **CSA**'s to that of **BSA**'s as we know the conjugacy of **CSA**'s in a solvable algebra. 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 7 *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 11 *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 14 *The restriction of κ to $\mathbf{g}_0 \times \mathbf{g}_0$ is non-degenerate.*

Our next intermediate step is to prove:

Proposition 15

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

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 8 *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 9 *$\mathbf{g}_0$ 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 $\mathbf{g_0}$. 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 10

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

We next prove

Lemma 11 $\mathbf{g}_0$ *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 $x \in \mathbf{h}$ 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 semi-simple 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 $sl(2)_\alpha \subset \mathbf{m}$. 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
$$\beta - r\alpha, \beta - (r-1)\alpha, \dots \beta + q\alpha$$

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 $\alpha_1, \dots, \alpha_\ell$ of $\mathbf{h}^*$ consisting of roots. This is possible because the roots span $\mathbf{h}^*$. Any root β 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

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

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 16 *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 α 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

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

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 $\gamma \in E, |\ (\gamma, \beta) \neq 0\ \forall\ \beta \in \Phi$. 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 12 *$\Delta(\gamma)$ 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 $\beta \notin \Delta(\gamma), (\gamma, \beta) = (\gamma, \beta_1) + (\gamma, \beta_2)$ 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

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

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 12 *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 $\beta_{i-1} := \alpha_i,\ \beta_j := s_{j+1}\cdots s_{i-1}\alpha_i,\quad j < i-1$. 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

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

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 (nonregular) $\gamma' \perp \alpha$, 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 13 *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 $(\operatorname{ad} t)^k n = 0$ 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

Chapter 6

The simple finite dimensional algebras.

In this chapter we classify all possible root systems of simple Lie algebras. A consequence, as we shall see, is the classification of the simple Lie algebras themselves. The amazing result - due to Killing with some repair work by Élie Cartan - is that with only five exceptions, the root systems of the classical algebras that we studied in Chapter III exhaust all possibilities.

The logical structure of this chapter is as follows: We first show that the root system of a simple Lie algebra is irreducible (definition below). We then develop some properties of the of the root structure of an irreducible root system, in particular we will introduce its extended Cartan matrix. We then use the Perron-Frobenius theorem to classify all possible such matrices. (For the expert, this means that we first classify the Dynkin diagrams of the affine algebras of the simple Lie algebras. Surprisingly, this is simpler and more efficient than the classification of the diagrams of the finite dimensional simple Lie algebras themselves.) From the extended diagrams it is an easy matter to get all possible bases of irreducible root systems. We then develop a few more facts about root systems which allow us to conclude that an isomorphism of irreducible root systems implies an isomorphism of the corresponding Lie algebras. We postpone the the proof of the existence of the exceptional Lie algebras until Chapter VIII, where we prove Serre's theorem which gives a unified presentation of all the simple Lie algebras in terms of generators and relations derived directly from the Cartan integers of the simple root system.

Throughout this chapter we will be dealing with semi-simple Lie algebras over the complex numbers.

6.1 Simple Lie algebras and irreducible root systems.

We choose a Cartan subalgebra $\mathbf{h}$ of a semi-simple Lie algebra $\mathbf{g}$, so we have the corresponding set Φ of roots and the real (Euclidean) space E that they span. We say that Φ is **irreducible** if Φ can *not* be partitioned into two disjoint subsets

$$\Phi = \Phi_1 \cup \Phi_2$$

such that every element of Φ_1 is orthogonal to every element of Φ_2.

Proposition 17 *If* $\mathbf{g}$ *is simple then* Φ *is irreducible.*

Proof. Suppose that Φ is not irreducible, so we have a decomposition as above. If $\alpha \in \Phi_1$ and $\beta \in \Phi_2$ then

$$(\alpha + \beta, \alpha) = (\alpha, \alpha) > 0 \quad \text{and} \quad (\alpha + \beta, \beta) = (\beta, \beta) > 0$$

which means that $\alpha + \beta$ can not belong to either Φ_1 or Φ_2 and so is not a root. This means that

$$[\mathbf{g}_\alpha, \mathbf{g}_\beta] = 0.$$

In other words, the subalgebra $\mathbf{g}_1$ of $\mathbf{g}$ generated by all the $\mathbf{g}_\alpha$, $\alpha \in \Phi_1$ is centralized by all the $\mathbf{g}_\beta$, so $\mathbf{g}_1$ is a proper subalgebra of $\mathbf{g}$, since if $\mathbf{g}_1 = \mathbf{g}$ this would say that $\mathbf{g}$ has a non-zero center, which is not true for any semi-simple Lie algebra. The above equation also implies that the normalizer of $\mathbf{g}_1$ contains all the $\mathbf{g}_\gamma$ where γ ranges over all the roots. But these $\mathbf{g}_\gamma$ generate $\mathbf{g}$. So $\mathbf{g_1}$ is a proper ideal in $\mathbf{g}$, contradicting the assumption that $\mathbf{g}$ is simple. QED

Let us choose a base Δ for the root system Φ of a semi-simple Lie algebra. We say that Δ is irreducible if we can not partition Δ into two non-empty mutually orthogonal sets as in the definition of irreducibility of Φ as above.

Proposition 18 Φ *is irreducible if and only if* Δ *is irreducible.*

Proof. Suppose that Φ is not irreducible, so has a decomposition as above. This induces a partition of Δ which is non-trivial unless Δ is wholly contained in Φ_1 or Φ_2. If $\Delta \subset \Phi_1$ say, then since E is spanned by Δ, this means that all the elements of Φ_2 are orthogonal to E which is impossible. So if Δ is irreducible so is Φ. Conversely, suppose that

$$\Delta = \Delta_1 \cup \Delta_2$$

is a partition of Δ into two non-empty mutually orthogonal subsets. We have proved that every root is conjugate to a simple root by an element of the Weyl group W which is generated by the simple reflections. Let Φ_1 consist of those roots which are conjugate to an element of Δ_1 and Φ_2 consist of those roots which are conjugate to an element of Δ_2. The reflections $s_\beta, \beta \in \Delta_2$ commute with the reflections s_α, $\alpha \in \Delta_1$, and furthermore

$$s_\beta(\alpha) = \alpha$$

since $(\alpha, \beta) = 0$. So any element of Φ_1 is conjugate to an element of Δ_1 by an element of the subgroup W_1 generated by the $s_\alpha,\ \alpha \in \Delta_1$. But each such reflection adds or subtracts α. So Φ_1 is in the subspace E_1 of E spanned by Δ_1 and so is orthogonal to all the elements of Φ_2. So if Φ_1 is irreducible so is Δ. QED

We are now into the business of classifying irreducible bases.

6.2 The maximal root and the minimal root.

Suppose that Φ is an irreducible root system and Δ a base, so irreducible. Recall that once we have chosen Δ, every root β is an integer combination of the elements of Δ with all coefficients non-negative, or with all coefficients non-positive. We write $\beta \succ 0$ in the first case, and $\beta \prec 0$ in the second case. This defines a partial order on the elements of E by

$$\mu \prec \lambda \text{ if and only if } \lambda - \mu = \sum_{\alpha \in \Delta} k_\alpha \alpha, \tag{6.1}$$

where the k_α are non-negative integers. This partial order will prove very important to us in representation theory.

Also, for any $\beta = \sum k_\alpha \alpha \in \Phi^+$ we define its **height** by

$$\text{ht } \beta = \sum_\alpha k_\alpha. \tag{6.2}$$

Proposition 19 *Suppose that Φ is an irreducible root system and Δ a base. Then*

- *There exists a unique $\beta \in \Phi^+$ which is maximal relative to the ordering $\prec$.*
- *This $\beta = \sum k_\alpha \alpha$ where all the k_α are positive.*
- *$(\beta, \alpha) \geq 0$ for all $\alpha \in \Delta$ and $(\beta, \alpha) > 0$ for at least one $\alpha \in \Delta$.*

Proof. Choose a $\beta = \sum k_\alpha \alpha$ which is maximal relative to the ordering. At least one of the $k_\alpha > 0$. We claim that *all* the $k_\alpha > 0$. Indeed, suppose not. This partitions Δ into Δ_1, the set of α for which $k_\alpha > 0$ and Δ_2, the set of α for which $k_\alpha = 0$. Now the scalar product of any two distinct simple roots is ≤ 0. (Recall that this followed from the fact that if $(\alpha_1, \alpha_2) > 0$, then $s_2(\alpha_1) = \alpha_1 - \langle \alpha_1, \alpha_2 \rangle \alpha_2$ would be a root whose α_1 coefficient is positive and whose α_2 coefficient is negative which is impossible.) In particular, all the $(\alpha_1, \alpha_2) \leq 0,\ \alpha_1 \in \Delta_1,\ \alpha_2 \in \Delta_2$ and so

$$(\beta, \alpha_2) \leq 0, \quad \forall\ \alpha_2 \in \Delta_2.$$

The irreducibility of Δ implies that $(\alpha_1, \alpha_2) \neq 0$ for at least one pair $\alpha_1 \in \Delta_1,\ \alpha_2 \in \Delta_2$. But this scalar product must then be negative. So

$$(\beta, \alpha_2) < 0$$

and hence

$$s_{\alpha_2}\beta = \beta - \langle\beta, \alpha_2\rangle\alpha_2$$

is a root with

$$s_{\alpha_2}\beta - \beta \succ 0$$

contradicting the maximality of β. So we have proved that all the k_α are positive. Furthermore, this same argument shows that $(\beta, \alpha) \geq 0$ for all $\alpha \in \Delta$. Since the elements of Δ form a basis of E, at least one of the scalar products must not vanish, and so be positive. We have established the second and third items in the proposition for any maximal β. We will now show that this maximal weight is unique.

Suppose there were two, β and β'. Write $\beta' = \sum k'_\alpha \alpha$ where all the $k'_\alpha > 0$. Then $(\beta, \beta') > 0$ since $(\beta, \alpha) \geq 0$ for all α and > 0 for at least one. Since $s_\beta\beta'$ is a root, this would imply that $\beta - \beta'$ is a root, unless $\beta = \beta'$. But if $\beta - \beta'$ is a root, it is either positive or negative, contradicting the maximality of one or the other. QED

Let us label the elements of Δ as $\alpha_1, \ldots, \alpha_\ell$, and let us set

$$\alpha_0 := -\beta$$

so that α_0 is the minimal root. From the second and third items in the proposition we know that

$$\alpha_0 + k_1\alpha_1 + \cdots + k_\ell\alpha_\ell = 0 \tag{6.3}$$

and that

$$\langle\alpha_0, \alpha_i\rangle \leq 0$$

for all i and < 0 for some i.

Let us take the left hand side (call it γ) of (6.3) and successively compute $\langle\gamma, \alpha_i\rangle,\ i = 0, 1, \ldots, \ell$. We obtain

$$\begin{pmatrix} 2 & \langle\alpha_1, \alpha_0\rangle & \cdots & \langle\alpha_\ell, \alpha_0\rangle \\ \langle\alpha_0, \alpha_1\rangle & 2 & \cdots & \langle\alpha_\ell, \alpha_1\rangle \\ \vdots & \vdots & \cdots & \vdots \\ \langle\alpha_0, \alpha_\ell\rangle & \cdots & \langle\alpha_{\ell-1}, \alpha_\ell\rangle & 2 \end{pmatrix} \begin{pmatrix} 1 \\ k_1 \\ \vdots \\ k_\ell \end{pmatrix} = 0.$$

This means that if we write the matrix on the left of this equation as $2I - A$, then A is a matrix with 0 on the diagonal and whose i, j entry is $-\langle\alpha_j, \alpha_i\rangle$.

So A a non-negative matrix with integer entries with the properties

- if $A_{ij} \neq 0$ then $A_{ji} \neq 0$,
- The diagonal entries of A are 0,
- A is irreducible in the sense that we can not partition the indices into two non-empty subsets I and J such that $A_{ij} = 0\ \forall\ i \in I,\ j \in J$ and
- A has an eigenvector of eigenvalue 2 with all its entries positive.

We will show that the Perron-Frobenius theorem allows us to classify all such matrices. From here it is an easy matter to classify all irreducible root systems and then all simple Lie algebras. For this it is convenient to introduce the language of graph theory.

6.3 Graphs.

An **undirected graph** $\Gamma = (N, E)$ consists of a set N (for us finite) and a subset E of the set of subsets of N of cardinality two. We call elements of N "nodes" or "vertices" and the elements of E "edges". If $e = \{i, j\} \in E$ we say that the "edge" e joins the vertices i and j or that "i and j are adjacent". Notice that in this definition our edges are "undirected": $\{i, j\} = \{j, i\}$, and we do not allow self-loops. An example of a graph is the "cycle" $A_\ell^{(1)}$ with $\ell + 1$ vertices, so $N = \{0, 1, 2, \ldots, \ell\}$ with 0 adjacent to ℓ and to 1, with 1 adjacent to 0 and to 2 etc.

The **adjacency matrix** A of a graph Γ is the (symmetric) $0 - 1$ matrix whose rows and columns are indexed by the elements of N and whose i, j-th entry $A_{ij} = 1$ if i is adjacent to j and zero otherwise.

For example, the adjacency matrix of the graph $A_3^{(1)}$ is

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

We can think of A as follows: Let V be the vector space with basis given by the nodes, so we can think of the i-th coordinate of a vector $x \in V$ as assigning the value x_i to the node i. Then $y = Ax$ assigns to i the sum of the values x_j summed over all nodes j adjacent to i.

A **path** of length r is a sequence of nodes $x_{i_1}, x_{i_2}, \ldots, x_{i_r}$ where each node is adjacent to the next. So, for example, the number of paths of length 2 joining i to j is the i, j-th entry in A^2 and similarly, the number of paths of length r joining i to j is the i, j-th entry in A^r. The graph is said to be **connected** if there is a path (of some length) joining every pair of vertices. In terms of the adjacency matrix, this means that for every i and j there is some r such that the i, j entry of A^r is non-zero. In terms of the theory of non-negative matrices (see below) this says that the matrix A is *irreducible*.

Notice that if $\mathbf{1}$ denotes the column vector all of whose entries are 1, then $\mathbf{1}$ is an eigenvector of the adjacency matrix of $A_\ell^{(1)}$, with eigenvalue 2, and all the entries of $\mathbf{1}$ are positive. In view of the Perron-Frobenius theorem to be stated below, this implies that 2 is the maximum eigenvalue of this matrix.

We modify the notion of the adjacency matrix as follows: We start with a connected graph Γ as before, but modify its adjacency matrix by replacing some of the ones that occur by positive integers a_{ij}. If, in this replacement $a_{ij} > 1$, we redraw the graph so that there is an arrow with a_{ij} lines pointing towards the node i. For example, the graph labeled $A_1^{(1)}$ in Table **Aff 1** corresponds to the matrix

$$\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$$

which clearly has $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ as an positive eigenvector with eigenvalue 2.

Similarly, diagram $A_2^{(2)}$ in Table **Aff 2** corresponds to the matrix

$$\begin{pmatrix} 0 & 4 \\ 1 & 0 \end{pmatrix}$$

which has $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ as eigenvector with eigenvalue 2. In the diagrams, the coefficient next to a node gives the coordinates of the eigenvector with eigenvalue 2, and it is immediate to check from the diagram that this is indeed an eigenvector with eigenvalue 2. For example, the 2 next to a node with an arrow pointing toward it in $C_\ell^{(1)}$ satisfies $2 \cdot 2 = 2 \cdot 1 + 2$ etc.

It will follow from the Perron Frobenius theorem to be stated and proved below, that these are the only possible connected diagrams with maximal eigenvector two.

All the graphs so far have zeros along the diagonal. If we relax this condition, and allow for any non-negative integer on the diagonal, then the only new possibilities are those given in Figure 4.

Let us call a matrix *symmetrizable* if $A_{ij} \neq 0 \Rightarrow A_{ji} \neq 0$. The main result of this chapter will be to show that the lists in the Figures 1-4 exhaust all irreducible matrices with non-negative integer matrices, which are symmetrizable and have maximum eigenvalue 2.

6.4 Perron-Frobenius.

We say that a real matrix T is **non-negative** (or **positive**) if all the entries of T are non-negative (or positive). We write $T \geq 0$ or $T > 0$. We will use these definitions primarily for square $(n \times n)$ matrices and for column vectors $= (n \times 1)$ matrices. We let

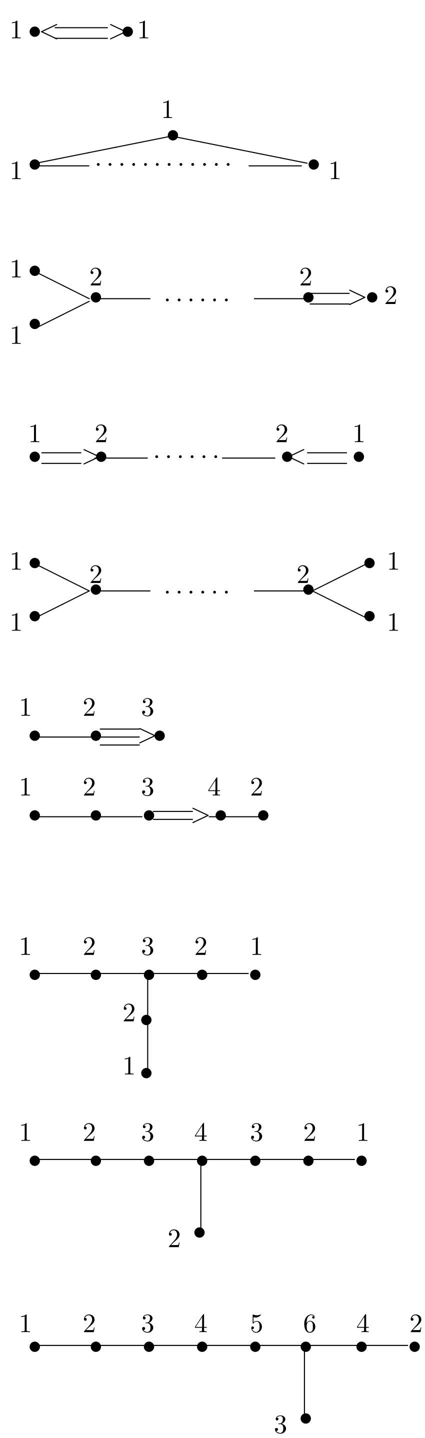
$$Q := \{x \in \mathbf{R}^n : x \geq 0, \quad x \neq 0\}$$

so Q is the non-negative "orthant" excluding the origin. Also let

$$C := \{x \geq 0 : \|x\| = 1\}.$$

So C is the intersection of the orthant with the unit sphere.

A non-negative matrix square T is called **primitive** if there is a k such that all the entries of T^k are positive. It is called **irreducible** if for any i, j there is a $k = k(i, j)$ such that $(T^k)_{ij} > 0$. For example, as mentioned above, the adjacency matrix of a connected graph is irreducible.



$A_1^{(1)}$

$A_\ell^{(1)},\ \ell \geq 2$

$B_\ell^{(1)}\ \ell \geq 3$

$C_\ell^{(1)}\ \ell \geq 2$

$D_\ell^{(1)}\ \ell \geq 4$

$G_2^{(1)}$

$F_4^{(1)}$

$E_6^{(1)}$

$E_7^{(1)}$

$E_8^{(1)}$

Figure 6.1: Aff 1.

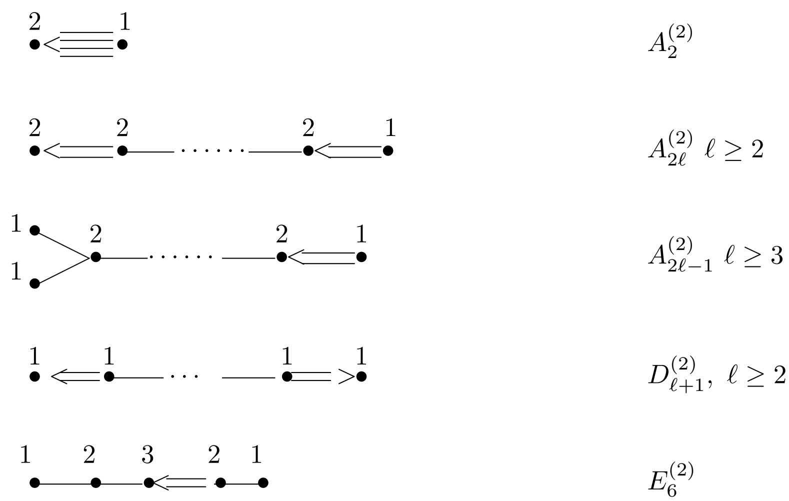
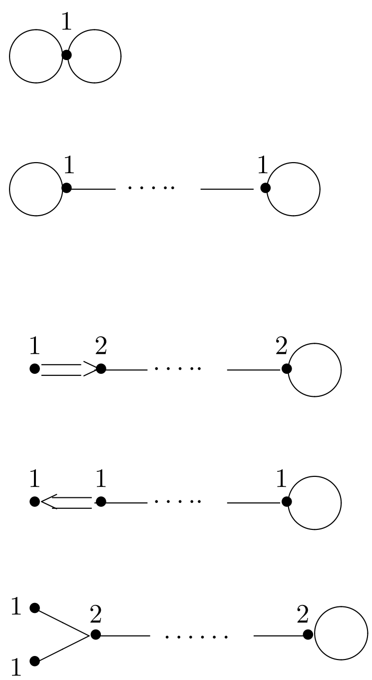


Figure 6.2: Aff 2

Figure 6.3: Aff 3



L_0

$L_\ell \quad \ell \geq 1$

$LC_\ell \ \ell \geq 1$

$LB_\ell \ \ell \geq 1$

$LD_\ell \ \ell \geq 2$

Figure 6.4: Loops allowed

If T is irreducible then $I + T$ is primitive.
In this section we will assume that T is non-negative and irreducible.

Theorem 13 Perron-Frobenius.

1. *T has a positive (real) eigenvalue $\lambda_{\max}$ such that all other eigenvalues of T satisfy*
$$|\lambda| \leq \lambda_{\max}.$$
2. *Furthermore $\lambda_{\max}$ has algebraic and geometric multiplicity one, and has an eigenvector x with $x > 0$.*
3. *Any non-negative eigenvector is a multiple of x.*
4. *More generally, if $y \geq 0,\ y \neq 0$ is a vector and μ is a number such that*
$$Ty \leq \mu y$$
then
$$y > 0, \quad \text{and} \quad \mu \geq \lambda_{\max}$$
with $\mu = \lambda_{\max}$ if and only if y is a multiple of x.
5. *If $0 \leq S \leq T,\ S \neq T$ then every eigenvalue σ of S satisfies $|\sigma| < \lambda_{\max}$.*
6. *In particular, all the diagonal minors $T_{(i)}$ obtained from T by deleting the i-th row and column have eigenvalues all of which have absolute value $< \lambda_{\max}$.*

We will present a proof of this theorem after first showing how it classifies the possible connected diagrams with maximal eigenvalue two. But first let us clarify the meaning of the last two assertions of the theorem. The matrix $T_{(i)}$ is usually thought of as an $(n-1) \times (n-1)$ matrix obtained by "striking out" the i-th row and column. But we can also consider the matrix T_i obtained from T by replacing the i-th row and column by all zeros. If x is an n-vector which is an eigenvector of T_i, then the $n-1$ vector y obtained from x by omitting the (0) i-th entry of x is then an eigenvector of $T_{(i)}$ with the same eigenvalue (unless the vector x only had non-zero entries in the i-th position). Conversely, if y is an eigenvector of $T_{(i)}$ then inserting 0 at the i-th position will give an n-vector which is an eigenvector of T_i with with the same eigenvalue as that of y.

More generally, suppose that S is obtained from T by replacing a certain number of rows and the corresponding columns by all zeros. Then we may apply item 5) of the theorem to this $n \times n$ matrix, S, or the "compressed version" of S obtained by eliminating all these rows and columns.

We will want to apply this to the following special case. A subgraph Γ' of a graph Γ is the graph obtained by eliminating some nodes, and all edges emanating from these nodes. Thus, if A is the adjacency matrix of Γ and A' is the adjacency matrix of A, then A' is obtained from A by striking out some rows and their corresponding columns. Thus if Γ is irreducible, so that we may apply the Perron Frobenius theorem to A, and if Γ' is a proper subgraph (so we have actually deleted some rows and columns of A to obtain A'), then the maximum eigenvalue of A' is strictly less than the maximum eigenvalue of A' is strictly less than the maximum eigenvalue of A. Similarly, if an entry A_{ij} is > 1, the matrix A' obtained from A by decreasing this entry while still keeping it positive will have a strictly smaller maximal eigenvalue.

We now apply this theorem to conclude that the diagrams listed in Figures Aff 1, Aff2, and Aff 3 are all possible connected diagrams with maximal eigenvalue two. A direct check shows that the vector whose coordinate at each node is the integer attached to that node given in the figure is an eigenvector with eigenvalue 2. Perron-Frobenius then guarantees 2 is the maximal eigenvalue. But now that we have shown that for each of these diagrams the maximal eigenvalue is two, any "larger" diagram must have maximal eigenvalue strictly greater than two and any "smaller" diagram must have maximal eigenvalue strictly less than two.

To get started, this argument shows that $A_1^{(1)}$ is the only diagram for which there is an i, j for which *both* a_{ij} and a_{ji} are > 1. Indeed, if A were such a matrix, by striking out all but the i and j rows and columns, we would obtain a two by two matrix whose off diagonal entries are both ≥ 2. If there were strict inequality, the maximum eigenvalue of this matrix would have to be bigger than 2 (and hence also the original diagram) by Perron Frobenius.

So other than $A_1^{(1)}$, we may assume that if $a_{ij} > 1$ then $a_{ji} = 1$.

Since any diagram with some entry $a_{ij} \geq 4$ must contain $A_2^{(2)}$ we see that this is the only diagram with this property and with maximum eigenvalue 2.

So other than this case, all $a_{ij} \leq 3$.

Diagram $G_2^{(1)}$ shows that a diagram with only two vertices and a triple bond has maximum eigenvalue strictly less than 2, since it is contained in $G_2^{(1)}$ as a subdiagram. So any diagram with a triple bond must have at least three vertices. But then it must "contain" either $G_2^{(1)}$ or $D_4^{(3)}$. But as both of these have maximal eigenvalue 2, it can not strictly contain either. So $G_2^{(1)}$ and $D_4^{(3)}$.are the only possibilities with a triple bond.

Since $A_\ell^{(1)}, \ell \geq 2$ is a cycle with maximum eigenvalue 2, no graph can contain a cycle without actually being a cycle, i.e. being $A_\ell^{(1)}$. On the other hand, a simple chain with only single bonds is contained in $A_\ell^{(1)}$, and so must have maximum eigenvalue strictly less than 2, So other than $A_\ell^{(1)}$, every candidate must contain at least one branch point or one double bond.

If the graph contains two double bonds, there are three possibilities as to the mutual orientation of the arrows, they could point toward one another as in $C_\ell^{(1)}$, away from one another as in $D_{\ell+1}^{(2)}$ or in the same direction as in $A_{2\ell}^{(2)}$. But then these are the only possibilities for diagrams with two double bonds, as no diagram can strictly contain any of them.

Also, striking off one end vertex of $C_\ell^{(1)}$ yields a graph with one extreme vertex with a double bound, with the arrow pointing away from the vertex, and no branch points. Striking out one of the two vertices at the end opposite the double bond in $B_\ell^{(1)}$ yields a graph with one extreme vertex with with a double bound and with the arrow pointing toward this vertex. So either diagram must have maximum eigenvalue < 2.

Thus if there are no branch points, there must be at least one double bond and at least two vertices on either side of the double bond. The graph with exactly two vertices on either side is strictly contained in $F_4^{(1)}$ and so is excluded. So there must be at least three vertices on one side and two on the other of the double bond. But then $F_4^{(1)}$ and $E_6^{(2)}$ exhaust the possibilities for one double bond and no branch points.

If there is a double bond and a branch point then either the double bond points toward the branch, as in $A_{2\ell-1}^{(2)}$ or away from the branch as in $B_\ell^{(1)}$. But then these exhaust the possibilities for a diagram containing both a double bond and a branch point.

If there are two branch points, the diagram must contain $D_\ell^{(1)}$ and hence must coincide with $D_\ell^{(1)}$.

So we are left with the task of analyzing the possibilities for diagrams with no double bonds and a single branch point. Let m denote the minimum number of vertices on some leg of a branch (excluding the branch point itself). If $m \geq 2$, then the diagram contains $E_6^{(1)}$ and hence must coincide with $E_6^{(1)}$. So we may assume that $m = 1$. If two branches have only one vertex emanating, then the diagram is strictly contained in $D_\ell^{(1)}$ and hence excluded. So each of the two other legs have at least two or more vertices. If both legs have more than two vertices on them, the graph must contain, and hence coincide with $E_7^{(1)}$. We are left with the sole possibility that one of the legs emanating from the branch point has one vertex and a second leg has two vertices. But then either the graph contains or is contained in $E_8^{(1)}$ so $E_8^{(1)}$ is the only such possibility.

We have completed the proof that the diagrams listed in Aff 1, Aff 2 and Aff 3 are the only diagrams without loops with maximum eigenvalue 2.

If we allow loops, an easy extension of the above argument shows that the only new diagrams are the ones in the table "Loops allowed".

6.5 Classification of the irreducible Δ.

Notice that if we remove a vertex labeled 1 (and the bonds emanating from it) from any of the diagrams in **Aff 2** or **Aff 3** we obtain a diagram which can also be obtained by removing a vertex labeled 1 from one of the diagrams in**Aff 1**. (In the diagram so obtained we ignore the remaining labels.) Indeed, removing the right hand vertex labeled 1 from $D_4^{(3)}$ yields A_2 which is obtained from $A_2^{(1)}$ by removing a vertex. Removing the left vertex marked 1 gives G_2, the diagram obtained from $G_2^{(1)}$ by removing the vertex marked 1.

Removing a vertex from $A_2^{(2)}$ gives A_1. Removing the vertex labeled 1 from $A_{2\ell}^{(2)}$ yields $B_{2\ell}$, obtained by removing one of the vertices labeled 1 from $B_\ell^{(1)}$.

Removing a vertex labeled 1 from $A^{(2)}_{2\ell-1}$ yields $D_{2\ell}$ or $C_{2\ell}$, removing a vertex labeled 1 from $D^{(2)}_{\ell+1}$ yields $B_{\ell+1}$ and removing a vertex labeled 1 from $E^{(2)}_6$ yields F_4 or C_4.

Thus all irreducible Δ correspond to graphs obtained by removing a vertex labeled 1 from the table**Aff 1**. So we have classified all possible Dynkin diagrams of all irreducible Δ . They are given in the table labeled Dynkin diagrams.

6.6 Classification of the irreducible root systems.

It is useful to introduce here some notation due to Bourbaki: A subset Φ of a Euclidean space E is called a **root system** if the following axioms hold:

- Φ is finite, spans E and does not contain 0.
- If $\alpha \in \Phi$ then the only multiples of α which are in Φ are $\pm\alpha$.
- If $\alpha \in \Phi$ then the reflection s_α in the hyperplane orthogonal to α sends Φ into itself.
- If $\alpha, \beta \in \Phi$ then $\langle \beta, \alpha \rangle \in \mathbf{Z}$,

Recall that

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

so that the reflection s_α is given by

$$s_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha.$$

We have shown that each semi-simple Lie algebra gives rise to a root system, and derived properties of the root system. If we go back to the various arguments, we will find that most of them apply to a “general” root system according to the above definition. The one place where we used Lie algebra arguments directly, was in showing that if $\beta \neq \pm\alpha$ is a root then the collection of j such that $\beta + j\alpha$ is a root forms an unbroken chain going from $-r$ to q where $r - q = \langle \beta, \alpha \rangle$. For this we used the representation theory of $sl(2)$. So we now pause to give an alternative proof of this fact based solely on that axioms above, and in the process derive some additional useful information about roots.

For any two non-zero vectors α and β in E, the cosine of the angle between them is given by

$$\|\alpha\|\|\beta\| \cos\theta = (\alpha, \beta).$$

So

$$\langle \beta, \alpha \rangle = 2\frac{\|\beta\|}{\|\alpha\|} \cos\theta.$$

Interchanging the role of α and β and multiplying gives

$$\langle \beta, \alpha \rangle \langle \alpha, \beta \rangle = 4\cos^2\theta.$$

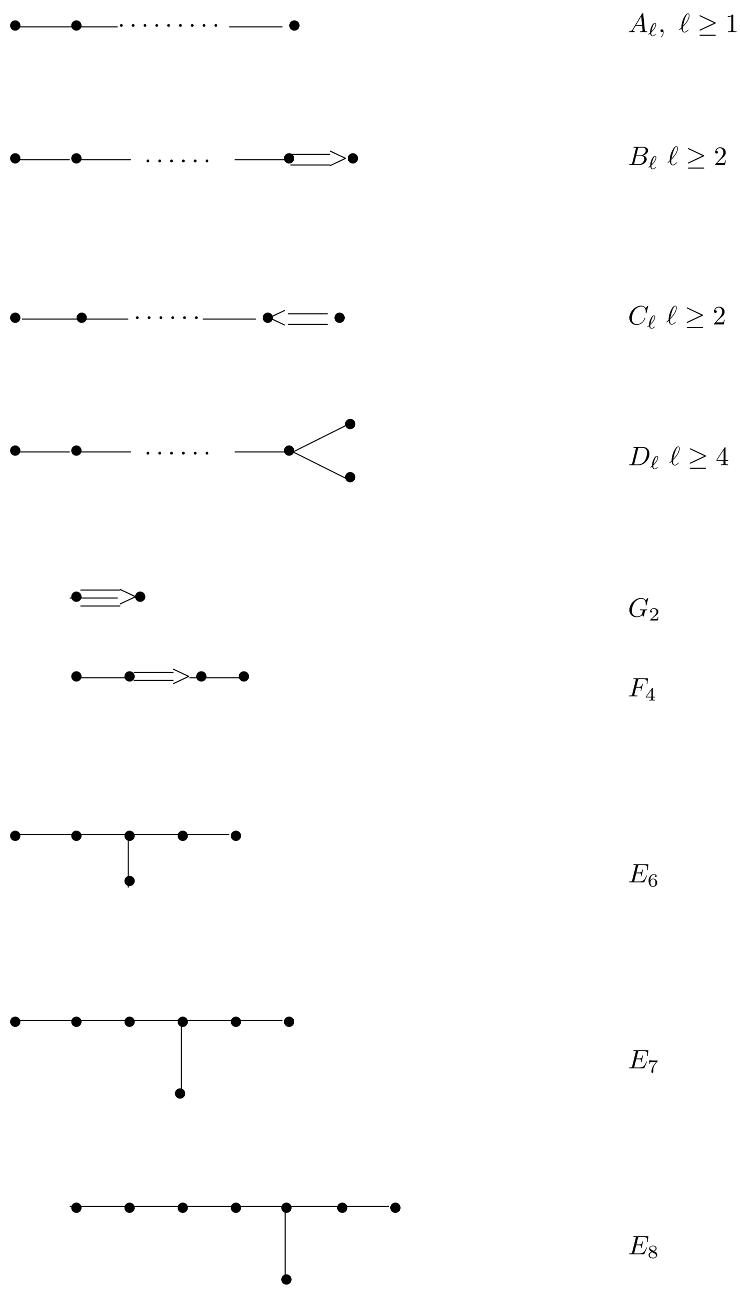


Figure 6.5: Dynkin diagrams.

The right hand side is a non-negative integer between 0 and 4. So assuming that $\alpha \neq \pm\beta$ and $\|\beta\| \geq \|\alpha\|$ The possibilities are listed in the following table:

$\langle\alpha,\beta\rangle$	$\langle\beta,\alpha\rangle$	$0 \leq \theta \leq \pi$	$\|\beta\|^2/\|\alpha\|^2$
0	0	$\pi/2$	undetermined
1	1	$\pi/3$	1
-1	-1	$2\pi/3$	1
1	2	$\pi/4$	2
-1	-1	$3\pi/4$	2
1	3	$\pi/6$	3
-1	-3	$5\pi/6$	3

Proposition 20 *If $\alpha \neq \pm\beta$ and if $(\alpha,\beta) > 0$ then $\alpha-\beta$ is a root. If $(\alpha,\beta) < 0$ then $\alpha+\beta$ is a root.*

Proof. The second assertion follows from the first by replacing β by $-\beta$. So we need to prove the first assertion. From the table, one or the other of $\langle\beta,\alpha\rangle$ or $\langle\alpha,\beta\rangle$ equals one. So either $s_\alpha\beta = \beta-\alpha$ is a root or $s_\beta\alpha = \alpha-\beta$ is a root. But roots occur along with their negatives so in either event $\alpha-\beta$ is a root. QED

Proposition 21 *Suppose that $\alpha \neq \pm\beta$ are roots. Let r be the largest integer such that $\beta - r\alpha$ is a root, and let q be the largest integer such that $\beta + q\alpha$ is a root. Then $\beta + i\alpha$ is a root for all $-r \leq i \leq q$. Furthermore $r - q = \langle\beta,\alpha\rangle$ so in particular $|q-r| \leq 3$.*

Proof. Suppose not. Then we can find a p and an s such that $-r \leq p < s \leq q$ such that $\beta + p\alpha$ is a root, but $\beta + (p+1)\alpha$ is not a root, and $\beta + s\alpha$ is a root but $\beta + (s-1)\alpha$ is not. The preceding proposition then implies that

$$\langle\beta + p\alpha, \alpha\rangle \geq 0 \quad \text{while} \quad \langle\beta + s\alpha, \alpha\rangle \leq 0$$

which is impossible since $\langle\alpha,\alpha\rangle = 2 > 0$.

Now s_α adds a multiple of α to any root, and so preserves the string of roots $\beta - r\alpha, \beta-(r-1)\alpha, \ldots, \beta + q\alpha$. Furthermore

$$s_\alpha(\beta + i\alpha) = \beta - (\langle\beta,\alpha\rangle + i)\alpha$$

so s_α reverses the order of the string. In particular

$$s_\alpha(\beta + q\alpha) = \beta - r\alpha.$$

The left hand side is $\beta - (\langle\beta,\alpha\rangle + q)\alpha$ so $r - q = \langle\beta,\alpha\rangle$ as stated in the proposition. QED

We can now apply all the preceding definitions and arguments to conclude that the Dynkin diagrams above classify all the irreducible bases Δ of root systems.

Since every root is conjugate to a simple root, we can use the Dynkin diagrams to conclude that in an irreducible root system, either all roots have the same length (cases A, D, E) or there are two root lengths - the remaining cases. Furthermore, if β denotes a long root and α a short root, the ratios $\|\beta\|^2/\|\alpha\|^2$ are 2 in the cases B, C, and F_4, and 3 for the case G_2.

Proposition 22 *In an irreducible root system, the Weyl group W acts irreducibly on E. In particular, the W-orbit of any root spans E.*

Proof. Let E' be a proper invariant subspace. Let E'' denote its orthogonal complement, so

$$E = E' \oplus E''.$$

For any root α, If $e \in E'$ then $s_\alpha e = e - \langle e, \alpha \rangle \alpha \in E'$. So either $(e, \alpha) = 0$ for all e, and so $\alpha \in E''$ or $\alpha \in E'$. Since the roots span, they can't all belong to the same subspace. This contradicts the irreducibility. QED

Proposition 23 *If there are two distinct root lengths in an irreducible root system, then all roots of the same length are conjugate under the Weyl group. Also, the maximal weight is long.*

Proof. Suppose that α and β have the same length. We can find a Weyl group element W such that $w\beta$ is not orthogonal to α by the preceding proposition. So we may assume that $\langle \beta, \alpha \rangle \neq 0$. Since α and β have the same length, by the table above we have $\langle \beta, \alpha \rangle = \pm 1$. Replacing β by $-\beta = s_\beta \beta$ we may assume that $\langle \beta, \alpha \rangle = 1$. Then

$$\begin{aligned} (s_\beta s_\alpha s_\beta)(\alpha) &= (s_\beta s_\alpha)(\alpha - \beta) \\ &= s_\beta(-\alpha - \beta + \alpha) \\ &= s_\beta(-\beta) \\ &= \beta. \text{ QED} \end{aligned}$$

Let (E, Φ) and (E', Φ') be two root systems. We say that a linear map $f : E \to E'$ is an **isomorphism** from the root system (E, Φ) to the root system (E', Φ') if f is a linear isomorphism of E onto E' with $f(\Phi) = \Phi'$ and

$$\langle f(\beta), f(\alpha) \rangle = \langle \beta, \alpha \rangle$$

for all $\alpha, \beta \in \Phi$.

Theorem 14 *Let $\Delta = \{\alpha_1, \dots, \alpha_\ell\}$ be a base of Φ. Suppose that (E', Φ') is a second root system with base $\Delta' = \{\alpha'_1, \dots, \alpha'_\ell\}$ and that*

$$\langle \alpha_i, \alpha_j \rangle = \langle \alpha'_i, \alpha'_j \rangle, \quad \forall\ 1 \leq i, j \leq \ell.$$

Then the bijection

$$\alpha_i \mapsto \alpha'_i$$

extends to a unique isomorphism $f : (E, \Phi) \to (E', \Phi')$. *In other words, the Cartan matrix* A *of* Δ *determines* Φ *up to isomorphism. In particular, The Dynkin diagrams characterize all possible irreducible root systems.*

Proof. Since Δ is a basis of E and Δ' is a basis of E', the map $\alpha_i \mapsto \alpha'_i$ extends to a unique linear isomorphism of E onto E'. The equality in the theorem implies that for $\alpha, \beta \in \Delta$ we have

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

Since the Weyl groups are generated by these simple reflections, this implies that the map

$$w \mapsto f \circ w \circ f^{-1}$$

is an isomorphism of W onto W'. Every $\beta \in \Phi$ is of the form $w(\alpha)$ where $w \in W$ and α is a simple root. Thus

$$f(\beta) = f \circ w \circ f^{-1} f(\alpha) \in \Phi'$$

so $f(\Phi) = \Phi'$. Since $s_\alpha(\beta) = \beta - \langle \beta, \alpha\rangle \alpha$, the number $\langle \beta, \alpha\rangle$ is determined by the reflection s_α acting on β. But then the corresponding formula for Φ' together with the fact that

$$s_{f(\alpha)} = f \circ s_\alpha \circ f^{-1}$$

implies that

$$\langle f(\beta), f(\alpha)\rangle = \langle \beta, \alpha\rangle.$$

QED

6.7 The classification of the possible simple Lie algebras.

Suppose that $\mathbf{g}, \mathbf{h}$, is a pair consisting of a semi-simple Lie algebra $\mathbf{g}$, and a Cartan subalgebra $\mathbf{h}$. This determines the corresponding Euclidean space E and root system Φ. Suppose we have a second such pair $(\mathbf{g}', \mathbf{h}')$. We would like to show that an isomorphism of (E, Φ) with (E', Φ') determines a Lie algebra isomorphism of $\mathbf{g}$ with $\mathbf{g}'$. This would then imply that the Dynkin diagrams classify all possible simple Lie algebras. We would still be left with the problem of showing that the exceptional Lie algebras exist. We will defer this until Chapter VIII where we prove Serre's theorem with gives a direct construction of all the simple Lie algebras in terms of generators and relations determined by the Cartan matrix.

We need a few preliminaries.

Proposition 24 *Every positive root can be written as a sum of simple roots*

$$\alpha_{i_1} + \cdots \alpha_{i_k}$$

in such a way that every partial sum is again a root.

Proof. By induction (on say the height) it is enough to prove that for every positive root β which is not simple, there is a simple root α such that $\beta-\alpha$ is a root. We can not have $(\beta,\alpha)\leq 0$ for all $\alpha\in\Delta$ for this would imply that the set $\{\beta\}\cup\Delta$ is independent (by the same method that we used to prove that Δ was independent). So $(\beta,\alpha)>0$ for some $\alpha\in\Delta$ and so $\beta-\alpha$ is a root. Since β is not simple, its height is at least two, and so subtracting α will not be zero or a negative root, hence positive. QED

Proposition 25 *Let* $\mathbf{g},\mathbf{h}$ *be a semi-simple Lie algebra with a choice of Cartan subalgebra, Let* Φ *be the corresponding root system, and let* Δ *be a base. Then* $\mathbf{g}$ *is generated as a Lie algebra by the subspaces* $\mathbf{g}_\alpha,\mathbf{g}_{-\alpha},\alpha\in\Delta$.

From the representation theory of $sl(2)_\alpha$ we know that $[\mathbf{g}_\alpha,\mathbf{g}_\beta]=\mathbf{g}_{\alpha+\beta}$ if $\alpha+\beta$ is a root. Thus from the preceding proposition, we can successively obtain all the $\mathbf{g}_\beta$ for β positive by bracketing the $\mathbf{g}_\alpha,\ \alpha\in\Delta$. Similarly we can get all the $\mathbf{g}_\beta$ for β negative from the $\mathbf{g}_{-\alpha}$. So we can get all the root spaces. But $[\mathbf{g}_\alpha,\mathbf{g}_{-\alpha}]=\mathbf{C}h_\alpha$ so we can get all of $\mathbf{h}$. The decomposition

$$\mathbf{g}=\mathbf{h}\oplus\bigoplus_{\gamma\in\Phi}\mathbf{g}_\gamma$$

then shows that we have generated all of $\mathbf{g}$.

Here is the big theorem:

Theorem 15 *Let* $\mathbf{g},\mathbf{h}$ *and* $\mathbf{g}',\mathbf{h}'$ *be simple Lie algebras with choices of Cartan subalgebras, and let* Φ,Φ' *be the corresponding root systems. Suppose there is an isomorphism*

$$f:(E,\Phi)\to(E',\Phi')$$

which is an isometry of Euclidean spaces. Extend f *to an isomorphism of*

$$\mathbf{h}^*\to\mathbf{h}'^*$$

via complexification. Let $f:\mathbf{h}\to\mathbf{h}'$ *denote the corresponding isomorphism on the Cartan subalgebras obtained by identifying* $\mathbf{h}$ *and* $\mathbf{h}'$ *with their duals using the Killing form.*

Fix a base Δ *of* Φ *and* Δ' *of* Φ'. *Choose* $0\neq x_\alpha\in\mathbf{g}_\alpha,\alpha\in\Delta$ *and* $0\neq x'_{\alpha'}\in\mathbf{g}'_{\alpha'}$. *Extend* f *to a linear map*

$$f:\mathbf{h}\oplus\bigoplus_{\alpha\in\Delta}\mathbf{g}_\alpha\to\mathbf{h}'\oplus\bigoplus_{\alpha'\in\Delta'}\mathbf{g}_{\alpha'}$$

by

$$f(x_\alpha)=x'_{\alpha'}.$$

Then f *extends to a unique isomorphism of* $\mathbf{g}\to\mathbf{g}'$.

Proof. The uniqueness is easy. Given x_α there is a unique $y_\alpha\in\mathbf{g}_{-\alpha}$ for which $[x_\alpha,y_\alpha]=h_\alpha$ so f, if it exists, is determined on the y_α and hence on all of $\mathbf{g}$ since the x_α and y_α generate $\mathbf{g}$ be the preceding proposition.

To prove the existence, we will construct the graph of this isomorphism. That is, we will construct a subalgebra $\mathbf{k}$ of $\mathbf{g} \oplus \mathbf{g}'$ whose projections onto the first and onto the second factor are isomorphisms:

Use the x_α and y_α as above, with the corresponding elements $x'_{\alpha'}$ and $y'_{\alpha'}$ in $\mathbf{g}'$. Let

$$\overline{x}_\alpha := x_\alpha \oplus x'_{\alpha'} \in \mathbf{g} \oplus \mathbf{g}'$$

and similarly define

$$\overline{y}_\alpha := y_\alpha \oplus y'_{\alpha'},$$

and

$$\overline{h_\alpha} := h_\alpha \oplus h'_{\alpha'}.$$

Let β be the (unique) maximal root of $\mathbf{g}$, and choose $x \in \mathbf{g}_\beta$. Make a similar choice of $x' \in \mathbf{g}'_{\beta'}$ where β' is the maximal root of $\mathbf{g}'$. Set

$$\overline{x} := x \oplus x'.$$

Let $\mathbf{m} \subset \mathbf{g} \oplus \mathbf{g}'$ be the subspace spanned by all the

$$\operatorname{ad} \overline{y}_{\alpha_{i_1}} \cdots \operatorname{ad} \overline{y}_{\alpha_{i_m}} \overline{x}.$$

The element $\operatorname{ad} \overline{y}_{\alpha_{i_1}} \cdots \operatorname{ad} \overline{y}_{\alpha_{i_m}} \overline{x}$ belongs to $\mathbf{g}_{\beta - \sum \alpha_{i_j}} \oplus \mathbf{g}'_{\beta' - \sum \alpha'_{i_j}}$ so

$$\mathbf{m} \cap (\mathbf{g}_\beta \oplus \mathbf{g}'_{\beta'}) \quad \text{is one dimensional .}$$

In particular $\mathbf{m}$ is a proper subspace of $\mathbf{g} \oplus \mathbf{g}'$.

Let $\mathbf{k}$ denote the subalgebra of $\mathbf{g} \oplus \mathbf{g}'$ generated by the $\overline{x}_\alpha$ the $\overline{y}_\alpha$ and the $\overline{h}_\alpha$. We claim that

$$[\mathbf{k}, \mathbf{m}] \subset \mathbf{m}.$$

Indeed, it is enough to prove that $\mathbf{m}$ is invariant under the adjoint action of the generators of $\mathbf{k}$. For the $\operatorname{ad} \overline{y}_\alpha$ this follows from the definition. For the $\operatorname{ad} \overline{h}_\alpha$ we use the fact that

$$[h, y_\alpha] = -\alpha(h) y_\alpha$$

to move the $\operatorname{ad} \overline{h}_\alpha$ past all the $\operatorname{ad} \overline{y}_\gamma$ at the cost of introducing some scalar multiple, while

$$\operatorname{ad} \overline{h}_\alpha \overline{x} = \langle \beta, \alpha \rangle x_\beta + \langle \beta', \alpha' \rangle x'_{\beta'} = \langle \beta, \alpha \rangle \overline{x}$$

because f is an isomorphism of root systems.

Finally $[x_{\alpha_1}, y_{\alpha_2}] = 0$ if $\alpha_1 \neq \alpha_2 \in \Delta$ since $\alpha_1 - \alpha_2$ is not a root. On the other hand $[x_\alpha, y_\alpha] = h_\alpha$. So we can move the $\operatorname{ad} \overline{x}_\alpha$ past the $\operatorname{ad} \overline{y}_\gamma$ at the expense of introducing an $\operatorname{ad} \overline{h}_\alpha$ every time $\gamma = \alpha$. Now $\alpha + \beta$ is not a root, since β is the maximal root. So $[x_\alpha, x_\beta] = 0$. Thus $\operatorname{ad} \overline{x}_\alpha \overline{x} = 0$, and we have proved that $[\mathbf{k}, \mathbf{m}] \subset \mathbf{m}$. But since $\mathbf{m}$ is a proper subspace of $\mathbf{g} \oplus \mathbf{g}'$, this implies that $\mathbf{k}$ is a proper subalgebra, since otherwise $\mathbf{m}$ would be a proper ideal, and the only proper ideals in $\mathbf{g} \oplus \mathbf{g}'$ are $\mathbf{g}$ and $\mathbf{g}'$.

Now the subalgebra $\mathbf{k}$ can not contain any element of the form $z \oplus 0, z \neq 0$, for it if did, it would have to contain all of the elements of the form $u \oplus 0$ since we could repeatedly apply $\mathrm{ad}\, x_\alpha$'s until we reached the maximal root space and then get all of $\mathbf{g} \oplus 0$, which would mean that $\mathbf{k}$ would also contain all of $0 \oplus \mathbf{g}'$ and hence all of $\mathbf{g} \oplus \mathbf{g}'$ which we know not to be the case. Similarly $\mathbf{k}$ can not contain any element of the form $0 \oplus z'$. So the projections of $\mathbf{k}$ onto $\mathbf{g}$ and onto $\mathbf{g}'$ are linear isomorphisms. By construction they are Lie algebra homomorphisms. Hence the inverse of the projection of $\mathbf{k}$ onto $\mathbf{g}$ followed by the projection of $\mathbf{k}$ onto $\mathbf{g}'$ is a Lie algebra isomorphism of $\mathbf{g}$ onto $\mathbf{g}'$. By construction it sends x_α to $x'_{\alpha'}$ and h_α to $h_{\alpha'}$ and so is an extension of f. QED

Chapter 7

Cyclic highest weight modules.

In this chapter, $\mathbf{g}$ will denote a semi-simple Lie algebra for which we have chosen a Cartan subalgebra, $\mathbf{h}$ and a base Δ for the roots $\Phi = \Phi^+ \cup \Phi^-$ of $\mathbf{g}$.

We will be interested in describing its finite dimensional irreducible representations. If W is a finite dimensional module for $\mathbf{g}$, then $\mathbf{h}$ has at least one simultaneous eigenvector; that is there is a $\mu \in \mathbf{h}^*$ and a $w \neq 0 \in W$ such that

$$hw = \mu(h)w \quad \forall\ h \in \mathbf{h}. \tag{7.1}$$

The linear function μ is called a **weight** and the vector v is called a **weight vector**. If $x \in \mathbf{g}_\alpha$,

$$hxw = [h,x]w + xhw = (\mu + \alpha)(h)xw.$$

This shows that the space of all vectors w satisfying an equation of the type (7.1) (for varying μ) spans an invariant subspace. If W is irreducible, then the weight vectors (those satisfying an equation of the type (7.1)) must span all of W. Furthermore, since W is finite dimensional, there must be a vector v and a linear function λ such that

$$hv = \lambda(h)v \ \ \forall\ h \in \mathbf{h}, \quad e_\alpha v = 0,\ \forall \alpha \in \Phi^+. \tag{7.2}$$

Using irreducibility again, we conclude that

$$W = U(\mathbf{g})v.$$

The module is **cyclic** generated by v. In fact we can be more precise: Let $h_1, \ldots, h_\ell$ be the basis of $\mathbf{h}$ corresponding to the choice of simple roots, let $e_i \in \mathbf{g}_{\alpha_i},\ f_i \in \mathbf{g}_{-\alpha_i}$ where $\alpha_1, \ldots, \alpha_m$ are all the positive roots. (We can choose them so that each e and f generate a little $sl(2)$.) Then

$$\mathbf{g} = \mathbf{n}_- \oplus \mathbf{h} \oplus \mathbf{n}_+,$$

where $e_1,\dots,e_m$ is a basis of $\mathbf{n}_+$, where $h_1,\dots,h_\ell$ is a basis of $\mathbf{h}$, and $f_1,\dots,f_m$ is a basis of $\mathbf{n}_-$. The Poincaré-Birkhoff-Witt theorem says that monomials of the form

$$f_1^{i_1}\cdots f_m^{i_m}h_1^{j_1}\cdots h_\ell^{j_\ell}e_1^{k_1}\cdots e_m^{k_m}$$

form a basis of $U(\mathbf{g})$. Here we have chosen to place all the e's to the extreme right, with the h's in the middle and the f's to the left. It now follows that the elements

$$f_1^{i_1}\cdots f_m^{i_m}v$$

span W. Every such element, if non-zero, is a weight vector with weight

$$\lambda-(i_1\alpha_1+\cdots+i_m\alpha_m).$$

Recall that

$$\mu\prec\lambda\quad\text{ means that }\lambda-\mu=\sum k_i\alpha_i,\quad\alpha_i>0,$$

where the k_i are non-negative integers. We have shown that every weight μ of W satisfies

$$\mu\prec\lambda.$$

So we make the definition: A cyclic highest weight module for $\mathbf{g}$ is a module (not necessarily finite dimensional) which has a vector v_+ such that

$$x_+v_+=0,\ \ \forall\ x_+\in\mathbf{n}_+,\quad hv_+=\lambda(h)v_+\ \forall h\in\mathbf{h}$$

and

$$V=U(\mathbf{g})v_+.$$

In any such cyclic highest weight module every submodule is a direct sum of its weight spaces (by van der Monde). The weight spaces V_μ all satisfy

$$\mu\prec\lambda$$

and we have

$$V=\bigoplus V_\mu.$$

Any proper submodule can not contain the highest weight vector, and so the sum of two proper submodules is again a proper submodule. Hence any such V has a unique maximal submodule and hence a unique irreducible quotient. The quotient of any highest weight module by an invariant submodule, if not zero, is again a cyclic highest weight module with the same highest weight.

7.1 Verma modules.

There is a "biggest" cyclic highest weight module, associated with any $\lambda\in\mathbf{h}^*$ called the **Verma module**. It is defined as follows: Let us set

$$\mathbf{b}:=\mathbf{h}\oplus\mathbf{n}_+.$$

Given any $\lambda \in \mathbf{h}^*$ let $\mathbf{C}_\lambda$ denote the one dimensional vector space $\mathbf{C}$ with basis z_+ and with the action of $\mathbf{b}$ given by

$$(h+\sum_{\beta\succ 0} x_\beta)z_+ := \lambda(h)z_+.$$

So it is a left $U(\mathbf{b})$ module. By the Poincaré Birkhoff Witt theorem, $U(\mathbf{g})$ is a free right $U(\mathbf{b})$ module with basis $\{f_1^{i_1}\cdots f_\ell^{i_m}\}$, and so we can form the **Verma module**

$$\mathrm{Verm}(\lambda) := U(\mathbf{g})\otimes_{U(\mathbf{b})}\mathbf{C}_\lambda$$

which is a cyclic module with highest weight vector $v_+ := 1\otimes z_+$.

Furthermore, any two *irreducible* cyclic highest weight modules with the same highest weight are isomorphic. Indeed, if V and W are two such with highest weight vector v_+, u_+, consider $V\oplus W$ which has (v_+,u_+) as a maximal weight vector with weight λ, and hence $Z := U(\mathbf{g})(v_+,u_+)$ is cyclic and of highest weight λ. Projections onto the first and second factors give non-zero homomorphisms which must be surjective. But Z has a unique irreducible quotient. Hence these must induce isomorphisms on this quotient, V and W are isomorphic.

Hence, up to isomorphism, there is a unique irreducible cyclic highest weight module with highest weight λ. We call it

$$\mathrm{Irr}(\lambda).$$

In short, we have constructed a "largest" highest weight module $\mathrm{Verm}(\lambda)$ and a "smallest" highest weight module $\mathrm{Irr}(\lambda)$.

7.2 When is $\dim \mathrm{Irr}(\lambda) < \infty$?

If $\mathrm{Irr}(\lambda)$ is finite dimensional, then it is finite dimensional as a module over any subalgebra, in particular over any subalgebra isomorphic to $sl(2)$. Applied to the subalgebra $sl(2)_i$ generated by e_i, h_i, f_i we conclude that

$$\lambda(h_i)\in\mathbf{Z}.$$

Such a weight is is called **integral**. Furthermore the representation theory of $sl(2)$ says that the maximal weight for any finite dimensional representation must satisfy

$$\lambda(h_i) = \langle\lambda,\alpha_i\rangle \geq 0$$

so that λ lies in the closure of the fundamental Weyl chamber. Such a weight is called **dominant**. So a necessary condition for $\mathrm{Irr}(\lambda)$ to be finite dimensional is that λ be dominant integral. We now show that conversely, $\mathrm{Irr}(\lambda)$ is finite dimensional whenever λ is dominant integral.

For this we recall that in the universal enveloping algebra $U(\mathbf{g})$ we have

1. $[e_j, f_i^{k+1}] = 0$, if $i\neq j$

2. $[h_j, f_i^{k+1}] = -(k+1)\alpha_i(h_j) f_i^{k+1}$
3. $[e_i, f_i^{k+1}] = -(k+1) f_i^k (k \cdot 1 - h_i)$

where the first two equations are consequences of the fact that ad is a derivation and

$$[e_i, f_j] = 0 \text{ if } i \neq j \text{ since } \alpha_i - \alpha_j \text{ is not a root}$$

and

$$[h_j, f_j] = -\alpha_j(h_i) f_j.$$

The last is a the fact about $sl(2)$ which we have proved in Chapter II. Notice that it follows from 1.) that $e_j(f_i^k)v_+ = 0$ for all k and all $i \neq j$ and from 3.) that

$$e_i f_i^{\lambda(h_i)+1} v_+ = 0$$

so that $f_i^{\lambda(h_i)+1} v_+$ is a maximal weight vector. If it were non-zero, the cyclic module it generates would be a proper submodule of $\mathrm{Irr}(\lambda)$ contradicting the irreducibility. Hence

$$f_i^{\lambda(h_i)+1} v_+ = 0.$$

So for each i the subspace spanned by $v_+, f_i v_+, \cdots, f_i^{\lambda(h_i)} v_+$ is a finite dimensional $sl(2)_i$ module. In particular $\mathrm{Irr}(\lambda)$ contains some finite dimensional $sl(2)_i$ modules. Let V' denote the sum of all such. If W is a finite dimensional $sl(2)_i$ module, then $e_\alpha W$ is again finite dimensional, thus so their sum, which is a finite dimensional $sl(2)_i$ module. Hence V' is $\mathbf{g}$- stable, hence all of $\mathrm{Irr}(\lambda)$.

In particular, the e_i and the f_i act as locally nilpotent operators on $\mathrm{Irr}(\lambda)$. So the operators $\tau_i := (\exp e_i)(\exp -f_i)(\exp e_i)$ are well defined and

$$\tau_i(\mathrm{Irr}(\lambda))_\mu = \mathrm{Irr}(\lambda)_{s_i\mu}$$

so

$$\dim \mathrm{Irr}(\lambda)_{w\mu} = \dim \mathrm{Irr}(\lambda)_\mu \quad \forall w \in \mathcal{W} \tag{7.3}$$

where $\mathcal{W}$ denotes the Weyl group. These are all finite dimensional subspaces: Indeed their dimension is at most the corresponding dimension in the Verma module $\mathrm{Verm}(\lambda)$, since $\mathrm{Irr}(\lambda)_\mu$ is a quotient space of $\mathrm{Verm}(\lambda)_\mu$. But $\mathrm{Verm}(\lambda)_\mu$ has a basis consisting of those $f_1^{k_1} \cdots f_m^{k_m} v_+$. The number of such elements is the number of ways of writing

$$\lambda - \mu = k_1\alpha_1 + \cdots k_m\alpha_m.$$

So $\dim \mathrm{Verm}(\lambda)_\mu$ is the number of m-tuplets of non-negative integers $(k_1, \ldots, k_m)$ such that the above equation holds. This number is clearly finite, and is known as $P_K(\lambda - \mu)$, the Kostant partition function of $\lambda - \mu$, which will play a central role in what follows.

Now every element of E is conjugate under W to an element of the closure of the fundamental Weyl chamber, i.e. to a μ satisfying

$$(\mu, \alpha_i) \geq 0$$

i.e. to a μ that is dominant. We claim that there are only finitely many dominant weights μ which are $\prec \lambda$, which will complete the proof of finite dimensionality. Indeed, the sum of two dominant weights is dominant, so $\lambda + \mu$ is dominant. On the other hand, $\lambda - \mu = \sum k_i \alpha_i$ with the $k_i \geq 0$. So

$$(\lambda, \lambda) - (\mu, u) = (\lambda + \mu, \lambda - \mu) = \sum k_i(\lambda + \mu, \alpha_i) \geq 0.$$

So μ lies in the intersection of the ball of radius $\sqrt{(\lambda, \lambda)}$ with the discrete set of weights $\prec \lambda$ which is finite.

We record a consequence of (7.3) which is useful under very special circumstances. Suppose we are given a finite dimensional representation of $\mathbf{g}$ with the property that each weight space is one dimensional and all weights are conjugate under $\mathcal{W}$. Then this representation must be irreducible. For example, take $\mathbf{g} = sl(n+1)$ and consider the representation of $\mathbf{g}$ on $\wedge^k(\mathbf{C}^{n+1}), 1 \leq k \leq n$. In terms of the standard basis $e_1, \ldots, e_{n+1}$ of $\mathbf{C}^{n+1}$ the elements $e_{i_1} \wedge \cdots \wedge e_{i_k}$ are weight vectors with weights $L_{i_1} + \cdots + L_{i_k}$, Where $\mathbf{h}$ consists of all diagonal traceless matrices and L_i is the linear function which assigns to each diagonal matrix its i-th entry.

These weight spaces are all one dimensional and conjugate under the Weyl group. Hence these representations are irreducible with highest weight

$$\omega_i := L_1 + \cdots + L_k$$

in terms of the usual choice of base, $h_1, \ldots, h_n$ where h_j is the diagonal matrix with 1 in the j-th position, -1 in the $j+1$-st position and zeros elsewhere. Notice that

$$\omega_i(h_j) = \delta_{ij}$$

so that the ω_i form a basis of the "weight lattice" consisting of those $\lambda \in \mathbf{h}^*$ which take integral values on $h_1, \ldots, h_n$.

7.3 The value of the Casimir.

Recall that our basis of $U(\mathbf{g})$ consists of the elements

$$f_1^{i_1} \cdots f_m^{i_m} h_1^{j_1} \cdots h_\ell^{j_\ell} e_1^{k_1} \cdots e_m^{k_m}.$$

The elements of $U(\mathbf{h})$ are then the ones with no e or f component in their expression. So we have a vector space direct sum decomposition

$$U(\mathbf{g}) = U(\mathbf{h}) \oplus (U(\mathbf{g})\mathbf{n}_+ + \mathbf{n}_- U(\mathbf{g}))\,,$$

where $\mathbf{n}_+$ and $\mathbf{n}_-$ are the corresponding nilpotent subalgebras. Let γ denote projection onto the first factor in this decomposition. Now suppose $z \in Z(\mathbf{g})$, the center of the universal enveloping algebra. In particular, $z \in U(\mathbf{g})^{\mathbf{h}}$. The eigenvalues of the monomial above under the action of $h \in \mathbf{h}$ are

$$\sum_{s=1}^m (k_s - i_s)\alpha_s(h).$$

So any monomial in the expression for z can not have f factors alone. We have proved that

$$z - \gamma(z) \in U(\mathbf{g})\mathbf{n}_+, \quad \forall\, z \in Z(\mathbf{g}). \tag{7.4}$$

For any $\lambda \in \mathbf{h}^*$, the element $z \in Z(\mathbf{g})$ acts as a scalar, call it $\chi_\lambda(z)$ on the Verma module associated to λ.

In particular, if λ is a dominant integral weight, it acts by this same scalar on the irreducible finite dimensional module associated to λ.

On the other hand, the linear map $\lambda : \mathbf{h} \to \mathbf{C}$ extends to a homomorphism, which we will also denote by λ of $U(\mathbf{h}) = S(\mathbf{h}) \to \mathbf{C}$. Explicitly, if we think of elements of $U(\mathbf{h}) = S(\mathbf{h})$ as polynomials on $\mathbf{h}^*$, then $\lambda(P) = P(\lambda)$ for $P \in S(\mathbf{h})$. Since $\mathbf{n}_+ v = 0$ if v is the maximal weight vector, we conclude from (7.4) that

$$\chi_\lambda(z) = \lambda(\gamma(z)) \quad \forall\, z \in Z(\mathbf{g}). \tag{7.5}$$

We want to apply this formula to the second order Casimir element associated to the Killing form κ. So let $k_1, \dots, k_\ell \in \mathbf{h}$ be the dual basis to $h_1, \dots, h_\ell$ relative to κ, i.e.

$$\kappa(h_i, k_j) = \delta_{ij}.$$

Let $x_\alpha \in \mathbf{g}_\alpha$ be a basis (i.e. non-zero) element and $z_\alpha \in \mathbf{g}_{-\alpha}$ be the dual basis element to x_α under the Killing form, so the second order Casimir element is

$$\mathrm{Cas}^\kappa = \sum h_i k_i + \sum_\alpha x_\alpha z_\alpha.$$

where the second sum on the right is over *all* roots. We might choose the $x_\alpha = e_\alpha$ for positive roots, and then the corresponding z_α is some multiple of the f_α. (And, for present purposes we might even choose $f_\alpha = z_\alpha$ for positive α.) The problem is that the z_α for positive α in the above expression for Cas^κ are written to the right, and we must move them to the left. So we write

$$\mathrm{Cas}^\kappa = \sum_i h_i k_i + \sum_{\alpha>0} [x_\alpha, z_\alpha] + \sum_{\alpha>0} z_\alpha x_\alpha + \sum_{\alpha<0} x_\alpha z_\alpha.$$

This expression for Cas^κ has all the $\mathbf{n}^+$ elements moved to the right; in particular, all of the summands in the last two sums annihilate v_λ. Hence

$$\gamma(\mathrm{Cas}^\kappa) = \sum_i h_i k_i + \sum_{\alpha>0} [x_\alpha, z_\alpha]$$

and

$$\chi_\lambda(\mathrm{Cas}^\kappa) = \sum_i \lambda(h_i)\lambda(k_i) + \sum_{\alpha>0} \lambda([x_\alpha, z_\alpha]).$$

For any $h \in \mathbf{h}$ we have

$$\kappa(h, [x_\alpha, z_\alpha]) = \kappa([h, x_\alpha], z_\alpha) = \alpha(h)\kappa(x_\alpha, z_\alpha) = \alpha(h)$$

so

$$[x_\alpha, z_\alpha] = t_\alpha$$

where $t_\alpha \in \mathbf{h}$ is uniquely determined by

$$\kappa(t_\alpha, h) = \alpha(h) \ \ \forall \ h \in \mathbf{h}.$$

Let $(\ ,\ )_\kappa$ denote the bilinear form on $\mathbf{h}^*$ obtained from the identification of $\mathbf{h}$ with $\mathbf{h}^*$ given by κ. Then

$$\sum_{\alpha>0} \lambda([x_\alpha, z_\alpha]) = \sum_{\alpha>0} \lambda(t_\alpha) = \sum_{\alpha>0} (\lambda, \alpha)_\kappa = 2(\lambda, \rho)_\kappa \tag{7.6}$$

where

$$\rho := \frac{1}{2} \sum_{\alpha>0} \alpha.$$

On the other hand, let the constants a_i be defined by

$$\lambda(h) = \sum_i a_i \kappa(h_i, h) \ \ \forall \ h \in \mathbf{h}.$$

In other words λ corresponds to $\sum a_i h_i$ under the isomorphism of $\mathbf{h}$ with $\mathbf{h}^*$ so

$$(\lambda, \lambda)_\kappa = \sum_{i,j} a_i a_j \kappa(h_i, h_j).$$

Since $\kappa(h_i, k_j) = \delta_{ij}$ we have

$$\lambda(k_i) = a_i.$$

Combined with $\lambda(h_i) = \sum_j a_j \kappa(h_j, h_i)$ this gives

$$(\lambda, \lambda)_\kappa = \sum_i \lambda(h_i) \lambda(k_i). \tag{7.7}$$

Combined with (7.6) this yields

$$\chi_\lambda(\mathrm{Cas}^\kappa) = (\lambda + \rho, \lambda + \rho)_\kappa - (\rho, \rho)_\kappa. \tag{7.8}$$

We now use this innocuous looking formula to prove the following: We let $\mathbf{L} = \mathbf{L}_\mathbf{g} \subset \mathbf{h}^*_\mathbf{R}$ denote the lattice of integral linear forms on $\mathbf{h}$, i.e.

$$\mathbf{L} = \{\mu \in \mathbf{h}^* | 2\frac{(\mu, \phi)}{(\phi, \phi)} \in \mathbf{Z} \ \forall \phi \in \Delta\}. \tag{7.9}$$

$\mathbf{L}$ is called the **weight lattice** of $\mathbf{g}$.

For $\mu, \lambda \in \mathbf{L}$ recall that

$$\mu \prec \lambda$$

if $\lambda - \mu$ is a sum of positive roots. Then

Proposition 26 *Any cyclic highest weight module* $Z(\lambda), \lambda \in \mathbf{L}$ *has a composition series whose quotients are irreducible modules,* $\mathrm{Irr}(\mu)$ *where* $\mu \prec \lambda$ *satisfies*

$$(\mu+\rho, \mu+\rho)_\kappa = (\lambda+\rho, \lambda+\rho)_\kappa. \tag{7.10}$$

In fact, if

$$d = \sum \dim Z(\lambda)_\mu$$

where the sum is over all μ *satisfying (7.10) then there are at most* d *steps in the composition series.*

Remark. There are only finitely many $\mu \in \mathbf{L}$ satisfying (7.10) since the set of all μ satisfying (7.10) is compact and $\mathbf{L}$ is discrete. Each weight is of finite multiplicity. Therefore d is finite.

Proof by induction on d**.** We first show that if $d = 1$ then $Z(\lambda)$ is irreducible. Indeed, if not, any proper submodule W, being the sum of its weight spaces, must have a highest weight vector with highest weight μ, say. But then

$$\chi_\lambda(\mathrm{Cas}^\kappa) = \chi_\mu(\mathrm{Cas}^\kappa)$$

since W is a submodule of $Z(\lambda)$ and Cas^κ takes on the constant value $\chi_\lambda(\mathrm{Cas}^\kappa)$ on $Z(\lambda)$. Thus μ and λ both satisfy (7.10) contradicting the assumption $d = 1$. In general, suppose that $Z(\lambda)$ is not irreducible, so has a submodule, W and quotient module $Z(\lambda)/W$. Each of these is a cyclic highest weight module, and we have a corresponding composition series on each factor. In particular, $d = d_W + d_{Z(\lambda)/W}$ so that the d's are strictly smaller for the submodule and the quotient module. Hence we can apply induction. QED

For each $\lambda \in \mathbf{L}$ we introduce a formal symbol, $e(\lambda)$ which we want to think of as an "exponential" and so the symbols are multiplied according to the rule

$$e(\mu) \cdot e(\nu) = e(\mu + \nu). \tag{7.11}$$

The *character* of a module N is defined as

$$\mathrm{ch}_N = \sum \dim N_\mu \cdot e(\mu).$$

In all cases we will consider (cyclic highest weight modules and the like) all these dimensions will be finite, so the coefficients are well defined, but (in the case of Verma modules for example) there may be infinitely many terms in the (formal) sum. Logically, such a formal sum is nothing other than a function on $\mathbf{L}$ giving the "coefficient" of each $e(\mu)$.

In the case that N is finite dimensional, the above sum is finite. If

$$f = \sum f_\mu e(\mu) \quad \text{and} \quad g = \sum g_\nu e(\nu)$$

are two finite sums, then their product (using the rule (7.11)) corresponds to convolution:

$$\left(\sum f_\mu e(\mu)\right) \cdot \left(\sum g_\nu e(\nu)\right) = \sum (f \star g)_\lambda e(\lambda)$$

where

$$(f \star g)_\lambda := \sum_{\mu+\nu=\lambda} f_\mu g_\nu.$$

So we let $\mathbf{Z}_{\text{fin}}(\mathbf{L})$ denote the set of $\mathbf{Z}$ valued functions on $\mathbf{L}$ which vanish outside a finite set. It is a commutative ring under convolution, and we will oscillate in notation between writing an element of $\mathbf{Z}_{\text{fin}}(\mathbf{L})$ as an "exponential sum" thinking of it as a function of finite support.

Since we also want to consider infinite sums such as the characters of Verma modules, we enlarge the space $\mathbf{Z}_{\text{fin}}(\mathbf{L})$ by defining $\mathbf{Z}_{\text{gen}}(\mathbf{L})$ to consist of $\mathbf{Z}$ valued functions whose supports are contained in finite unions of sets of the form $\lambda - \sum_{\alpha \succ 0} k_\alpha \alpha$. The convolution of two functions belonging to $\mathbf{Z}_{\text{gen}}(\mathbf{L})$ is well defined, and belongs to $\mathbf{Z}_{\text{gen}}(\mathbf{L})$. So $\mathbf{Z}_{\text{gen}}(\mathbf{L})$ is again a ring.

It now follows from Prop.26 that

$$\text{ch}_{Z(\lambda)} = \sum \text{ch}_{\text{Irr}(\mu)}$$

where the sum is over the finitely many terms in the composition series. In particular, we can apply this to $Z(\lambda) = \text{Verm}(\lambda)$, the Verma module. Let us order the $\mu_i \prec \lambda$ satisfying (7.10) in such a way that $\mu_i \prec \mu_j \Rightarrow i \leq j$. Then for each of the finitely many μ_i occurring we get a corresponding formula for $\text{ch}_{\text{Verm}(\mu_i)}$ and so we get collection of equations

$$\text{ch}_{\text{Verm}(\mu_j)} = \sum a_{ij} \text{ch}_{\ \text{Irr}(\mu_i)}$$

where $a_{ii} = 1$ and $i \leq j$ in the sum. We can invert this upper triangular matrix and therefore conclude that there is a formula of the form

$$\text{ch}_{\text{Irr}(\lambda)} = \sum b(\mu) \text{ch}_{\text{Verm}(\mu)} \tag{7.12}$$

where the sum is over $\mu \prec \lambda$ satisfying (7.10) with coefficients $b(\mu)$ that we shall soon determine. But we do know that $b(\lambda) = 1$.

7.4 The Weyl character formula.

We will now prove

Proposition 27 *The non-zero coefficients in (7.12) occur only when*

$$\mu = w(\lambda + \rho) - \rho$$

where $w \in W$, the Weyl group of $\mathbf{g}$, and then

$$b(\mu) = (-1)^w.$$

Here

$$(-1)^w := \det w.$$

We will prove this by proving some combinatorial facts about multiplication of sums of exponentials.

We recall our notation: For $\lambda \in \mathbf{h}^*$, $\mathrm{Irr}(\lambda)$ denotes the unique irreducible module of highest weight, λ,and $\mathrm{Verm}(\lambda)$ denotes the Verma module of highest weight λ, and more generally, $Z(\lambda)$ denotes an arbitrary cyclic module of highest weight λ. Also

$$\rho := \frac{1}{2}\sum_{\phi\in\Phi^+}\phi$$

is one half the sum of the positive roots. Let $\lambda_i, i = 1, \ldots, \dim \mathbf{h}$ be the basis of the weight lattice, L dual to the base Δ. So

$$\lambda_i(h_{\alpha_j}) = \langle\lambda_i, \alpha_j\rangle = \delta_{ij}.$$

Since $s_i(\alpha_i) = -\alpha_i$ while keeping all the other positive roots positive, we saw that this implied that

$$s_i\rho = \rho - \alpha_i$$

and therefore

$$\langle\rho, \alpha_i\rangle = 1, \quad i = 1, \ldots, \ell := \dim(\mathbf{h}).$$

In other words

$$\rho = \frac{1}{2}\sum_{\phi\in\Phi^+}\phi = \lambda_1 + \cdots + \lambda_\ell. \tag{7.13}$$

The **Kostant partition function**, $P_K(\mu)$ is defined as the number of sets of non-negative integers, k_β such that

$$\mu = \sum_{\beta\in\Phi^+} k_\beta\beta.$$

(The value is zero if μ can not be expressed as a sum of positive roots.)

For any module N and any $\mu \in \mathbf{h}^*$, N_μ denotes the weight space of weight μ. For example, in the Verma module, $\mathrm{Verm}(\lambda)$, the only non-zero weight spaces are the ones where $\mu = \lambda - \sum_{\beta\in\Phi^+} k_\beta\beta$ and the multiplicity of this weight space, i.e. the dimension of $\mathrm{Verm}(\lambda)_\mu$ is the number of ways of expressing in this fashion, i.e.

$$\dim\ \mathrm{Verm}(\lambda)_\mu = P_K(\lambda - \mu). \tag{7.14}$$

In terms of the character notation introduced in the preceding section we can write this as

$$\mathrm{ch}_{\mathrm{Verm}(\lambda)} = \sum P_K(\lambda - \mu)e(\mu).$$

To be consistent with Humphreys' notation, define the *Kostant function* p by

$$p(\nu) = P_K(-\nu)$$

and then in succinct language

$$\mathrm{ch}_{\mathrm{Verm}(\lambda)} = p(\cdot - \lambda). \tag{7.15}$$

Observe that if

$$f = \sum f(\mu)e(\mu)$$

then

$$f \cdot e(\lambda) = \sum f(\mu)e(\lambda + \mu) = \sum f(\nu - \lambda)e(\nu).$$

We can express this by saying that

$$f \cdot e(\lambda) = f(\cdot - \lambda).$$

Thus, for example,

$$\mathrm{ch}_{\mathrm{Verm}(\lambda)} = p(\cdot - \lambda) = p \cdot e(\lambda).$$

Also observe that if

$$f_\alpha = \frac{1}{1 - e(-\alpha)} := 1 + e(-\alpha) + e(-2\alpha) + \cdots$$

then

$$(1 - e(-\alpha))f_\alpha = 1$$

and

$$\prod_{\alpha \in \Phi^+} f_\alpha = p$$

by the definition of the Kostant function.

Define the function q by

$$q := \prod_{\alpha \in \Phi^+} (e(\alpha/2) - e(-\alpha/2)) = e(\rho) \prod (1 - e(-\alpha))$$

since $e(\rho) = \prod_{\alpha \in \Phi^+} e(\alpha/2)$. Notice that

$$wq = (-1)^w q.$$

It is enough to check this on fundamental reflections, but they have the property that they make exactly one positive root negative, hence change the sign of q.

We have

$$qp = e(\rho). \tag{7.16}$$

Indeed,

$$\begin{aligned} qpe(-\rho) &= \left[\prod (1 - e(-\alpha))\right] e(\rho)pe(-\rho) \\ &= \left[\prod (1 - e(-\alpha))\right] p \\ &= \prod (1 - e(-\alpha)) \prod f_\alpha \\ &= 1. \end{aligned}$$

Therefore,

$$q\mathrm{ch}_{\mathrm{Verm}(\lambda)} = qpe(\lambda) = e(\rho)e(\lambda) = e(\lambda + \rho).$$

Let us now multiply both sides of (7.12) by q and use the preceding equation. We obtain

$$q\mathrm{ch}_{\mathrm{Irr}(\lambda)} = \sum b(\mu)e(\mu+\rho)$$

where the sum is over all $\mu \prec \lambda$ satisfying (7.10), and the $b(\mu)$ are coefficients we must determine.

Now ch $_{\mathrm{Irr}(\lambda)}$ is invariant under the Weyl group W, and q transforms by $(-1)^w$. Hence if we apply $w \in W$ to the preceding equation we obtain

$$(-1)^w q\mathrm{ch}_{\mathrm{Irr}(\lambda)} = \sum b(\mu)e(w(\mu+\rho)).$$

This shows that the set of $\mu+\rho$ with non-zero coefficients is stable under W and the coefficients transform by the sign representation for each W orbit. In particular, each element of the form $\mu = w(\lambda+\rho)-\rho$ has $(-1)^w$ as its coefficient. We can thus write

$$q\mathrm{ch}_{V(\lambda)} = \sum_{w\in W}(-1)^w e(w(\lambda+\rho)) + R$$

where R is a sum of terms corresponding to $\mu+\rho$ which are not of the form $w(\lambda+\rho)$. We claim that there are no such terms and hence $R = 0$. Indeed, if there were such a term, the transformation properties under W would demand that there be such a term with $\mu+\rho$ in the closure of the Weyl chamber, i.e.

$$\mu+\rho \in \Lambda := \mathbf{L}\cap D$$

where

$$D = D_{\mathbf{g}} = \{\lambda \in E | (\lambda,\phi) \geq 0 \ \ \forall \phi \in \Delta^+\}$$

and $E = \mathbf{h}^*_{\mathbf{R}}$ denotes the space of real linear combinations of the roots. But we claim that

$$\mu \prec \lambda, \ \ (\mu+\rho,\mu+\rho) = (\lambda+\rho,\lambda+\rho), \ \ \&\ \mu+\rho \in \Lambda \implies \mu = \lambda.$$

Indeed, write $\mu = \lambda - \pi, \ \ \pi = \sum k_\alpha \alpha, \ k_\alpha \geq 0$ so

$$\begin{aligned} 0 &= (\lambda+\rho,\lambda+\rho) - (\mu+\rho,\mu+\rho) \\ &= (\lambda+\rho,\lambda+\rho) - (\lambda+\rho-\pi,\lambda+\rho-\pi) \\ &= (\lambda+\rho,\pi) + (\pi,\mu+\rho) \\ &\geq (\lambda+\rho,\pi) \ \text{ since } \mu+\rho \in \Lambda \\ &\geq 0 \end{aligned}$$

since $\lambda+\rho \in \Lambda$ and in fact lies in the interior of D. But the last inequality is strict unless $\pi = 0$. Hence $\pi = 0$. We will have occasion to use this type of argument several times again in the future. In any event we have derived the fundamental formula

$$q\mathrm{ch}_{\mathrm{Irr}(\lambda)} = \sum_{w\in W}(-1)^w e(w(\lambda+\rho)). \tag{7.17}$$

Notice that if we take $\lambda = 0$ and so the trivial representation with character 1 for $V(\lambda)$, (7.17) becomes

$$q = \sum (-1)^w e(w\rho)$$

and this is precisely the denominator in the **Weyl character formula**:

$$\textbf{WCF} \ \mathrm{ch}_{\mathrm{Irr}(\lambda)} = \frac{\sum_{w\in W}(-1)^w e(w(\lambda+\rho))}{\sum_{w\in W}(-1)^w e(w\rho)} \tag{7.18}$$

7.5 The Weyl dimension formula.

For any weight, μ we define

$$A_\mu := \sum_{w\in W}(-1)^w e(w\mu).$$

Then we can write the Weyl character formula as

$$\mathrm{ch}_{\mathrm{Irr}(\lambda)} = \frac{A_{\lambda+\rho}}{A_\rho}.$$

For any weight μ define the homomorphism Ψ_μ from the ring $\mathbf{Z}_{\mathrm{fin}}(\mathbf{L})$ into the ring of formal power series in one variable t by the formula

$$\Psi_\mu(e(\nu)) = e^{(\nu,\mu)_\kappa t}$$

(and extend linearly). The left hand side of the Weyl character formula belongs to $\mathbf{Z}_{\mathrm{fin}}(\mathbf{L})$, and hence so does the right hand side which is a quotient of two elements of $\mathbf{Z}_{\mathrm{fin}}(\mathbf{L})$. Therefore for any μ we have

$$\Psi_\mu(\mathrm{ch}_{\mathrm{Irr}(\lambda)}) = \frac{\Psi_\mu(A_{\rho+\lambda})}{\Psi_\mu(A_\rho)}.$$

$$\Psi_\mu(A_\nu) = \Psi_\nu(A_\mu) \tag{7.19}$$

for any pair of weights. Indeed,

$$\begin{aligned}
\Psi_\mu(A_\nu) &= \sum_w (-1)^w e^{(\mu,w\nu)_\kappa t}\\
&= \sum_w (-1)^w e^{(w^{-1}\mu,\nu)_\kappa t}\\
&= \sum (-1)^w e^{(w\mu,\nu)_\kappa t}\\
&= \Psi_\nu(A_\mu).
\end{aligned}$$

In particular,

$$\begin{aligned}
\Psi_\rho(A_\lambda) &= \Psi_\lambda(A_\rho)\\
&= \Psi_\lambda(q)\\
&= \Psi_\lambda\left(\prod(e(\alpha/2)-e(-\alpha/2))\right)\\
&= \prod_{\alpha\in\Phi^+}\left(e^{(\lambda,\alpha)_\kappa t/2}-e^{-(\lambda,\alpha)_\kappa t/2)}\right)\\
&= \left(\prod(\lambda,\alpha)_\kappa\right)t^{\#\Phi^+} + \text{ terms of higher degree in } t.
\end{aligned}$$

Hence

$$\Psi_\rho(\mathrm{ch}_{\mathrm{Irr}(\lambda)}) = \frac{\Psi_\rho(A_{\lambda+\rho})}{\Psi_\rho(A_\rho)} = \frac{\prod(\lambda+\rho,\alpha)_\kappa}{\prod(\rho,\alpha)_\kappa} + \quad \text{terms of positive degree in } t.$$

Now consider the composite homomorphism: first apply Ψ_ρ and then set $t = 0$. This has the effect of replacing every $e(\mu)$ by the constant 1. Hence applied to the left hand side of the Weyl character formula this gives the dimension of the representation $\mathrm{Irr}(\lambda)$. The previous equation shows that when this composite homomorphism is applied to the right hand side of the Weyl character formula, we get the right hand side of the **Weyl dimension formula**:

$$\dim \mathrm{Irr}(\lambda) = \frac{\prod_{\alpha\in\Phi^+}(\lambda+\rho,\alpha)_\kappa}{\prod_{\alpha\in\Phi^+}(\rho,\alpha)_\kappa}. \tag{7.20}$$

7.6 The Kostant multiplicity formula.

Let us multiply the fundamental equation (7.17) by $pe(-\rho)$ and use the fact (7.16) that $qpe(-\rho) = 1$ to obtain

$$\mathrm{ch}_{\mathrm{Irr}(\lambda)} = \sum_{w\in W}(-1)^w pe(-\rho)e(w(\lambda+\rho)).$$

But

$$pe(-\rho)e(w(\lambda+\rho)) = p(\cdot - w(\lambda+\rho)+\rho)$$

or, in more pedestrian terms, the left hand side of this equation has, as its coefficient of $e(\mu)$ the value

$$p(\mu+\rho-w(\lambda+\rho)).$$

On the other hand, by definition,

$$\mathrm{ch}_{\mathrm{Irr}(\lambda)} = \sum \dim(\mathrm{Irr}(\lambda)_\mu e(\mu).$$

We thus obtain Kostant's formula for the multiplicity of a weight μ in the irreducible module with highest weight λ:

$$\textbf{KMF} \quad \dim\left(\mathrm{Irr}(\lambda)\right)_\mu = \sum_{w\in W}(-1)^w p(\mu+\rho-w(\lambda+\rho)). \tag{7.21}$$

It will be convenient to introduce some notation which simplifies the appearance of the Kostant multiplicity formula: For $w \in W$ and $\mu \in \mathbf{L}$ (or in E for that matter) define

$$w \odot \mu := w(\mu + \rho) - \rho. \tag{7.22}$$

This defines another action of W on E where the "origin of the orthogonal transformations w has been shifted from 0 to $-\rho$". Then we can rewrite the Kostant multiplicity formula as

$$\dim(\mathrm{Irr}(\lambda))_\mu = \sum_{w \in W} (-1)^w P_K(w \odot \lambda - \mu) \tag{7.23}$$

or as

$$\mathrm{ch}(\mathrm{Irr}(\lambda)) = \sum_{w \in W} \sum_{\mu} (-1)^w P_K(w \odot \lambda - \mu) e(\mu), \tag{7.24}$$

where P_K is the original Kostant partition function.

For the purposes of the next section it will be useful to record the following lemma:

Lemma 14 *If ν is a dominant weight and $e \neq w \in W$ then $w \odot \nu$ is not dominant.*

Proof. If ν is dominant, so lies in the closure of the positive Weyl chamber, then $\nu + \rho$ lies in the interior of the positive Weyl chamber. Hence if $w \neq e$, then $w(\nu + \rho)(h_i) < 0$ for some i, and so $w \odot \nu = w(\nu + \rho) - \rho$ is not dominant. QED

7.7 Steinberg's formula.

Suppose that λ' and λ'' are dominant integral weights. Decompose $\mathrm{Irr}(\lambda') \otimes \mathrm{Irr}(\lambda'')$ into irreducibles, and let $n(\lambda) = n(\lambda, \lambda' \otimes \lambda'')$ denote the multiplicity of $\mathrm{Irr}(\lambda)$ in this decomposition into irreducibles (with $n(\lambda) = 0$ if $\mathrm{Irr}(\lambda)$ does not appear as a summand in the decomposition). In particular, $n(\nu) = 0$ if ν is not a dominant weight since $\mathrm{Irr}(\nu)$ is infinite dimensional in this case, so can not appear as a summand in the decomposition. In terms of characters, we have

$$\mathrm{ch}(\mathrm{Irr}(\lambda'))\,\mathrm{ch}(\mathrm{Irr}(\lambda'')) = \sum_\lambda n(\lambda)\,\mathrm{ch}(\mathrm{Irr}(\lambda)).$$

Steinberg's formula is a formula for $n(\lambda)$. To derive it, use the Weyl character formula

$$\mathrm{ch}(\mathrm{Irr}(\lambda'')) = \frac{A_{\lambda''+\rho}}{A_\rho}, \quad \mathrm{ch}(\mathrm{Irr}(\lambda)) = \frac{A_{\lambda+\rho}}{A_\rho}$$

in the above formula to obtain

$$\mathrm{ch}(\mathrm{Irr}(\lambda'))A_{\lambda''+\rho} = \sum_\lambda n(\lambda) A_{\lambda+\rho}.$$

Use the Kostant multiplicity formula (7.24) for λ':

$$\mathrm{ch}(\mathrm{Irr}(\lambda')) = \sum_{w\in W}\sum_{\mu}(-1)^w P_K(w\odot\lambda'-\mu)e(\mu)$$

and the definition

$$A_{\lambda''+\rho} = \sum_{u\in W}(-1)^u e(u(\lambda''+\rho))$$

and the similar expression for $A_{\lambda+\rho}$ to get

$$\sum_{\mu}\sum_{u,w\in W}(-1)^{uw}P_K(w\odot\lambda'-\mu))e(u(\lambda''+\rho)+\mu) =$$

$$\sum_{\lambda}\sum_{w}n(\lambda)(-1)^w e(w(\lambda+\rho)).$$

Let us make a change of variables on the right hand side, writing

$$\nu = w\odot\lambda$$

so the right hand side becomes

$$\sum_{\nu}\sum_{w}(-1)^w n(w^{-1}\odot\nu)e(\nu+\rho).$$

If ν is a dominant weight, then by Lemma 14 $w^{-1}\odot\nu$ is not dominant if $w^{-1}\neq e$. So $n(w^{-1}\odot\nu)=0$ if $w\neq 1$ and so the coefficient of $e(\nu+\rho)$ is precisely $n(\nu)$ when ν is dominant.

On the left hand side let

$$\mu = \nu - u\odot\lambda''$$

to obtain

$$\sum_{\nu,u,w}(-1)^{uw}P_K(w\odot\lambda'+u\odot\lambda''-\nu)e(\nu+\rho).$$

Comparing coefficients for ν dominant gives

$$n(\nu) = \sum_{u,w}(-1)^{uw}P_K(w\odot\lambda'+u\odot\lambda''-\nu). \tag{7.25}$$

7.8 The Freudenthal - de Vries formula.

We return to the study of a semi-simple Lie algebra **g** and get a refinement of the Weyl dimension formula by looking at the next order term in the expansion we used to derive the Weyl dimension formula from the Weyl character formula.

By definition, the Killing form restricted to the Cartan subalgebra **h** is given by

$$\kappa(h,h') = \sum_{\alpha}\alpha(h)\alpha(h')$$

where the sum is over all roots. If $\mu, \lambda \in \mathbf{h}^*$ with t_μ, t_λ the elements of H corresponding to them under the Killing form, we have

$$(\lambda, \mu)_\kappa = \kappa(t_\lambda, t_\mu) = \sum_\alpha \alpha(t_\lambda)\alpha(t_\mu)$$

so

$$(\lambda, \mu)_\kappa = \sum_\alpha (\lambda, \alpha)_\kappa (\mu, \alpha)_\kappa. \tag{7.26}$$

For each λ in the weight lattice L we have let $e(\lambda)$ denote the "formal exponential" so $\mathbf{Z}_{fin}(L)$ is the space spanned by the $e(\lambda)$ and we have defined the homomorphism

$$\Psi_\rho : \mathbf{Z}_{fin}(\Lambda) \to \mathbf{C}[[t]], \quad e(\lambda) \mapsto e^{(\lambda.\rho)_\kappa t}.$$

Let N and D be the images under Ψ_ρ of the Weyl numerator and denominator. So

$$N = \Psi_\rho(A_{\rho+\lambda}) = \Psi_{\rho+\lambda}(A_\rho)$$

by (7.19) and

$$A_\rho = q = \prod_{\alpha \in \Phi^+} \left(e^{\alpha/2} - e^{-\alpha/2} \right) \tag{7.27}$$

and therefore

$$\begin{aligned} N(t) &= \prod_{\alpha>0} \left(e^{(\lambda+\rho,\alpha)_\kappa t/2} - e^{-(\lambda+\rho,\alpha)_\kappa t/2} \right) \\ &= \prod \left((\lambda+\rho,\alpha)_\kappa t[1 + \frac{1}{24}(\lambda+\rho,\alpha)_\kappa^2 t^2 + \cdots] \right) \end{aligned}$$

with a similar formula for D. Then $N/D \to d(\lambda) =$ the dimension of the representation as $t \to 0$ is the usual proof (that we reproduced above) of the Weyl dimension formula. Sticking this in to N/D gives

$$\frac{N}{D} = d(\lambda) \left(1 + \frac{1}{24} \sum_{\alpha>0} [(\lambda+\rho,\alpha)_\kappa^2 - (\rho,\alpha)_\kappa^2] t^2 + \cdots \right).$$

For any weight μ we have $(\mu,\mu)_\kappa = \sum (\mu,\alpha)_\kappa^2$ by (7.26), where the sum is over *all* roots so

$$\frac{N}{D} = d\left(1 + \frac{1}{48}[(\lambda+\rho,\lambda+\rho)_\kappa - (\rho,\rho)_\kappa] t^2 + \cdots \right),$$

and we recognize the coefficient of $\frac{1}{48}t^2$ in the above expression as $\chi_\lambda(\mathrm{Cas}^\kappa))$, the scalar giving the value of the Casimir associated to the Killing form in the representation with highest weight λ.

On the other hand, the image under Ψ_ρ of the character of the irreducible representation with highest weight λ is

$$\sum_\mu e^{(\mu,\rho)_\kappa t} = \sum_\mu (1 + (\mu,\rho)_\kappa t + \frac{1}{2}(\mu,\rho)_\kappa^2 t^2 + \cdots)$$

where the sum is over all weights in the irreducible representation counted with multiplicity. Comparing coefficients gives

$$\sum_\mu (\mu,\rho)_\kappa^2 = \frac{1}{24} d(\lambda)\chi_\lambda(\mathrm{Cas}^\kappa).$$

Applied to the adjoint representation the left hand side becomes $(\rho,\rho)_\kappa$ by (7.26), while $d(\lambda)$ is the dimension of the Lie algebra. On the other hand, $\chi_\lambda(\mathrm{Cas}^\kappa) = 1$ since $\mathrm{tr}\,\mathrm{ad}(\mathrm{Cas}^\kappa) = \dim(\mathbf{g})$ by the definition of Cas^κ. So we get

$$(\rho,\rho)_\kappa = \frac{1}{24}\dim\,\mathbf{g} \tag{7.28}$$

for any semisimple Lie algebra $\mathbf{g}$.

An algebra which is the direct sum a commutative Lie and a semi-simple Lie algebra is called reductive. The previous result of Freudenthal and deVries has been generalized by Kostant from a semi-simple Lie algebra to all reductive Lie algebras: Suppose that $\mathbf{g}$ is merely reductive, and that we have chosen a symmtric bilinear form on $\mathbf{g}$ which is invariant under the adjoint representation, and denote the associated Casimir element by $\mathrm{Cas}_{\mathbf{g}}$. We claim that (7.28) generalizes to

$$\frac{1}{24}\,\mathrm{tr}\,\mathrm{ad}(\mathrm{Cas}_{\mathbf{g}}) = (\rho,\rho). \tag{7.29}$$

(Notice that if $\mathbf{g}$ is semisimple and we take our symmetric bilinear form to be the Killing form $(\ ,\ )_\kappa$ (7.29) becomes (7.28).) To prove (7.29) observe that both sides decompose into sums as we decompose $\mathbf{g}$ into as sum of its center and its simple ideals, since this must be an orthogonal decomposition for our invariant scalar product. The contribution of the center is zero on both sides, so we are reduced to proving (7.29) for a simple algebra. Then our symmetric biinear form $(\ ,\ )$ must be a scalar multiple of the Killing form:

$$(\ ,\ ) = c^2(\ ,\ )_\kappa$$

for some non-zero scalar c. If $z_1,\dots,z_N$ is an orthonormal basis of $\mathbf{g}$ for $(\ ,\ )_\kappa$ then $z_1/c,\dots,z_N/c$ is an orthonormal basis for $(\ ,\ )$. Thus

$$\mathrm{Cas}_{\mathbf{g}} = \frac{1}{c^2}\,\mathrm{Cas}^\kappa\,.$$

So

$$\mathrm{tr}\,\mathrm{ad}(\mathrm{Cas}_{\mathbf{g}}) = \frac{1}{c^2}\,\mathrm{tr}\,\mathrm{ad}(\mathrm{Cas}^\kappa) = \frac{1}{c^2}\frac{1}{24}\dim\,\mathbf{g}.$$

But on $\mathbf{h}^*$ we have the dual relation

$$(\rho,\rho) = \frac{1}{c^2}(\rho,\rho)_\kappa.$$

Combining the last two equations shows that (7.29) becomes (7.28).

Notice that the same proof shows that we can generalize (7.8) as

$$\chi_\lambda(\mathrm{Cas}) = (\lambda+\rho,\lambda+\rho) - (\rho,\rho) \tag{7.30}$$

valid for any reductive Lie algebra equipped with a symmetric bilinear form invariant under the adjoint representation.

7.9 Fundamental representations.

We let ω_i denote the weight which satisfies

$$\omega_i(h_j) = \delta_{ij}$$

so that the ω_i form an integral basis of $\mathbf{L}$ and are dominant. We call these the **basic** weights.If (V,ρ) and (W,σ) are two finite dimensional irreducible representations with highest weights λ and σ, then $V \otimes W, \rho \otimes \sigma$ contains the irreducible representation with highest weight $\lambda + \mu$, and highest weight vector $v_\lambda \otimes w_\mu$, the tensor product of the highest weight vectors in V and W. Taking this "highest" component in the tensor product is known as the **Cartan product** of the two irreducible representations.

Let (V_i, ρ_i) be the irreducible representations corresponding to the basic weight ω_i. Then every finite dimensional irreducible representation of $\mathbf{g}$ can be obtained by Cartan products from these, and for that reason they are called the **fundamental representations**.

For the case of $A_n = sl(n+1)$ we have already verified that the fundamental representations are $\wedge^k(V)$ where $V = \mathbf{C}^{n+1}$ and where the basic weights are

$$\omega_i = L_1 + \cdots + L_i$$

We now sketch the results for the other classical simple algebras, leaving the details as an exercise in the use of the Weyl dimension formula.

For $C_n = sp(2n)$ it is immediate to check that these same expressions give the basic weights. However while $V = \mathbf{C}^{2n} = \wedge^1(V)$ is irreducible, the higher order exterior powers are not: Indeed, the symplectic form $\Omega \in \wedge^2(V^*)$ is preserved, and hence so is the the map

$$\wedge^j(V) \to \wedge^{j-2}(V)$$

given by contraction by Ω. It is easy to check that the image of this map is surjective (for $j = 2, \ldots, n$). the kernel is thus an invariant subspace of dimension

$$\binom{2n}{j} - \binom{2n}{2j-2}$$

and a (not completely trivial) application of the Weyl dimension formula will show that these are indeed the dimensions of the irreducible representations with highest weight ω_j. Thus these kernels are the fundamental representations of C_n. Here are some of the details:

We have

$$\rho = \omega_1 + \cdots + \omega_n = \sum (n - i + 1)L_i.$$

The most general dominant weight is of the form

$$\sum k_i\omega_i = a_1L_1 + \cdots + a_nL_n$$

where

$$a_1 = k_1 + \cdots + k_n, \quad a_2 = k_2 + \cdots + k_n, \cdots a_n = k_n$$

where the k_i are non-negative integers. So we can equally well use any decreasing sequence $a_1 \geq a_2 \geq \cdots \geq a_n \geq 0$ of integers to parameterize the irreducible representations. We have

$$(\rho, L_i - L_j) = j - i, \quad (\rho, L_i + L_j) = 2n + 2 - i - j.$$

Multiplying these all together gives the denominator in the Weyl dimension formula.

Similarly the numerator becomes

$$\prod_{i<j}(l_i - l_j) \prod_{i \leq j.}(l_i + l_j)$$

where

$$l_i := a_i + n - i + 1.$$

If we set $m_i := n - i + 1$ then we can write the Weyl dimension formula as

$$\dim V(a_1, \ldots, a_n) = \prod_{i<j} \frac{l_i^2 - l_j^2}{m_i^2 - m_j^2} \prod_i \frac{l_i}{m_i},$$

where for the case $i = j$ we have taken out a common factor of 2^n from the numerator and the denominator.

An easy induction shows that

$$\prod_{i<j}(m_i^2 - m_j^2) \prod_i mi = (2n-1)!(2n-3)! \cdots 1!.$$

so if we set

$$r_i = l_i - 1 = a + n - i$$

then

$$\dim V(a_1, \ldots, a_n) = \frac{\prod_{i<j}(r_i - r_j)(r_i + r_j + 2) \prod_i (r_i + 1)}{(2n-1)!(2n-3)! \cdots 1!}.$$

For example, suppose we want to compute the dimension of the fundamental representation corresponding to $\lambda_2 = L_1 + L_2$ so $a_1 = a_2 = 1, a_i = 0, i > 2$. In applying the preceding formula, all of the terms with $2 < i$ are the same as for the trivial representation, as is $r_1 - r_2$. The ratios of the remaining factors to those of the trivial representation are

$$\prod_{j=3}^{n} \frac{j}{j-1} \cdot \prod_{j=3}^{n} \frac{j}{j-2} = \prod_{j=3}^{n} \frac{j}{j-2}$$

coming from the $r_i - r_j$ terms, $i = 1, 2$. Similarly the $r_i + r_j$ terms give a factor

$$\frac{2n+1}{2n-1} \prod_{j=3}^{n} \frac{2n+2-j}{2n-j}$$

and the terms r_1+1, r_2+1 contribute a factor

$$\frac{n+1}{n-1}.$$

In multiplying all of these terms together there is a huge cancellation and what is left for the dimension of this fundamental representation is

$$\frac{(2n+1)(2n-2)}{2}.$$

Notice that this equals

$$\left(\begin{array}{c} 2n \\ 2 \end{array} \right) - 1 = \dim \ \wedge^2 V - 1.$$

More generally this dimension argument will show that the fundamental representations are the kernels of the contraction maps $i(\Omega) : \wedge^k \to (V) \wedge^{k-2} (V)$ where Ω is the symplectic form.

For B_n it is easy to check that $\omega_i := L_1 + \cdots + L_i \ (i \leq n-1)$, and $\omega_n = \frac{1}{2}(L_1 + \cdots + L_n)$ are the basic weights and the Weyl dimension formula gives the value $\binom{2n+1}{j}$ for $j \leq n-1$ as the dimensions of the irreducibles with these weight, so that they are $\wedge^j(V), \ j = 1, \ldots n-1$ while the dimension of the irreducible corresponding to ω_n is 2^n. This is the spin representation which we will study later.

Finally, for $D_n = o(2n)$ the basic weights are

$$\omega_j = L_1 + \cdots + L_j, \ j \leq n-2,$$

and

$$\omega_{n-1} := \frac{1}{2}(L_1 + \cdots + L_{n-1} + L_n) \text{ and } \omega_n := \frac{1}{2}(L_1 + \cdots + L_{n-1} - L_n).$$

The Weyl dimension formula shows that the the first $n-2$ fundamental representations are in fact the representation on $\wedge^j(V), \ j = 1, \ldots, n-2$ while the last two have dimension 2^{n-1}. These are the half spin representations which we will also study later.

7.10 Equal rank subgroups.

In this section we present a generalization of the Weyl character formula due to Ramond-Gross-Kostant-Sternberg. It depends on an interpretation of the Weyl denominator in terms of the spin representation of the orthogonal group $O(\mathbf{g}/\mathbf{h})$, and so on some results which we will prove in Chapter IX. But its logical place is in this chapter. So we will quote the results that we will need. You might prefer to read this section after Chapter IX.

Let $\mathbf{p}$ be an even dimensional space with a symmetric bilinear such that

$$\mathbf{p} = \mathbf{p}_+ \oplus \mathbf{p}_-$$

is a direct sum decomposition of $\mathbf{p}$ into two isotropic subspaces. In other words $\mathbf{p}_+$ and $\mathbf{p}_-$ are each half the dimension of $\mathbf{p}$, and the scalar product of any two vectors in $\mathbf{p}_+$ vanishes, as does the scalar product of any two elements of $\mathbf{p}_-$. For example, we might take $\mathbf{p} = \mathbf{n}_+ \oplus \mathbf{n}_-$ and the symmetric bilinear form to be the Killing form. Then $\mathbf{p}_\pm = \mathbf{n}_\pm$ is such a desired decomposition.

The symmetric bilinear form then puts $\mathbf{p}_\pm$ into duality, i.e. we may identify $\mathbf{p}_-$ with $\mathbf{p}_+^*$ and vice versa. Suppose that we have a commutative Lie algebra $\mathbf{h}$ acting on $\mathbf{p}$ as infinitesimal isometries, so as to preserve each $\mathbf{p}_\pm$, that the e_i^+ are weight vectors corresponding to weights β_i and that the e_i^- form the dual basis, corresponding to the negative of these weights $-\beta_i$. In particular, we have a Lie algebra homomorphism ν from $\mathbf{h}$ to $o(\mathbf{p})$, and the two spin representations of $o(\mathbf{p})$ give two representations of $\mathbf{h}$. By abuse of language, let us denote these two representations by $\mathrm{Spin}_{+\nu}$ and $\mathrm{Spin}_{-\nu}$. We can also consider the characters of these representations of $\mathbf{h}$. According to equation (9.22) (to be proved in Chapter IX) we have

$$\mathrm{ch}_{Spin_{+\nu}} - \mathrm{ch}_{Spin_{-\nu}} = \prod_j \left(e(\frac{1}{2}\beta_j) - e(-\frac{1}{2}\beta_j) \right).$$

In the case that $\mathbf{h}$ is the Cartan subalgebra of a semi-simple Lie algebra and and $\mathbf{p}_\pm = \mathbf{n}_\pm$ we recognize this expression as the Weyl denominator.

Now let $\mathbf{g}$ be a semi-simple Lie algebra and $\mathbf{r} \subset \mathbf{g}$ a reductive subalgebra of the same rank. This means that we can choose a Cartan subalgebra of $\mathbf{g}$ which is also a Cartan subalgebra of $\mathbf{r}$. The roots of $\mathbf{r}$ form a subset of the roots of $\mathbf{g}$. The Weyl group $W_{\mathbf{g}}$ acts simply transitively on the Weyl chambers of $\mathbf{g}$ each of which is contained in a Weyl chamber for $\mathbf{r}$. We choose a positive root system for $\mathbf{g}$, which then determines a positive root system for $\mathbf{r}$, and the positive Weyl chamber for $\mathbf{g}$ is contained in the positive Weyl chamber for $\mathbf{r}$.

Let

$$C \subset W_{\mathbf{g}}$$

denote the set of those elements of the Weyl group of $\mathbf{g}$ which map the positive Weyl chamber of $\mathbf{g}$ into the positive Weyl chamber for $\mathbf{r}$. By the simple transitivity of the Weyl group actions on chambers, we know that elements of C form coset representatives for the subgroup $W_{\mathbf{r}} \subset W_{\mathbf{g}}$. In particular, the number of elements of C is the same as the index of $W_{\mathbf{r}}$ in $W_{\mathbf{g}}$.

Let

$$\rho_{\mathbf{g}} \quad \text{and} \quad \rho_{\mathbf{r}}$$

denote half the sum of the positive roots of $\mathbf{g}$ and $\mathbf{r}$ respectively. For any dominant weight λ of $\mathbf{g}$ the weight $\lambda + \rho_{\mathbf{g}}$ lies in the interior of the positive Weyl chamber for $\mathbf{g}$. Hence for each $c \in C$, the element $c(\lambda + \rho_{\mathbf{g}})$ lies in the interior for $\mathbf{r}$ and hence

$$c \bullet \lambda := c(\lambda + \rho_{\mathbf{g}}) - \rho_{\mathbf{r}}$$

is a dominant weight for $\mathbf{r}$, and each of these is distinct.

Let V_λ denote the irreducible representation of $\mathbf{g}$ with highest weight λ. We can consider it as a representation of the subalgebra $\mathbf{r}$. Also the Killing form (or more generally any ad invariant symmetric bilinear form) on $\mathbf{g}$ induces an invariant form on $\mathbf{r}$. Let $\mathbf{p}$ denote the orthogonal complement of $\mathbf{r}$ in $\mathbf{g}$. We thus get a homomorphism of $\mathbf{r}$ into the orthogonal algebra $o(\mathbf{g}/\mathbf{r})$, which is an even dimensional orthogonal algebra, and hence has two spin representations. To specify which of these two spin representations we shall denote by S_+ and which by S_-, we note that there is a one dimensional weight space with weight $\rho_\mathbf{g} - \rho_\mathbf{r}$, and we let S_+ denote the spin representation which contains that one dimensional space. The spaces $S_\pm$ are $o(\mathbf{g}/\mathbf{r})$ modules, and via the homomorphism $\mathbf{r} \to o(\mathbf{g}/\mathbf{r})$ we can consider them as $\mathbf{r}$ modules.

Finally, for any dominant integral weight μ of $\mathbf{r}$ we let U_μ denote the irreducible module of $\mathbf{r}$ with highest weight μ.

With all this notation we can now state

Theorem 16 **[G-K-R-S]** *In the representation ring $R(\mathbf{r})$ we have*

$$V_\lambda \otimes S_+ - V_\lambda \otimes S_- = \sum_{c \in C} (-1)^c U_{c \bullet \lambda}. \tag{7.31}$$

Proof. To say that the above equation holds in the representation ring of $\mathbf{r}$ means that when we take the signed sums of the characters of the representations occurring on both sides we get equality. In the special case that $\mathbf{r} = \mathbf{h}$, we have observed that (7.31) is just the Weyl character formula:

$$\chi(\mathrm{Irr}(\lambda)(\chi(S_{+\mathbf{g}/\mathbf{h}}) - \chi(S_{-\mathbf{g}/\mathbf{h}})) = \sum_{w \in W_\mathbf{g}} (-1)^w e(w(\lambda + \rho_\mathbf{g})).$$

The general case follows from this special case by dividing both sides of this equation by $\chi(S_{+\mathbf{r}/\mathbf{h}}) - \chi(S_{-\mathbf{r}/\mathbf{h}})$. The left hand side becomes the character of the left hand side of (7.31) because the weights that go into this quotient via (9.22) are exactly those roots of $\mathbf{g}$ which are not roots of $\mathbf{r}$. The right hand side becomes the character of the right hand side of (9.22) by reorganizing the sum and using the Weyl character formula for $\mathbf{r}$. QED

Chapter 8

Serre's theorem.

We have classified all the possibilities for an irreducible Cartan matrix via the classification of the possible Dynkin diagrams. The four major series in our classification correspond to the classical simple algebras we introduced in Chapter III. The remaining five cases also correspond to simple algebras - the "exceptional algebras". Each deserves a discussion on its own. However a theorem of Serre guarantees that starting with any Cartan matrix, there is a corresponding semi-simple Lie algebra. Any root system gives rise to a Cartan matrix. So even before studying each of the simple algebras in detail, we know in advance that they exist, provided that we know that the corresponding root system exists. We present Serre's theorem in this chapter. At the end of the chapter we show that each of the exceptional root systems exists. This then proves the existence of the exceptional simple Lie algebras.

8.1 The Serre relations.

Recall that if α and β are roots,

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

and the string of roots of the form $\beta + j\alpha$ is unbroken and extends from

$$\beta - r\alpha \;\; \text{to} \;\; \beta + q\alpha \;\; \text{where} \;\; r - q = \langle \beta, \alpha \rangle.$$

In particular, if $\alpha, \beta \in \Delta$ so that $\beta - \alpha$ is not a root, the string is

$$\beta, \beta + \alpha, \ldots, \beta + q\alpha$$

where

$$q = -\langle \beta, \alpha \rangle.$$

Thus

$$(\mathrm{ad} e_\alpha)^{-\langle \beta, \alpha \rangle + 1} e_\beta = 0,$$

for $e_\alpha \in \mathbf{g}_\alpha,\ e_\beta \in \mathbf{g}_\beta$ but

$$(\mathrm{ad}\ e_\alpha)^k e_\beta \neq 0 \quad \text{for } 0 \leq k \leq -\langle \beta, \alpha \rangle,$$

if $e_\alpha \neq 0,\ e_\beta \neq 0$. So if $\Delta = \{\alpha_1, \ldots, \alpha_\ell\}$ we may choose

$$e_i \in \mathbf{g}_{\alpha_i},\ f_i \in \mathbf{g}_{-\alpha_i}$$

so that

$$e_1, \ldots, , e_\ell,\ f_1, \ldots, f_\ell$$

generate the algebra and

$$[h_i, h_j] = 0, \quad 1 \leq i, j, \leq \ell \tag{8.1}$$
$$[e_i, f_i] = h_i \tag{8.2}$$
$$[e_i, f_j] = 0 \quad i \neq j \tag{8.3}$$
$$[h_i, e_j] = \langle \alpha_j, \alpha_i \rangle e_j \tag{8.4}$$
$$[h_i, f_j] = -\langle \alpha_j, \alpha_i \rangle f_i \tag{8.5}$$
$$(\mathrm{ad}\ e_i)^{-\langle \alpha_j, \alpha_i \rangle + 1} e_j = 0 \quad i \neq j \tag{8.6}$$
$$(\mathrm{ad}\ f_i)^{-\langle \alpha_j, \alpha_i \rangle + 1} f_j = 0 \quad i \neq j. \tag{8.7}$$

Serre's theorem says that this is a presentation of a (semi-)simple Lie algebra. In particular, the Cartan matrix gives a presentation of a simple Lie algebra, showing that for every Dynkin diagram there exists a unique simple Lie algebra.

8.2 The first five relations.

Let $\mathbf{f}$ be the free Lie algebra on 3ℓ generators, $X_1, \ldots, X_\ell, Y_1, \ldots, Y_\ell, Z_1, \ldots, Z_\ell$. If $\mathbf{g}$ is a semi-simple Lie algebra with generators and relations (8.1)–(8.7), we have a unique homomorphism $\mathbf{f} \to \mathbf{g}$ where $X_i \to e_i,\ Y_i \to f_i, Z_i \to h_i$. We want to consider an intermediate algebra, $\mathbf{m}$, where we make use of all but the last two sets of relations. So let $\mathbf{I}$ be the ideal in $\mathbf{f}$ generated by the elements

$$[Z_i, Z_j],\ [X_i, Y_j] - \delta_{ij} Z_i,\ [Z_i, X_j] - \langle \alpha_j, \alpha_i \rangle X_j,\ [Z_i, Y_j] + \langle \alpha_j, \alpha_i \rangle Y_j.$$

We let $\mathbf{m} := \mathbf{f}/\mathbf{I}$ and denote the image of X_i in $\mathbf{m}$ by x_i etc.

We will first exhibit $\mathbf{m}$ as Lie subalgebra of the algebra of endomorphisms of a vector space. This will allow us to conclude that the x_i, y_j and z_k are linearly independent and from this deduce the structure of $\mathbf{m}$. We will then find that there is a homomorphism of $\mathbf{m}$ onto our desired semi-simple Lie algebra sending $x \mapsto e,\ y \mapsto f,\ z \mapsto h$.

So consider a vector space with basis $v_1, \ldots, v_\ell$ and let A be the tensor algebra over this vector space. We drop the tensor product signs in the algebra, so write

$$v_{i_1} v_{i_2} \cdots v_{i_t} := v_{i_1} \otimes \cdots v_{i_t}$$

for any finite sequence of integers with values from 1 to ℓ. We make A into an **f** module as follows: We let the Z_i act as derivations of A, determined by its actions on generators by

$$Z_i 1 = 0, \quad Z_j v_i = -\langle \alpha_i, \alpha_j \rangle v_j.$$

So if we define

$$c_{ij} := \langle \alpha_i, \alpha_j \rangle$$

we have

$$Z_j(v_{i_1} \cdots v_{i_t}) = -(c_{i_1 j} + \cdots + c_{i_t j})(v_{i_1} \cdots v_{i_t}).$$

The action of the Z_i is diagonal in this basis, so their actions commute. We let the Y_i act by left multiplication by v_i. So

$$Y_j v_{i_1} \cdots v_{i_t} := v_j v_{i_1} \cdots v_{i_t}$$

and hence

$$[Z_i, Y_j] = -c_{ji} Y_j = -\langle \alpha_j, \alpha_i \rangle Y_j$$

as desired. We now want to define the action of the X_i so that the relations analogous to (8.2) and (8.3) hold. Since $Z_i 1 = 0$ these relations will hold when applied to the element 1 if we set

$$X_j 1 = 0 \quad \forall j$$

and

$$X_j v_i = 0 \quad \forall i, j.$$

Suppose we define

$$X_j(v_p v_q) = -\delta_{jp} c_{qj} v_q.$$

Then

$$Z_i X_j(v_p v_q) = \delta_{jp} c_{qj} c_{qi} v_q = -c_{qi} X_j(v_p v_q)$$

while

$$X_j Z_i(v_p v_q) = \delta_{jp} c_{qj}(c_{pi} + c_{qi}) v_q = -(c_{pi} + c_{qi}) X_j(v_j v_q).$$

Thus

$$[Z_i, X_j](v_p v_q) = c_{ji} X_j(v_p v_q)$$

as desired.

In general, define

$$X_j(v_{p_1} \cdots v_{p_t}) := v_{p_1}\left(X_j(v_{p_2} \cdots v_{p_t})\right) - \delta_{p_1 j}(c_{p_2 j} + \cdots + c_{p_t j})(v_{p_2} \cdots v_{p_t}) \quad (8.8)$$

for $t \geq 2$. We claim that

$$Z_i X_j\left(v_{p_1} \cdots v_{p_t}\right) = -(c_{p_1 i} + \cdots + c_{p_t i} - c_{ji}) X_j(v_{p_1} \cdots v_{p_t}).$$

Indeed, we have verified this for the case $t = 2$. By induction, we may assume that $X_j(v_{p_2} \cdots v_{p_t})$ is an eigenvector of Z_i with eigenvalue $c_{p_2 i} + \cdots + c_{p_t i} -$

c_{ji}. Multiplying this on the left by v_{p_1} produces the first term on the right of (8.8). On the other hand, this multiplication produces an eigenvector of Z_i with eigenvalue $c_{p_1 i} + \cdots + c_{p_t i} - c_{ji}$. As for the second term on the right of (8.8), if $j \neq p_1$ it does not appear. If $j = p_1$ then $c_{p_1 i} + \cdots + c_{p_t i} - c_{ji} = c_{p_2 i} + \cdots + c_{p_t i}$. So in either case, the right hand side of (8.8) is an eigenvector of Z_i with eigenvalue $c_{p_1 i} + \cdots + c_{p_t i} - c_{ji}$. But then

$$[Z_i, X_j] = \langle \alpha_j, \alpha_i \rangle X_j$$

as desired. We have defined an action of **f** on A whose kernel contains **I**, hence descends to an action of **m** on A.

Let $\phi : \mathbf{m} \to \text{End } A$ denote this action. Suppose that $z := a_1 z_1 + \cdots + a_\ell z_\ell$ for some complex numbers $a_1, \ldots, a_\ell$ and that $\phi(z) = 0$. The operator $\phi(z)$ has eigenvalues

$$-\sum a_j c_{ij}$$

when acting on the subspace V of A. All of these must be zero. But the Cartan matrix is non-singular. Hence all the $a_i = 0$. This shows that the space spanned by the z_i is in fact ℓ-dimensional and spans an ℓ-dimensional abelian subalgebra of **m**. Call this subalgebra **z**.

Now consider the 3ℓ-dimensional subspace of **f** spanned by the X_i, Y_i and Z_i, $i = 1, \ldots, \ell$. We wish to show that it projects onto a 3ℓ dimensional subspace of **m** under the natural passage to the quotient $\mathbf{f} \to \mathbf{m} = \mathbf{f}/\mathbf{i}$. The image of this subspace is spanned by x_i, y_i and z_i. Since $\phi(x_i) \neq 0$ and $\phi(y_i) \neq 0$ we know that $x_i \neq 0$ and $y_i \neq 0$. Suppose we had a linear relation of the form

$$\sum a_i x_i + \sum b_i y_i + z = 0.$$

Choose some $z' \in \mathbf{z}$ such that $\alpha_i(z') \neq 0$ and $\alpha_i(z') \neq \alpha_j(z')$ for any $i \neq j$. This is possible since the α_i are all linearly independent. Bracketing the above equation by z' gives

$$\sum \alpha_(z') a_i x_i - \sum \alpha_i(z') b_i y_i = 0$$

by the relations (8.4) and (8.5). Repeated bracketing by z' and using the van der Monde (or induction) argument shows that $a_i = 0, b_i = 0$ and hence that $z = 0$.

We have proved that the elements x_i, y_j, z_k in **m** are linearly independent.

The element

$$[x_{i_1}, [x_{i_2}, [\cdots [x_{i_{t-1}}, x_{i_t}] \cdots]]]$$

is an eigenvector of z_i with eigenvalue

$$c_{i_1 i} + \cdots + c_{i_t i}.$$

For any pair of elements μ and λ of $\mathbf{z}^*$ (or of $\mathbf{h}^*$) recall that

$$\mu \prec \lambda$$

denotes the fact that $\lambda - \mu = \sum k_i \alpha_i$ where the k_i are all non-negative integers.

For any $\lambda \in \mathbf{z}^*$ let $\mathbf{m}_\lambda$ denote the set of all $m \in \mathbf{m}$ satisfying

$$[z, m] = \lambda(z)m \quad \forall z \in \mathbf{z}.$$

Then we have shown that the subalgebra $\mathbf{x}$ of $\mathbf{m}$ generated by $x_1, \ldots, x_\ell$ is contained in

$$\mathbf{m}_+ := \bigoplus_{0 \prec \lambda} \mathbf{m}_\lambda.$$

Similarly, the subalgebra $\mathbf{y}$ of $\mathbf{m}$ generated by the y_i lies in

$$\mathbf{m}_- := \bigoplus_{\lambda \prec 0} \mathbf{m}_\lambda.$$

In particular, the vector space sum

$$\mathbf{y} + \mathbf{z} + \mathbf{x}$$

is direct since $\mathbf{z} \subset \mathbf{m}_0$. We claim that this is in fact all of $\mathbf{m}$. First of all, observe that it is a subalgebra. Indeed, $[y_i, x_j] = -\delta_{ij} z_i$ lies in this subspace, and hence

$$[y_i, [x_{j_1}, [\cdots [x_{j_{t-1}}, x_{j_t}] \cdots] \in \mathbf{x} \quad \text{for } t \geq 2.$$

Thus the subspace $\mathbf{y} + \mathbf{z} + \mathbf{x}$ is closed under ad y_i and hence under any product of these operators. Similarly for ad x_i. Since these generate the algebra $\mathbf{m}$ we see that $\mathbf{y} + \mathbf{z} + \mathbf{x} = \mathbf{m}$ and hence

$$\mathbf{x} = \mathbf{m}_+ \quad \text{and} \quad \mathbf{y} = \mathbf{m}_-.$$

We have shown that

$$\mathbf{m} = \mathbf{m}_- \oplus \mathbf{z} \oplus \mathbf{m}_+$$

where $\mathbf{z}$ is an abelian subalgebra of dimension ℓ, where the subalgebra $\mathbf{m}_+$ is generated by $x_1, \ldots, x_\ell$, where the subalgebra $\mathbf{m}_-$ is generated by $y_1, \ldots, y_\ell$, and where the 3ℓ elements $x_1, \ldots, x_\ell, y_1 \ldots, y_\ell, z_1, \ldots, z_\ell$ are linearly independent.

There is a further property of $\mathbf{m}$ which we want to use in the next section in the proof of Serre's theorem. For all $i \neq j$ between 1 and ℓ define the elements x_{ij} and y_{ij} by

$$x_{ij} := (\text{ad } x_i)^{-c_{ji}+1}(x_j), \quad y_{ij} := (\text{ad } y_i)^{-c_{ji}+1}(y_j).$$

Conditions (8.6) and (8.7) amount to setting these elements, and hence the ideal that they generate equal to zero. We claim that for all k and all $i \neq j$ between 1 and ℓ we have

$$\text{ad } x_k(y_{ij}) = 0 \tag{8.9}$$

and

$$\text{ad } y_k(x_{ij}) = 0. \tag{8.10}$$

By symmetry, it is enough to prove the first of these equations. If $k \neq i$ then $[x_k, y_i] = 0$ by (8.3) and hence

$$\text{ad } x_k(y_{ij}) = (\text{ad } y_i)^{-c_{ji}+1}[x_k, y_j] = (\text{ad } y_i)^{-c_{ji}+1}\delta_{kj}h_j$$

by (8.2) and (8.3). If $k \neq j$ this is zero. If $k = j$ we can write this as

$$(\text{ad } y_i)^{-c_{ji}}(\text{ad } y_i)h_j = (\text{ad } y_i)^{-c_{ji}}c_{ij}y_i.$$

If $c_{ij} = 0$ there is nothing to prove. If $c_{ij} \neq 0$ then $c_{ji} \neq 0$ and in fact is strictly negative since the angles between all elements of a base are obtuse. But then

$$(\text{ad } y_i)^{-c_{ji}}y_i = 0.$$

It remains to consider the case where $k = i$. The algebra generated by $x_i, y_, z_i$ is isomorphic to $sl(2)$ with $[x_i, y_i] = z_i, [z_i, x_i] = 2x_i, [z_i, y_i] = -2y_i$. We have a decomposition of $\mathbf{m}$ into weight spaces for all of $\mathbf{z}$, in particular into weight spaces for this little $sl(2)$. Now $[x_i, y_j] = 0$ (from (8.3)) so y_j is a maximal weight vector for this $sl(2)$ with weight $-c_{ji}$ and (8.9) is just a standard property of a maximal weight module for $sl(2)$ with non-negative integer maximal weight.

8.3 Proof of Serre's theorem.

Let $\mathbf{k}$ be the ideal of $\mathbf{m}$ generated by the x_{ij} and y_{ij} as defined above. We wish to show that

$$\mathbf{g} := \mathbf{m}/\mathbf{k}$$

is a semi-simple Lie algebra with Cartan subalgebra $\mathbf{h} = \mathbf{z}/\mathbf{k}$ and root system Φ. For this purpose, let $\mathbf{i}$ now denote the ideal in $\mathbf{m}_+$ generated by the x_{ij} and $\mathbf{j}$ be the ideal in $\mathbf{m}_-$ generated by the y_{ij} so that

$$\mathbf{i} + \mathbf{j} \subset \mathbf{k}.$$

We claim that $\mathbf{j}$ is an ideal of $\mathbf{m}$. Indeed, each y_{ij} is a weight vector for $\mathbf{z}$, and $[\mathbf{z}, \mathbf{m}_-] \subset \mathbf{m}_-$, hence $[\mathbf{z}, \mathbf{j}] \subset \mathbf{j}$. On the other hand, we know that $[x_k, \mathbf{m}_-] \subset \mathbf{m}_- + \mathbf{z}$ and $[x_k, y_{ij}] = 0$ by (8.9). So $(\text{ad } x_k)\mathbf{j} \subset \mathbf{j}$ by Jacobi. Since the x_k generate $\mathbf{m}_+$ Jacobi then implies that $[\mathbf{m}_+, \mathbf{j}] \subset \mathbf{j}$ as well, hence $\mathbf{j}$ is an ideal of $\mathbf{m}$. Similarly, $\mathbf{i}$ is an ideal of $\mathbf{m}$. Hence $\mathbf{i} + \mathbf{j}$ is an ideal of $\mathbf{m}$, and since it contains the generators of $\mathbf{k}$, it must coincide with $\mathbf{k}$, i.e.

$$\mathbf{k} = \mathbf{i} + \mathbf{j}.$$

In particular, $\mathbf{z} \cap \mathbf{k} = \{0\}$ and so $\mathbf{z}$ projects isomorphically onto an ℓ-dimensional abelian subalgebra of $\mathbf{g} = \mathbf{m}/\mathbf{k}$. Furthermore, since $\mathbf{j} \cap \mathbf{m}_+ = \{0\}$ and $\mathbf{i} \cap \mathbf{m}_- = \{0\}$ we have

$$\mathbf{g} = \mathbf{n}_- \oplus \mathbf{h} \oplus \mathbf{n}_+ \tag{8.11}$$

as a vector space where

$$\mathbf{n}_- = \mathbf{m}_-/\mathbf{j}, \quad \text{and} \quad \mathbf{n}_+ = \mathbf{m}_+/\mathbf{i},$$

and $\mathbf{n}_+$ is a a sum of weight spaces of $\mathbf{h}$, summed over $\lambda \succ 0$ while $\mathbf{n}_-$ is a sum of weight spaces of $\mathbf{h}$ with $\lambda \prec 0$. We have to see which weight spaces survive the passage to the quotient. The $sl(2)$ generated by x_i, y_i, z_i is not sent into zero by the projection of $\mathbf{m}$ onto $\mathbf{g}$ since z_i is not sent into zero. Since $sl(2)$ is simple, this means that the projection map is an isomorphism when restricted to this $sl(2)$. Let us denote the images of x_i, y_i, z_i by e_i, f_i, h_i. Thus $\mathbf{g}$ is generated by the 3ℓ elements

$$e_1, \dots, e_\ell, f_1, \dots, f_\ell, h_1, \dots, h_\ell$$

and all the axioms (8.1)-(8.7) are satisfied.

We must show that $\mathbf{g}$ is finite dimensional, semi-simple, and has Φ as its root system.

First observe that ad e_i acts nilpotently on each of the generators of the algebra $\mathbf{g}$, and hence acts locally nilpotently on all of $\mathbf{g}$. Similarly for ad f_i. Hence the automorphism

$$\tau_i := (\exp \mathrm{ad}\ e_i)(\mathrm{ad}\ -f_i)(\exp \mathrm{ad}\ e_i)$$

is well defined on all of $\mathbf{g}$. So if s_i denotes the reflection in the Weyl group W corresponding to i, we have

$$\tau_i(\mathbf{g}_\lambda) = \mathbf{g}_{s_i \lambda}.$$

Notice that each of the $\mathbf{m}_\lambda$ is finite dimensional, since the dimension of $\mathbf{m}_\lambda$ for $\lambda \succ 0$ is at most the number of ways to write λ as a sum of successive α_i, each such sum corresponding to the element $[x_{i_1}, [x_{i_2}, [\cdots, x_{i_t}]\cdots]$. (In particular $\mathbf{m}_{k\alpha} = \{0\}$ for $k > 1$.) Similarly for $\lambda \prec 0$. So it follows that each of the $\mathbf{g}_\lambda$ is finite dimensional, that

$$\dim \mathbf{g}_{w\lambda} = \dim \mathbf{g}_\lambda \quad \forall w \in W$$

and that

$$\mathbf{g}_{k\lambda} = 0 \quad \text{for } k \neq -1, 0, 1.$$

Furthermore, $\mathbf{g}_{\alpha_i}$ is one dimensional, and since every root is conjugate to a simple root, we conclude that

$$\dim \mathbf{g}_\alpha = 1 \quad \forall \alpha \in \Phi.$$

We now show that

$$\mathbf{g}_\lambda = \{0\} \ \text{ for } \lambda \neq 0, \quad \lambda \notin \Phi.$$

Indeed, suppose that $\mathbf{g}_\lambda \neq \{0\}$. We know that λ is not a multiple of α for any $\alpha \in \Phi$, since we know this to be true for simple roots, and the dimensions of the $\mathbf{g}_\lambda$ are invariant under the Weyl group, each root being conjugate to a simple root. So $\lambda^\perp$ does not coincide with any hyperplane orthogonal to any root. So we can find a $\mu \in \lambda^\perp$ such that $(\alpha, \mu) \neq 0$ for all roots. We may find a $w \in W$ which maps μ into the positive Weyl chamber for Δ so that $(\alpha_i, \mu) \geq 0$ and hence $(\alpha_i, w\mu) > 0$ for $i = 1, \dots, \ell$. Now

$$\dim \mathbf{g}_{w\lambda} = \dim \mathbf{g}_\lambda$$

and for the latter to be non-zero, we must have

$$w\lambda = \sum k_i \alpha_i$$

with the coefficients all non-negative or non-positive integers. But

$$0 = (\lambda, \mu) = (w\lambda, w\mu) = \sum k_i (\alpha_i, \mu)$$

with $(\alpha_i, \mu) > 0 \ \forall i$. Hence all the $k_i = 0$.

So

$$\dim \mathbf{g} = \ell + \text{ Card } \Phi.$$

We conclude the proof if we show that $\mathbf{g}$ is semi-simple, i.e. contains no abelian ideals. So suppose that $\mathbf{a}$ is an abelian ideal. Since $\mathbf{a}$ is an ideal, it is stable under $\mathbf{h}$ and hence decomposes into weight spaces. If $\mathbf{g}_\alpha \cap \mathbf{a} \neq \{0\}$, then $\mathbf{g}_\alpha \subset \mathbf{a}$ and hence $[\mathbf{g}_{-\alpha}, \mathbf{g}_\alpha] \subset \mathbf{a}$ and hence the entire $sl(2)$ generated by $\mathbf{g}_\alpha$ and $\mathbf{g}_{-\alpha}$ is contained in $\mathbf{a}$ which is impossible since $\mathbf{a}$ is abelian and $sl(2)$ is simple. So $\mathbf{a} \subset \mathbf{h}$. But then $\mathbf{a}$ must be annihilated by all the roots, which implies that $\mathbf{a} = \{0\}$ since the roots span $\mathbf{h}^*$. QED

8.4 The existence of the exceptional root systems.

The idea of the construction is as follows. For each Dynkin diagram we will chose a lattice $\mathbb{L}$ in a Euclidean space V, and then let Φ consist of *all* vectors in this lattice having all the same length, or having one of two prescribed lengths. We then check that

$$\frac{2(\alpha_1, \alpha_2)}{(\alpha_1, \alpha_1)} \in \mathbb{Z} \qquad \forall \ \alpha_1, \ \alpha_2 \in \Phi.$$

This implies that reflection through the hyperplane orthogonal to α_1 preserves $\mathbb{L}$, and since reflections preserve length that these reflection s preserve Φ. This will show that Φ is a root system and then calculation shows that it is of the desired type.

(G_2). Let V be the plane in $\mathbb{R}^3$ consisting of all vectors $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ with

$$x + y + z = 0.$$

Let $\mathbb{L}$ be the intersection of the three dimensional standard lattice $\mathbb{Z}^2$ with V. Let L_1, L_2, L_3 denote the standard basis of $\mathbb{R}^3$. Let Φ consist of all vectors in $\mathbb{L}$ of squared length 2 or 6. So Φ consists of the six short vectors

$$\pm(L_i - L_j) \quad i < j$$

and the six long vectors

$$\pm(2L_i - L_j - L_k) \quad i = 1,2,3 \; j \neq i, \; j \neq k.$$

We may choose the base to consist of

$$\alpha_= L_1 - L_2, \quad \alpha_2 = -2L_1 + L_2 + L_3.$$

F_4. Let $V = \mathbb{R}^4$ and $\mathbb{L} = \mathbb{Z}^4 + \mathbb{Z}(\frac{1}{2}(L_1 + L_2 + L_3 + L_4)$. Let Φ consist of all vectors of $\mathbb{L}$ of squared length 1 or 2. So Φ consists of the 24 long roots

$$\pm L_i \pm L_j \quad i < j$$

and the 24 short roots

$$\pm L_1, \quad \frac{1}{2}(\pm L_1 \pm L_2 \pm L_3 \pm L_4).$$

For Δ we may take

$$\alpha_1 = L_2 - L_3, \quad \alpha_2 = L_3 - L_4, \quad \alpha_3 = L_4, \quad \alpha_4 = \frac{1}{2}(L_1 - L_2 - L_3 - L_4).$$

E_8. Let $V = \mathbb{R}^8$. Let

$$\mathbb{L}' := \ \{\sum c_i L_i, \quad c_i \in \mathbb{Z}, \quad \sum c_i \ \text{ even } \}.$$

Let

$$\mathbb{L} := \mathbb{L}' + \mathbb{Z}(\frac{1}{2}(L_1 + L_1 + L_2 + L_3 + L_4 + L_5 + L_6 + L_7 + L_8)).$$

Let Φ consist of all vectors in $\mathbb{L}$ of squared length 2. So Φ consists of the 240 roots

$$\pm L_i \pm L_j \ \ (i < j), \qquad \frac{1}{2}\sum_{i=1}^{8} \pm L_i, \ \text{(even number of + signs)}.$$

For Δ we may take

$$\begin{aligned} \alpha_1 &= \frac{1}{2}(L_1 - L_1 - L_2 - L_3 - L_4 - L_5 - L_6 - L_7 + L_8) \\ \alpha_2 &= L_1 + L_2 \\ \alpha_i &= L_{i-1} - L_{i-2} \ \ (3 \leq i \leq 8). \end{aligned}$$

E_7 is obtained from E_8 by letting V be the span of the first 7 α_i. E_6 is obtained from E_8 by letting V be the span of the first 6 α_i.

Chapter 9

Clifford algebras and spin representations.

9.1 Definition and basic properties

9.1.1 Definition.

Let $\mathbf{p}$ be a vector space with a symmetric bilinear form (,). The **Clifford algebra** associated to this data is the algebra

$$C(\mathbf{p}) := T(\mathbf{p})/I$$

where $T(\mathbf{p})$ denotes the tensor algebra

$$T(\mathbf{p}) = k \oplus \mathbf{p} \oplus (\mathbf{p} \otimes \mathbf{p}) \oplus \cdots$$

and where I denotes the ideal in $T(\mathbf{p})$ generated by all elements of the form

$$y_1 y_2 + y_2 y_1 - 2(y_1, y_2)\mathbf{1}, \quad y_1,\ y_2\ \in \mathbf{p}$$

and $\mathbf{1}$ is the unit element of the tensor algebra. The space $\mathbf{p}$ injects as a subspace of $C(\mathbf{p})$ and generates $C(\mathbf{p})$ as an algebra.

A linear map f of $\mathbf{p}$ to an associative algebra A with unit 1_A is called a **Clifford map** if

$$f(y_1)f(y_2) + f(y_2)f(y_1) = 2(y_1, y_2)1_A, \quad \forall y_1,\ y_2\ \in \mathbf{p}$$

or what amounts to the same thing (by polarization since we are not over a field of characteristic 2) if

$$f(y)^2 = (y, y)1_A \ \ \forall\ y \in \mathbf{p}.$$

Any Clifford map gives rise to a unique algebra homomorphism of $C(\mathbf{p})$ to A whose restriction to $\mathbf{p}$ is f. The Clifford algebra is "universal" with respect to this property.

If the bilinear form is identically zero, then $C(\mathbf{p}) = \wedge\mathbf{p}$, the exterior algebra. But we will be interested in the opposite extreme, the case where the bilinear form is non-degenerate.

9.1.2 Gradation.

The ideal I defining the Clifford algebra is not $\mathbf{Z}$ homogeneous (unless the bilinear form is identically zero) since its generators $y_1y_2 + y_2y_1 - 2(y_1, y_2)\mathbf{1}$ are "mixed", being a sum of terms of degree two and degree zero in $T(\mathbf{p})$. But these terms are both even. So the $\mathbf{Z}/2\mathbf{Z}$ gradation *is* preserved upon passing to the quotient. In other words, $C(\mathbf{p})$ is a $\mathbf{Z}/2\mathbf{Z}$ graded algebra:

$$C(\mathbf{p}) = C_{\mathbf{0}}(\mathbf{p}) \oplus C_{\mathbf{1}}(\mathbf{p})$$

where the elements of $C_{\mathbf{0}}(\mathbf{p})$ consist of sums of products of elements of $\mathbf{p}$ with an even number of factors and $C_{\mathbf{1}}(\mathbf{p})$ consist of sums of terms each a product of elements of $\mathbf{p}$ with an odd number of factors. The usual rules for multiplication of a graded algebra obtain:

$$C_{\mathbf{0}}(\mathbf{p}) \cdot C_{\mathbf{0}}(\mathbf{p}) \subset C_{\mathbf{0}}(\mathbf{p}),\ C_{\mathbf{0}}(\mathbf{p}) \cdot C_{\mathbf{1}}(\mathbf{p}) \subset C_{\mathbf{1}}(\mathbf{p}),\ C_{\mathbf{1}}(\mathbf{p}) \cdot C_{\mathbf{1}}(\mathbf{p}) \subset C_{\mathbf{0}}(\mathbf{p}).$$

9.1.3 $\wedge\mathbf{p}$ as a $C(\mathbf{p})$ module.

Let $\mathbf{p}$ be a vector space with a non-degenerate symmetric bilinear form. The exterior algebra, $\wedge\mathbf{p}$ inherits a bilinear form which we continue to denote by $(\ ,\ )$. Here the spaces $\wedge^k(\mathbf{p})$ and $\wedge^\ell(\mathbf{p})$ are orthogonal if $k \neq \ell$ while

$$(x_1 \wedge \cdots \wedge x_k, y_1 \wedge \cdots \wedge y_k) = \det\left((x_i, y_j)\right).$$

For $v \in \mathbf{p}$ let $\epsilon(v) \in \mathrm{End}(\wedge\mathbf{p})$ denote exterior multiplication by v and $\iota(v)$ be the transpose of $\epsilon(v)$ relative to this biinear form on $\wedge\mathbf{p}$.

So $\iota(v)$ is interior multiplication by the element of $\mathbf{p}^*$ corresponding to v under the map $\mathbf{p} \to \mathbf{p}^*$ induced by $(\ ,\ )_{\mathbf{p}}$. The map

$$\mathbf{p} \to\ \mathrm{End}(\wedge\mathbf{p}),\quad v \mapsto \epsilon(v) + \iota(v)$$

is a Clifford map, i.e. satisfies

$$(\epsilon(v) + \iota(v))^2 = (v, v)_{\mathbf{p}}\mathrm{id}$$

and so extends to a homomorphism of

$$C(\mathbf{p}) \to \mathrm{End}\wedge\mathbf{p}$$

making $\wedge\mathbf{p}$ into a $C(\mathbf{p})$ module. We let xy denote the product of x and y in $C(\mathbf{p})$.

9.1.4 Chevalley's linear identification of $C(\mathbf{p})$ with $\wedge\mathbf{p}$.

Consider the linear map

$$C(\mathbf{p}) \to \wedge\mathbf{p},\qquad x \mapsto x1$$

where $1 \in \wedge^0 \mathbf{p}$ under the identification of $\wedge^0 \mathbf{p}$ with the ground field. The element $x1$ on the extreme right means the image of 1 under the action of $x \in C(\mathbf{p})$.

For elements $v_1, \ldots, v_k \in \mathbf{p}$ this map sends

$$\begin{array}{rcl} v_1 & \mapsto & v_1 \\ v_1 v_2 & \mapsto & v_1 \wedge v_2 + (v_1, v_2)1 \\ v_1 v_2 v_3 & \mapsto & v_1 \wedge v_2 \wedge v_3 + (v_1, v_2)v_3 - (v_1, v_3)v_2 + (v_2, v_3)v_1 \\ v_1 v_2 v_3 v_4 & \mapsto & v_1 \wedge v_2 \wedge v_3 \wedge v_4 + (v_2, v_3)v_1 \wedge v_4 - (v_2, v_4)v_1 \wedge v_3 \\ & & +(v_3, v_4)v_1 \wedge v_2 + (v_1, v_2)v_3 \wedge v_4 - (v_1, v_3)v_1 \wedge v_4 \\ & & +(v_1, v_4)v_2 \wedge v_3 + (v_1, v_4)(v_2, v_3) - (v_1, v_3)(v_2, v_4) + (v_1, v_2)(v_3, v_4) \\ \vdots & & \vdots \end{array}$$

If the v's form an "orthonormal" basis of $\mathbf{p}$ then the products

$$v_{i_1} \cdots v_{i_k}, \quad i_1 < i_2 \cdots < i_k,\ k = 0, 1, \ldots, n$$

form a basis of $C(\mathbf{p})$ while the

$$v_{i_1} \wedge \cdots \wedge v_{i_k}, \quad i_1 < i_2 \cdots < i_k,\ k = 0, 1, \ldots, n$$

form a basis of $\wedge \mathbf{p}$, and in fact

$$v_1 \cdots v_k \mapsto v_1 \wedge \cdots \wedge v_k \quad \text{if } (v_i, v_j) = 0\ \forall i \neq j. \tag{9.1}$$

In particular, the map $C(\mathbf{p}) \to \wedge \mathbf{p}$ given above is an isomorphism of vector spaces, so we may identify $C(\mathbf{p})$ with $\wedge \mathbf{p}$ as a vector space if we choose, and then consider that $\wedge \mathbf{p}$ has two products: the Clifford product which we denote by juxtaposition and the exterior product which we denote with a $\wedge$.

Notice that this identification preserves the $\mathbf{Z}/2\mathbf{Z}$ gradation, an even element of the Clifford algebra is identified with an even element of the exterior algebra and an odd element is identified with an odd element.

9.1.5 The canonical antiautomorphism.

The Clifford algebra has a canonical anti-automorphism a which is the identity map on $\mathbf{p}$. In particular, for $v_i \in \mathbf{p}$ we have $a(v_1 v_2) = v_2 v_1$, $a(v_1 v_2 v_3) = v_3 v_2 v_1$, etc. By abuse of language, we use the same letter a to denote the similar anti-automorphism on $\wedge \mathbf{p}$ and observe from the above computations (in particular from the corresponding choice of bases) that a commutes with our identifying map $C(\mathbf{p}) \to \wedge \mathbf{p}$ so the notation is consistent. We have

$$a = (-1)^{\frac{1}{2}k(k-1)} \mathrm{id} \quad \text{on } \wedge^k (\mathbf{p}).$$

For small values of k we have

k	$(-1)^{\frac{1}{2}k(k-1)}$
0	1
1	1
2	-1
3	-1
4	1
5	1
6	-1.

We will use subscripts to denote the homogeneous components of elements of $\wedge\mathbf{p}$. Notice that if $u \in \wedge^2\mathbf{p}$ then $au = -u$ by the above table, while $a(u^2) = (au)^2 = u^2$. Since u^2 is even (and hence has only even homogeneous components) and since the maximum degree of the homogeneous component of u^2 is 4, we conclude that

$$u^2 = (u^2)_0 + (u^2)_4 \quad \forall\, u \in \wedge^2\mathbf{p}. \tag{9.2}$$

For the same reason

$$v^2 = (v^2)_0 + (v^2)_4 \quad \forall\, v \in \wedge^3\mathbf{p}. \tag{9.3}$$

We also claim the following:

$$(ww')_0 = (aw, w') = (-1)^{\frac{1}{2}k(k-1)}(w, w') \quad \forall\, w, w' \in \wedge^k(\mathbf{p}). \tag{9.4}$$

Indeed, it is sufficient to verify this for w, w' belonging to a basis of $\wedge\mathbf{p}$, say the basis given by all elements of the form (9.1), in which case both sides of (9.4) vanish unless $w = w'$. If $w = w' = v_1 \wedge \cdots \wedge v_k$ (say) then

$$(ww)_0 = \iota(v_1)\cdots\iota(v_k)v_1 \wedge \cdots \wedge v_k =$$

$$(-1)^{\frac{1}{2}k(k-1)}(v_1, v_1)\cdots(v_k, v_k) = (-1)^{\frac{1}{2}k(k-1)}(w, w)$$

proving (9.4).

As special cases that we will use later on, observe that

$$(uu')_0 = -(u, u') \;\; \forall\, u, u' \in \wedge^2\mathbf{p} \tag{9.5}$$

and

$$(vv')_0 = -(v, v') \;\; \forall\, v, v' \in \wedge^3\mathbf{p}. \tag{9.6}$$

9.1.6 Commutator by an element of p.

For any $y \in \mathbf{p}$ consider the linear map

$$w \mapsto [y, w] = yw - (-1)^k wy \;\; \text{for } w \in \wedge^k\mathbf{p}$$

which is (anti)commutator in the Clifford multiplication by y. We claim that

$$[y, w] = 2\iota(y)w. \tag{9.7}$$

In particular, $[y, \cdot]$, which is automatically a derivation for the Clifford multiplication, is also a derivation for the exterior multiplication. Alternatively, this equation says that $\iota(y)$, which is a derivation for the exterior algebra multiplication, is also a derivation for the Clifford multiplication.

To prove (9.7) write

$$wy = a(ya(w)).$$

Then

$$yw = y \wedge w + \iota(y)w, \quad wy = a(y \wedge a(w)) + a(\iota(y)aw) = w \wedge y + (a\iota(y)a)w.$$

We may assume that $w \in \wedge^k \mathbf{p}$. Then

$$y \wedge w - (-1)^k w \wedge y = 0,$$

so we must show that

$$a\iota(y)aw = (-1)^{k-1}\iota(y)w.$$

For this we may assume that $y \neq 0$ and we may write

$$w = u \wedge z + z',$$

where $\iota(y)u = 1$ and $\iota(y)z = \iota(y)z' = 0$. In fact, we may assume that z and z' are sums of products of linear elements all of which are orthogonal to y. Then $\iota(y)az = \iota(y)az' = 0$ so

$$\iota(y)aw = (-1)^{k-1}az$$

since z has degree one less than w and hence

$$a\iota(y)aw = (-1)^{k-1}z = (-1)^{k-1}\iota(y)w. \quad QED$$

9.1.7 Commutator by an element of $\wedge^2\mathbf{p}$.

Suppose that

$$u \in \wedge^2 \mathbf{p}.$$

Then for $y \in \mathbf{p}$ we have

$$[u, y] = -[y, u] = -2\iota(y)u. \tag{9.8}$$

In particular, if $u = y_i \wedge y_j$ where $y_i, y_j \in \mathbf{p}$ we have

$$[u, y] = 2(y_j, y)y_i - 2(y_i, y)y_j \quad \forall\ y \in \mathbf{p}. \tag{9.9}$$

If $(y_i, y_j) = 0$ this is an “infinitesimal rotation” in the plane spanned by y_i and y_j. Since $y_i \wedge y_j,\ i < j$ form a basis of $\wedge^2\mathbf{p}$ if $y_1, \ldots, y_n$ form an “orthonormal” basis of $\mathbf{p}$, we see that the map

$$u \mapsto [u, \cdot]$$

gives an isomorphism of $\wedge^2\mathbf{p}$ with the orthogonal algebra $o(\mathbf{p})$. This identification differs by a factor of two from the identification that we had been using earlier.

Now each element of $o(\mathbf{p})$ (in fact any linear transformation on $\mathbf{p}$) induces a derivation of $\wedge\mathbf{p}$. We claim that under the above identification of $\wedge^2\mathbf{p}$ with $o(\mathbf{p})$, the derivation corresponding to $u \in \wedge^2\mathbf{p}$ is Clifford commutation by u. In symbols, if θ_u denotes this induced derivation, we claim that

$$\theta_u(w) = [u, w] = uw - wu \quad \forall\ w \in \wedge\mathbf{p}. \tag{9.10}$$

To verify this, it is enough to check it on basis elements of the form (9.1), and hence by the derivation property for each v_j, where this reduces to (9.8).

We can now be more explicit about the degree four component of the Clifford square of an element of $\wedge^2\mathbf{p}$, i.e. the element $(u^2)_4$ occurring on the right of (9.2). We claim that for any three elements $y, y', y'' \ \in \mathbf{p}$

$$\frac{1}{2}\iota(y'')\iota(y')\iota(y)u^2 = (y\wedge y', u)\iota(y'')u+(y'\wedge y'', u)\iota(y)u+(y''\wedge y, u)\iota(y')u. \tag{9.11}$$

To prove this observe that

$$\begin{aligned}
\iota(y)u^2 &= (\iota(y)u)\,u + u\,(\iota(y)u) \\
\iota(y')\iota(y)u^2 &= (\iota(y')\iota(y)u)\,u - \iota(y)u\iota(y')u + \iota(y')u\iota(y)u + u\iota(y')\iota(y)u \\
&= 2\,((y\wedge y', u)u + \iota(y')u \wedge \iota(y)u) \\
\frac{1}{2}\iota(y'')\iota(y')\iota(y)u^2 &= (y\wedge y', u)\iota(y'')u + \iota(y'')\iota(y')u \wedge \iota(y)u - \iota(y')u \wedge \iota(y'')\iota(y)u \\
&= (y\wedge y', u)\iota(y'')u + (y'\wedge y'', u)\iota(y)u + (y''\wedge y, u)\iota(y')u
\end{aligned}$$

as required.

We can also be explicit about the degree zero component of u^2. Indeed, it follows from (9.9) that if $u = y_i \wedge y_j,\ i < j$ where $y_1, \ldots, y_n$ form an "orthonormal" basis of $\mathbf{p}$ then

$$\operatorname{tr}(\operatorname{ad}_{\mathbf{p}} u)^2 = -8(y_i, y_i)(y_j, y_j),$$

where $\operatorname{ad}_{\mathbf{p}} u$ denotes the (commutator) action of u on $\mathbf{p}$ under our identification of $\wedge^2\mathbf{p}$ with $o(\mathbf{p})$. But

$$(y_i \wedge y_j, y_i \wedge y_j) = (y_i, y_i)(y_j, y_j)(= \pm 1).$$

So using (9.5) we see that

$$(u^2)_0 = \frac{1}{8}\operatorname{tr}(\operatorname{ad}_{\mathbf{p}} u)^2 = -(u, u) \tag{9.12}$$

for $u \in \wedge^2\mathbf{p}$.

9.2 Orthogonal action of a Lie algebra.

Let $\mathbf{r}$ be a Lie algebra. Suppose that we have a representation of $\mathbf{r}$ acting as infinitesimal orthogonal transformations of $\mathbf{p}$ which means, in view of the identification of $\wedge^2\mathbf{p}$ with $o(\mathbf{p})$ that we have a map

$$\nu : \mathbf{r} \to \wedge^2\mathbf{p}$$

such that

$$x \cdot y = -2\iota(y)\nu(x) \tag{9.13}$$

where $x \cdot y$ denotes the action of $x \in \mathbf{r}$ on $y \in \mathbf{p}$.

9.2.1 Expression for ν in terms of dual bases.

It will be useful for us to write equation (9.13) in terms of a basis. So let $y_1, \dots, y_n$ be a basis of $\mathbf{p}$ and let $z_1, \dots, z_n$ be the dual basis relative to $(\ ,\ ) = (\ ,\ )_{\mathbf{p}}$. We claim that

$$\nu(x) = -\frac{1}{4}\sum_j y_j \wedge (x \cdot z_j). \tag{9.14}$$

Indeed, it suffices to verify (9.13) for each of the elements z_i. Now

$$\iota(z_i)\left(-\frac{1}{4}\sum_j y_j \wedge x \cdot z_j\right) = -\frac{1}{4}x \cdot z_i + \frac{1}{4}\sum_j (z_i, x \cdot z_j)y_j.$$

But

$$(z_i, x \cdot z_j) = -(x \cdot z_i, z_j)$$

since x acts as an infinitesimal orthogonal transformation relative to $(\ ,\ )$. So we can write the sum as

$$\frac{1}{4}\sum_j (z_i, x \cdot z_j)y_j = -\frac{1}{4}\sum_j (x \cdot z_i, z_j)y_j = -\frac{1}{4}x \cdot z_i$$

yielding

$$\iota(z_i)\left(-\frac{1}{4}\sum_j y_j \wedge x \cdot z_j\right) = -\frac{1}{2}x \cdot z_i$$

which is (9.13).

9.2.2 The adjoint action of a reductive Lie algebra.

For future use we record here a special case of (9.14): Suppose that $\mathbf{p} = \mathbf{r} = \mathbf{g}$ is a reductive Lie algebra with an invariant symmetric bilinear form, and the action is the adjoint action, i.e. $x \cdot y = [x, y]$. Let $\mathbf{h}$ be a Cartan subalgebra of $\mathbf{g}$

and let Φ denote the set of roots and suppose that we have chosen root vectors $e_\phi, e_{-\phi},\ \phi \in \Phi$ so that

$$(e_\phi, e_{-\phi}) = 1.$$

Let $h_1, \ldots, h_s$ be a basis of $\mathbf{h}$ and $k_1, \ldots k_s$ the dual basis. Let

$$\psi : \mathbf{g} \to \wedge^2 \mathbf{g}$$

be the map ν when applied to this adjoint action. Then (9.14) becomes

$$\psi(x) = \frac{1}{4}\left(\sum_{i=1}^{s} h_i \wedge [k_i, x] + \sum_{\phi \in \Phi} e_{-\phi} \wedge [e_\phi, x]\right). \tag{9.15}$$

In case $x = h \in \mathbf{h}$ this formula simplifies. The $[k_i, h] = 0$, and in the second sum we have

$$e_{-\phi} \wedge [e_\phi, h] = -\phi(h) e_{-\phi} \wedge e_\phi$$

which is invariant under the interchange of ϕ and $-\phi$. So let us make a choice Φ^+ of positive roots. Then we can write (9.15) as

$$\psi(h) = -\frac{1}{2} \sum_{\phi \in \Phi^+} \phi(h) e_{-\phi} \wedge e_\phi, \quad h \in \mathbf{h}. \tag{9.16}$$

Now

$$e_{-\phi} \wedge e_\phi = -1 + e_{-\phi} e_\phi.$$

So if

$$\rho := \frac{1}{2} \sum_{\phi \in \Phi^+} \phi \tag{9.17}$$

is one half the sum of the positive roots we have

$$\psi(h) = \rho(h) - \frac{1}{2} \sum_{\phi \in \Phi^+} \phi(h) e_{-\phi} e_\phi, \quad h \in \mathbf{h}. \tag{9.18}$$

In this equation, the multiplication on the right is in the Clifford algebra.

9.3 The spin representations.

If

$$\mathbf{p} = \mathbf{p}_1 \oplus \mathbf{p}_2$$

is a direct sum decomposition of a vector space $\mathbf{p}$ with a symmetric bilinear form into two orthogonal subspaces then it follows from the definition of the Clifford algebra that

$$C(\mathbf{p}) = C(\mathbf{p}_1) \otimes C(\mathbf{p}_2)$$

where the multiplication on the tensor product is taken in the sense of superalgebras, that is

$$(a_1 \otimes a_2)(b_1 \otimes b_2) := a_1 b_1 \otimes a_2 b_2$$

if either a_2 or b_1 are even, but

$$(a_1 \otimes a_2)(b_1 \otimes b_2) := -a_1 b_1 \otimes a_2 b_2$$

if both a_2 and b_1 are odd. It costs a sign to move one odd symbol past another.

9.3.1 The even dimensional case.

Suppose that $\mathbf{p}$ is even dimensional. If the metric is split (which is always the case if the metric is non-degenerate and we are over the complex numbers) then $\mathbf{p}$ is a direct sum of two dimensional mutually orthogonal split spaces, $\mathbf{W}_i$, so let us examine first the case of a two dimensional split space $\mathbf{p}$, spanned by ι, ϵ with $(\iota, \iota) = (\epsilon, \epsilon) = 0,\ (\iota, \epsilon) = \frac{1}{2}$. Let T be a one dimensional space with basis t and consider the linear map of $\mathbf{p} \to \mathrm{End}\ (\wedge T)$ determined by

$$\epsilon \mapsto \epsilon(t), \quad \iota \mapsto \iota(t^*)$$

where $\epsilon(t)$ denotes exterior multiplication by t and $\iota(t^*)$ denotes interior multiplication by t^*, the dual element to t in T^*. This is a Clifford map since

$$\epsilon(t)^2 = 0 = \iota(t^*)^2, \quad \epsilon(t)\iota(t^*) + \iota(t^*)\epsilon(t) = \mathrm{id}.$$

This therefore extends to a map of $C(\mathbf{p}) \to \mathrm{End}(\wedge T)$. Explicitly, if we use $1 \in \wedge^0 T,\ t \in \wedge^1 T$ as a basis of $\wedge T$ this map is given by

$$\begin{aligned} 1 &\mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \iota &\mapsto \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ \epsilon &\mapsto \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \\ \iota\epsilon &\mapsto \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

This shows that the map is an isomorphism. If now

$$\mathbf{p} = \mathbf{W_1} \oplus \cdots \oplus \mathbf{W}_m$$

is a direct sum of two dimensional split spaces, and we write

$$T = T_1 \oplus \cdots \oplus T_m$$

where the $C(\mathbf{W}_i) \cong \mathrm{End}(\wedge T_i)$ as above, then since

$$\wedge T = \wedge T_1 \otimes \cdots \otimes \wedge T_m$$

we see that

$$C(\mathbf{p}) \cong \mathrm{End}\ (\wedge T).$$

In particular, $C(\mathbf{p})$ is isomorphic to the full $2^m \times 2^m$ matrix algebra and hence has a unique (up to isomorphism) irreducible module. One model of this is

$$S = \wedge T.$$

We can write

$$S = S_+ \oplus S_-$$

as a supervector space, where we choose the standard $\mathbf{Z}_2$ grading on $\wedge T$ to determine the grading on S if m is even, but use the opposite grading (for reasons which will become apparent in a moment) if m is odd.

The even part, $C_{\mathbf{0}}(\mathbf{p})$ of $C(\mathbf{p})$ acts irreducibly on each of $S_\pm$. Since $\wedge^2\mathbf{p}$ together with the constants generates $C_{\mathbf{0}}(\mathbf{p})$ we see that the action of $\wedge^2\mathbf{p}$ on each of $S_\pm$ is irreducible. Since $\wedge^2\mathbf{p}$ under Clifford commutation is isomorphic to $o(\mathbf{p})$ the two modules $S_\pm$ give irreducible modules for the even orthogonal algebra $o(\mathbf{p})$. These are the half spin representations of the even orthogonal algebras.

We can identify $S = S_+ \oplus S_-$ as a left ideal in $C(\mathbf{p})$ as follows: Suppose that we write

$$\mathbf{p} = \mathbf{p}_+ \oplus \mathbf{p}_-$$

where $\mathbf{p}_\pm$ are complementary isotropic subspaces. Choose a basis $e_1^+, \dots, e_m^+$ of $\mathbf{p}_+$ and let

$$e_+ := e_1^+ \cdots e_m^+ = e_1^+ \wedge \cdots \wedge e_m^+ \in \wedge^m \mathbf{p}_+.$$

We have

$$y_+ e_+ = 0, \quad \forall\ y_+ \in \mathbf{p}_+$$

and hence

$$(\wedge \mathbf{p}_+)_+ e_+ = 0.$$

In other words

$$\wedge \mathbf{p}_+ e_+$$

consists of all scalar multiples of e_+.

Since

$$\wedge \mathbf{p}_- \otimes \wedge \mathbf{p}_+ \to C(\mathbf{p}), \quad w_- \otimes w_+ \mapsto w_- w_+$$

is a linear bijection, we see that

$$C(\mathbf{p}) e_+ = \wedge \mathbf{p}_- e_+.$$

This means that the left ideal generated by e_+ in $C(\mathbf{p})$ has dimension 2^m, and hence must be isomorphic as a left $C(\mathbf{p})$ module to S. In particular it is a minimal left ideal.

Let $e_1^-, \dots e_m^-$ be a basis of $\mathbf{p}_-$ and for any subset $J = \{i_1, \dots, i_j\}$, $i_1 < i_2 \cdots < i_j$ of $\{1, \dots, m\}$ let

$$e_-^J := e_{i_1}^- \wedge \cdots \wedge e_{i_j}^- = e_{i_1}^- \cdots e_{i_j}^-.$$

Then the elements

$$e_-^J e_+$$

form a basis of this model of S as J ranges over all subsets of $\{1, \dots, m\}$.

For example, suppose that we have a commutative Lie algebra $\mathbf{h}$ acting on $\mathbf{p}$ as infinitesimal isometries, so as to preserve each $\mathbf{p}_\pm$, that the e_i^+ are weight vectors corresponding to weights β_i and that the e_i^- form the dual basis, corresponding to the negative of these weights $-\beta_i$. Then it follows from (9.14) that the image, $\nu(h) \in \wedge^2(\mathbf{p}) \subset C(\mathbf{p})$ of an element $h \in \mathbf{h}$ is given by

$$\nu(h) = \frac{1}{2} \sum \beta_i(h) e_i^+ \wedge e_i^- = \frac{1}{2} \sum \beta_i(h)(1 - e_i^- e_i^+).$$

Thus

$$\nu(h)e_+ = \rho_{\mathbf{p}}(h)e_+ \tag{9.19}$$

where

$$\rho_{\mathbf{p}} := \frac{1}{2}(\beta_1 + \cdots + \beta_m). \tag{9.20}$$

For a subset J of $\{1, \dots, m\}$ let us set

$$\beta_J := \sum_{j \in J} \beta_j.$$

Then we have

$$[\nu(h), e_-^J] = -\beta_J(h) e_-^J$$

and so

$$\nu(h)(e_-^J e_+) = [\nu(h), e_-^J]e_+ + e_-^J \nu(h) e_+ = (\rho_{\mathbf{p}}(h) - \beta_J(h)) e_-^J e_+.$$

So if we denote the action of $\nu(h)$ on $S_\pm$ by $\mathrm{Spin}_\pm\, \nu(h)$ and the action of $\nu(h)$ on $S = S_+ \oplus S_-$ by Spin $\nu(h)$ we have proved that

$$\text{The } e_-^J e_+ \text{ are weight vectors of Spin } \nu \text{ with weights } \rho_{\mathbf{p}} - \beta_J. \tag{9.21}$$

It follows from (9.21) that the difference of the characters of $\mathrm{Spin}_+\nu$ and $\mathrm{Spin}_-\nu$ is given by

$$\mathrm{ch}_{Spin_+\nu} - \mathrm{ch}_{Spin_-\nu} = \prod_j \left(e(\frac{1}{2}\beta_j) - e(-\frac{1}{2}\beta_j) \right) = e(\rho_{\mathbf{p}}) \prod_j \left(1 - e(-\beta_j)\right). \tag{9.22}$$

There are two special cases which are of particular importance: First, this applies to the case where we take $\mathbf{h}$ to be a Cartan subalgebra of $o(\mathbf{p}) = o(\mathbf{C}^{2k})$

itself, say the diagonal matrices in the block decomposition of $o(\mathbf{p})$ given by the decomposition

$$\mathbf{C}^{2k} = \mathbf{C}^k \oplus \mathbf{C}^k$$

into two isotropic subspaces. In this case the β_i is just the i-th diagonal entry and (9.22) yields the standard formula for the difference of the characters of the spin representations of the even orthogonal algebras.

A second very important case is where we take $\mathbf{h}$ to be the Cartan subalgebra of a semi-simple Lie algebra $\mathbf{g}$, and take

$$\mathbf{p} := \mathbf{n}_+ \oplus \mathbf{n}_-$$

relative to a choice of positive roots. Then the β_j are just the positive roots, and we see that the right hand side of (9.22) is just the Weyl denominator, the denominator occurring in the Weyl character formula. This means that we can write the Weyl character formula as

$$\mathrm{ch}(\mathrm{Irr}(\lambda) \otimes S_+) - \mathrm{ch}(\mathrm{Irr}(\lambda) \otimes S_-) = \sum_{w \in W} (-1)^w e(w \bullet \lambda)$$

where

$$w \bullet \lambda := w(\lambda + \rho).$$

If we let U_μ denote the one dimensional module for $\mathbf{h}$ given by the weight μ we can drop the characters from the preceding equation and simply write the Weyl character formula as an equation in virtual representations of $\mathbf{h}$:

$$\mathrm{Irr}(\lambda) \otimes S_+ - \mathrm{Irr}(\lambda) \otimes S_- = \sum_{w \in W} (-1)^w U_{w \bullet \lambda}. \tag{9.23}$$

The reader can now go back to the preceding chapter and to Theorem 16 where this version of the Weyl character formula has been generalized from the Cartan subalgebra to the case of a reductive subalgebra of equal rank. In the next chapter we shall see the meaning of this generalization in terms of the Kostant Dirac operator.

9.3.2 The odd dimensional case.

Since every odd dimensional space with a non-singular bilinear form can be written as a sum of a one dimensional space and an even dimensional space (both non-degenerate), we need only look at the Clifford algebra of a one dimensional space with a basis element x such that $(x, x) = 1$ (since we are over the complex numbers). This Clifford algebra is two dimensional, spanned by 1 and x with $x^2 = 1$, the element x being odd. This algebra clearly has itself as a canonical module under left multiplication and is irreducible as a $\mathbf{Z}/2\mathbf{Z}$ module. We may call this the spin representation of Clifford algebra of a one dimensional space. Under the even part of the Clifford algebra (i.e. under the scalars) it splits into two isomorphic (one dimensional) spaces corresponding to the basis $1, x$ of the Clifford algebra. Relative to this basis $1, x$ we have the left multiplication representation given by

$$1 \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad x \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Let us use $C(\mathbf{C})$ to denote the Clifford algebra of the one dimensional orthogonal vector space just described, and $S(\mathbf{C})$ its canonical module. Then if

$$\mathbf{q} = \mathbf{p} \oplus \mathbf{C}$$

is an orthogonal decomposition of an odd dimensional vector space into a direct sum of an even dimensional space and a one dimensional space (both non-degenerate) we have

$$C(\mathbf{q}) \cong C(\mathbf{p}) \otimes C(\mathbf{C}) \cong \mathrm{End}(S(\mathbf{q}))$$

where

$$S(\mathbf{q}) := S(\mathbf{p}) \otimes S(\mathbf{C})$$

all tensor products being taken in the sense of superalgebra. We have a decomposition

$$S(\mathbf{q}) = S_+(\mathbf{q}) \oplus S_-(\mathbf{q})$$

as a super vector space where

$$S_+(\mathbf{q}) = S_+(\mathbf{p}) \oplus xS_-(\mathbf{p}),\ S_-(\mathbf{q}) = S_-(\mathbf{p}) \oplus xS_+(\mathbf{p}).$$

These two spaces are equivalent and irreducible as $C_{\mathbf{0}}(\mathbf{q})$ modules. Since the even part of the Clifford algebra is generated by $\wedge^2\mathbf{q}$ together with the scalars, we see that either of these spaces is a model for the irreducible spin representation of $o(\mathbf{q})$ in this odd dimensional case.

Consider the decomposition $\mathbf{p} = \mathbf{p}_+ \oplus \mathbf{p}_-$ that we used to construct a model for $S(\mathbf{p})$ as being the left ideal in $C(\mathbf{p})$ generated by $\wedge^m\mathbf{p}_+$ where $m = \dim \mathbf{p}_+$. We have

$$\wedge(\mathbf{C} \oplus \mathbf{p}_-) = \wedge(\mathbf{C}) \otimes \wedge\mathbf{p}_-,$$

and

Proposition 28 *The left ideal in the Clifford algebra generated by $\wedge^m\mathbf{p}_+$ is a model for the spin representation.*

Notice that this description is valid for both the even and the odd dimensional case.

9.3.3 Spin ad and V_ρ.

We want to consider the following situation: $\mathbf{g}$ is a simple Lie algebra and we take (,) to be the Killing form. We have

$$\Phi : \mathbf{g} \to \wedge^2\mathbf{g} \subset C(\mathbf{g})$$

which is the map ν associated to the adjoint representation of $\mathbf{g}$. Let $\mathbf{h}$ be a Cartan subalgebra and Φ the collection of roots. We choose root vectors $e_\phi, \phi \in \Phi$ so that

$$(e_\phi, e_{-\phi}) = 1.$$

Then it follows from (9.14) that

$$\Phi(x) = \frac{1}{4}\left(\sum h_i \wedge [k_i, x]_{\mathbf{g}} + \sum_{\phi\in\Phi} e_{-\phi} \wedge [e_\phi, x]_{\mathbf{g}}\right) \tag{9.24}$$

where the brackets are the Lie brackets of $\mathbf{g}$, where the h_i range over a basis of $\mathbf{h}$ and the k_i over a dual basis. This equation simplifies in the special cases where $x = h \in \mathbf{h}$ and in the case where $x = e_\psi,\ \psi \in \Phi^+$ relative to a choice, Φ^+ of positive roots. In the case that $x = h \in \mathbf{h}$ we have seen that $[k_i, h] = 0$ and the equation simplifies to

$$\Phi(h) = \rho(h)1 - \frac{1}{2}\sum_{\phi\in\Phi^+} \phi(h)e_{-\phi}e_\phi \tag{9.25}$$

where

$$\rho = \frac{1}{2}\sum_{\phi\in\Phi^+}\phi$$

is one half the sum of then positive roots.

We claim that for $\psi \in \Phi^+$ we have

$$\Phi(e_\psi) = \sum x_{\gamma'} e_{\psi'} \tag{9.26}$$

where the sum is over pairs (γ', ψ') such that either

1. $\gamma' = 0, \psi' = \psi$ and $x_{\gamma'} \in \mathbf{h}$ or
2. $\gamma' \in \Phi,\ \psi' \in \Phi_+$ and $\gamma' + \psi' = \psi,$ and $x_{\gamma'} \in \mathbf{g}_{\gamma'}$.

To see this, first observe that this first sum on the right of (9.24) gives

$$\sum \psi(k_i)h_i \wedge e_\psi$$

and so all these summands are of the form 1). For each summand

$$e_{-\phi} \wedge [e_\phi, e_\psi]$$

of the second sum, we may assume that either $\phi + \psi = 0$ or that $\phi + \psi \in \Phi$ for otherwise $[e_\phi, e_\psi] = 0$. If $\phi + \psi = 0$, so $\psi = -\phi \neq 0$, we have $[e_\phi, e_\psi] \in \mathbf{h}$ which is orthogonal to $e_{-\phi}$ since $\phi \neq 0$. So

$$e_{-\phi} \wedge [e_\phi, e_\psi] = -[e_\phi, e_\psi]e_\psi$$

again has the form 1).

If $\phi + \psi = \tau \neq 0$ is a root, then $(e_{-\phi}, e_\tau) = 0$ since $\phi \neq \tau$. If $\tau \in \Phi_+$ then

$$e_{-\phi} \wedge [e_\phi, e_\psi] = e_{-\phi} y_\tau,$$

where y_τ is a multiple of e_τ so this summand is of the form 2). If τ is a negative root, the ϕ must be a negative root so $-\phi$ is a positive root, and we can switch the order of the factors in the preceding expression at the expense of introducing a sign. So again this is of the form 2), completing the proof of (9.26).

Let $\mathbf{n}_+$ be the subalgebra of $\mathbf{g}$ generated by the positive root vectors and similarly $\mathbf{n}_-$ the subalgebra generated by the negative root vectors so

$$\mathbf{g} = \mathbf{n}_+ \oplus \mathbf{b}_-, \quad \mathbf{b}_- := \mathbf{n}_- \oplus \mathbf{h}$$

is an $\mathbf{h}$ stable decomposition of $\mathbf{g}$ into a direct sum of the nilradical and its opposite Borel subalgebra.

Let N be the number of positive roots and let

$$0 \neq n \in \wedge^N \mathbf{n}_+.$$

Clearly

$$yn = 0 \quad \forall y \in \mathbf{n}_+.$$

Hence by (9.26) we have

$$\Phi(\mathbf{n}_+)n = 0$$

while by (9.25)

$$\Phi(h)n = \rho(h)n \ \ \forall\, h \in \mathbf{h}.$$

This implies that the cyclic module

$$\Phi(U(\mathbf{g}))n$$

is a model for the irreducible representation V_ρ of $\mathbf{g}$ with highest weight ρ. Left multiplication by $\Phi(x), \ \ x \in \mathbf{g}$ gives the action of $\mathbf{g}$ on this module.

Furthermore, if $nc \neq 0$ for some $c \in C(\mathbf{g})$ then nc has the same property:

$$\Phi(\mathbf{n}_+)nc = 0, \quad \Phi(h)nc = \rho(h)nc, \ \ \forall\, h \in \mathbf{h}.$$

Thus every $nc \neq 0$ also generates a $\mathbf{g}$ module isomorphic to V_ρ.

Now the map

$$\wedge \mathbf{n}_+ \otimes \wedge \mathbf{b}_- \to C(\mathbf{g}), \quad x \otimes b \to xb$$

is a linear isomorphism and right Clifford multiplication of $\wedge^N \mathbf{n}_+$ by $\wedge \mathbf{n}_+$ is just $\wedge^N \mathbf{n}_+$, all the elements of of $\wedge^+ \mathbf{n}_+$ yielding 0. So we have the vector space isomorphism

$$nC(\mathbf{g}) = \wedge^N \mathbf{n}_+ \otimes \wedge \mathbf{b}_-.$$

In other words,

$$\Phi(U(\mathbf{g}))nC(\mathbf{g})$$

is a direct sum of irreducible modules all isomorphic to V_ρ with multiplicity equal to

$$\dim \wedge \mathbf{b}_- = 2^{s+N}$$

where $s = \dim \mathbf{h}$ and $N = \dim \mathbf{n}_- = \dim \mathbf{n}_+$. Let us compute the dimension of V_ρ using the Weyl dimension formula which asserts that for any irreducible finite dimensional representation V_λ with highest weight λ we have

$$\dim V_\lambda = \frac{\prod_{\phi \in \Phi_+} (\rho + \lambda, \phi)}{\prod_{\phi \in \Phi_+} (\rho, \phi)}.$$

If we plug in $\lambda = \rho$ we see that each factor in the numerator is twice the corresponding factor in the denominator so

$$\dim V_\rho = 2^N. \tag{9.27}$$

But then

$$\dim \Phi(U(\mathbf{g})) n C(\mathbf{g}) = 2^{s+2N} = \dim C(\mathbf{g}).$$

This implies that

$$C(\mathbf{g}) = \Phi(U(\mathbf{g})) n C(\mathbf{g}) = \Phi(U(\mathbf{g})) n (\wedge \mathbf{b}_-), \tag{9.28}$$

proving that $C(\mathbf{g})$ is primary of type V_ρ with multiplicity 2^{s+N} as a representation of $\mathbf{g}$ under the left multiplication action of $\Phi(\mathbf{g})$.

This implies that any submodule for this action, in particular any left ideal of $C(\mathbf{g})$, is primary of type V_ρ. Since we have realized the spin representation of $C(\mathbf{g})$ as a left ideal in $C(\mathbf{g})$ we have proved the important

Theorem 17 Spin ad *is primary of type* V_ρ.

One consequence of this theorem is the following:

Proposition 29 *The weights of* V_ρ *are*

$$\rho - \phi_J \tag{9.29}$$

where J *ranges over subsets of the positive roots and*

$$\phi_J = \sum_{\phi \in J} \phi_J$$

each occurring with multiplicity equal to the number of subsets J *yielding the same value of* ϕ_J.

Indeed, (9.21) gives the weights of Spin ad, but several of the β_J are equal due to the trivial action of ad($\mathbf{h}$) on itself. However this contribution to the multiplicity of each weight occurring in (9.21) is the same, and hence is equal to the multiplicity of V_ρ in Spin ad. So each weight vector of V_ρ must be of the form (9.29) each occurring with the multiplicity given in the proposition.

Chapter 10

The Kostant Dirac operator

Let $\mathbf{p}$ be a vector space with a non-degenerate symmetric bilinear form. We have the Clifford algebra $C(\mathbf{p})$ and the identification of $o(\mathbf{p}) = \wedge^2(\mathbf{p})$ inside $C(\mathbf{p})$.

10.1 Antisymmetric trilinear forms.

Let ϕ be an antisymmetric trilinear form on $\mathbf{p}$. Then ϕ defines an antisymmetric map

$$b = b_\phi : \mathbf{p} \otimes \mathbf{p} \to \mathbf{p}$$

by the formula

$$(b(y, y'), y'') = \phi(y, y', y'') \;\; \forall \; y, y', y'' \in \mathbf{p}.$$

This bilinear map "leaves (,) invariant" in the sense that

$$(b(y, y'), y'') = (y, b(y', y'')).$$

Conversely, any antisymmetric map $b : \mathbf{p} \otimes \mathbf{p} \to \mathbf{p}$ satisfying this condition defines an antisymmetric form ϕ. Finally either of these two objects defines an element $v \in \wedge^3\mathbf{p}$ by

$$-2(v, y \wedge y' \wedge y'') = (b(y, y'), y'') = \phi(y, y', y''). \tag{10.1}$$

We can write this relation in several alternative ways: Since

$$-2(v, y \wedge y' \wedge y'') = -2(\iota(y')\iota(y)v, y'') = 2(\iota(y)\iota(y')v, y'')$$

we have

$$b(y, y') = 2\iota(y)\iota(y')v. \tag{10.2}$$

Also, $\iota(y)v \in \wedge^2\mathbf{p}$ and so is identified with an element of $o(\mathbf{p})$ by commutator in the Cliford algebra:

$$\mathrm{ad}(\iota(y)v)(y') = [\iota(y)v, y'] = -2\iota(y')\iota(y)v$$

so

$$\mathrm{ad}(\iota(y)v)(y') = [\iota(y)v, y'] = b(y, y'). \tag{10.3}$$

10.2 Jacobi and Clifford.

Given an antisymmetric bilinear map $b : \mathbf{p} \otimes \mathbf{p} \to \mathbf{p}$ we may define

$$\mathrm{Jac}(b) : \mathbf{p} \otimes \mathbf{p} \otimes \mathbf{p} \to \mathbf{p}$$

by

$$\mathrm{Jac}(b)(y, y', y'') = b(b(y, y'), y'') + b(b(y', y''), y) + b(b(y'', y), y')$$

so that the vanishing of $\mathrm{Jac}(b)$ is the usual Jacobi identity. It is easy to check that $\mathrm{Jac}(b)$ is antisymmetric and that if b satisfies $(b(y, y'), y'') = (y, b(y', y''))$ then the four form

$$y, y', y'', y''' \mapsto (\mathrm{Jac}(b)(y, y', y''), y''')$$

is antisymmetric. We claim that if $v \in \wedge^3\mathbf{p}$ as in the preceding subsection, then

$$\iota(y'')\iota(y')\iota(y)v^2 = \frac{1}{2}\,\mathrm{Jac}(b)(y, y', y''). \tag{10.4}$$

To prove this observe that

$$\begin{aligned}
\iota(y)v^2 &= (\iota(y)v)v - v(\iota(y)v)\\
\iota(y')\iota(y)v^2 &= (\iota(y')\iota(y)v)v + (\iota(y)v)(\iota(y')v) - (\iota(y')v)(\iota(y)v) + v(\iota(y')\iota(y)v)\\
\iota(y'')\iota(y')\iota(y)v^2 &= -(\iota(y')\iota(y)v)\iota(y'')v + (\iota(y'')v)(\iota(y')\iota(y)v) + (\iota(y'')\iota(y)v)\iota(y')v\\
&\quad +(\iota(y)v)(\iota(y'')\iota(y')v - (\iota(y'')\iota(y')v)(\iota(y)v) - (\iota(y')v)(\iota(y'')\iota(y)v))\\
&= [\iota(y'')v, \iota(y')\iota(y)v] + [\iota(y')v, \iota(y)\iota(y'')v] + [\iota(y)v, \iota(y')\iota(y'')v]\\
&= \frac{1}{2}\,\mathrm{Jac}(b)(y, y', y'')
\end{aligned}$$

by (10.2) and (10.3).

Equation (10.4) describes the degree four component of v^2 in terms of $\mathrm{Jac}(b)$. We can be explicit about the degree zero component of v^2. We claim that

$$(v^2)_0 = \frac{1}{24}\,\mathrm{tr}\sum_{j=1}^{n}[y \to \epsilon_j b(y_j, b(y_j, y))], \quad \epsilon_j := (y_j, y_j). \tag{10.5}$$

Indeed, by (9.6) we know that $(v^2)_0 = -(v, v)$ and since $y_i \wedge y_j \wedge y_k,\ i < j < k$

form an "orthonormal" basis of $\wedge^3\mathbf{p}$ we have

$$\begin{aligned}
-(v,v) &= -\sum_{1\le i<j<k\le n} \pm(v, y_i\wedge y_j\wedge y_k)^2,\ \pm=\epsilon_i\epsilon_j\epsilon_k\\
&= -\frac{1}{6}\sum_{i=1,j=1,k=1}^{n,n,n} \pm(v, y_i\wedge y_j\wedge y_k)^2\\
&= -\frac{1}{6}\sum_{i=1,j=1,k=1}^{n,n,n} \pm(\iota(y_k)\iota(y_j)v, y_i)^2\\
&= -\frac{1}{24}\sum_{i=1,j=1,k=1}^{n,n,n} \pm(b(y_j,y_k), y_i)^2\\
&= -\frac{1}{24}\sum_{j=1,k=1}^{n,n} \epsilon_j\epsilon_k(b(y_j,y_k), b(y_j,y_k))\\
&= \frac{1}{24}\sum_{j=1,k=1}^{n,n} \epsilon_j\epsilon_k(b(y_j,b(y_j,y_k)), y_k)
\end{aligned}$$

proving (10.5).

10.3 Orthogonal extension of a Lie algebra.

Let us get back to the general case of a Lie algebra $\mathbf{r}$ acting as infinitesimal orthogonal transformations on $\mathbf{p}$ and the map $\nu : \mathbf{r} \to \wedge^2\mathbf{p}$ given by (9.13). Suppose that the Lie algebra $\mathbf{r}$ has a non-degenerate invariant symmetric bilinear form $(\ ,\ )_{\mathbf{r}}$. We have the transpose map

$$\nu^\dagger : \wedge^2\mathbf{p} \to \mathbf{r}$$

since both $\mathbf{r}$ and $\wedge^2\mathbf{p}$ have non-degenerate symmetric bilinear forms. For y and y' in $\mathbf{p}$, let us define

$$[y,y']_{\mathbf{r}} := -2\nu^\dagger(y\wedge y'),.$$

This map is an $\mathbf{r}$ morphism which says that

$$[x,[y,y']_{\mathbf{r}}] = [x\cdot y, y']_{\mathbf{r}} + [y, x\cdot y']_{\mathbf{r}}, \tag{10.6}$$

where the bracket on the left denotes the Lie bracket on $\mathbf{r}$. Also, we have

$$\begin{aligned}
(x,[y,y']_{\mathbf{r}})_{\mathbf{r}} &= -2(x,\nu^\dagger(y\wedge y'))_{\mathbf{r}}\\
&= -2(\nu(x), y\wedge y')_{\mathbf{p}}\\
&= -2(\iota(y)\nu(x), y')_{\mathbf{p}}\\
&= (x\cdot y, y')_{\mathbf{p}}.
\end{aligned}$$

So we have proved
$$(x, [y, y']_{\mathbf{r}})_{\mathbf{r}} = (x \cdot y, y')_{\mathbf{p}}. \tag{10.7}$$
This has the following significance: Suppose that we want to make $\mathbf{r} \oplus \mathbf{p}$ into a Lie algebra with an invariant symmetric bilinear form $(\ ,\ )$ such that

- $\mathbf{r}$ and $\mathbf{p}$ are orthogonal under $(\ ,\ )$,
- the restriction of $(\ ,\ )$ to $\mathbf{r}$ is $(\ ,\ )_{\mathbf{r}}$ and the restriction of $(\ ,\ )$ to $\mathbf{p}$ is $(\ ,\ )_{\mathbf{p}}$, and
- $[\mathbf{r}, \mathbf{p}] \subset \mathbf{p}$ and the bracket of an element of $\mathbf{r}$ with an element of $\mathbf{p}$ is given by $[x, y] = x \cdot y$.

Then
$$\text{the } \mathbf{r} \text{ component of } [y, y'] \text{ must be given by } [y, y']_{\mathbf{r}}.$$

Thus to define a Lie algebra structure on $\mathbf{r} \oplus \mathbf{p}$ we must specify the $\mathbf{p}$ component of the bracket of two elements of $\mathbf{p}$. This amounts to specifying a $v \in \wedge^3\mathbf{p}$ as we have seen, and the condition that the Jacobi identity hold for x, y, y' with $x \in \mathbf{r}$ and $y, y' \in \mathbf{p}$ amounts to the condition that $v \in \wedge^3\mathbf{p}$ be invariant under the action of $\mathbf{r}$. It then follows that if we try to define $[\ ,\ ] = [\ ,\ ]_v$ by
$$[y, y'] = [y, y']_{\mathbf{r}} + 2\iota(y)\iota(y')v$$
then
$$([z, z'], z'') = (z, [z', z''])$$
for any three elements of $\mathbf{g} := \mathbf{r} \oplus \mathbf{p}$, and the Jacobi identity is satisfied if at least one of these elements belongs to $\mathbf{r}$. Furthermore, for any $x \in \mathbf{r}$ we have

$$\begin{aligned}
([[y, y'], y''], x) &= ([[y, y'], y''], x)_{\mathbf{r}} \\
&= ([y, y'], [y'', x])_{\mathbf{p}} \quad \text{by (10.7)} \\
&= ([x, [y, y']], y'')_{\mathbf{p}} \\
&= ([[x, y], y'], y'')_{\mathbf{p}} + ([y, [x, y']], y'')_{\mathbf{p}} \\
&= ([x, y], [y', y''])_{\mathbf{p}} + ([x, y'], [y'', y])_{\mathbf{p}} \\
&= (x, [y, [y', y'']])_{\mathbf{r}} + (x, [y', [y'', y]])_{\mathbf{r}}
\end{aligned}$$

or
$$([[y, y'], y''] + [[y', y''], y] + [[y'', y], y'], x) = 0.$$

In other words, the $\mathbf{r}$ component of the Jacobi identity holds for three elements of $\mathbf{p}$.

So what remains to be checked is the $\mathbf{p}$ component of the Jacobi identity for three elements of $\mathbf{p}$. This is the sum
$$\mathrm{Jac}(b)(y, y', y'') + [y, y']_{\mathbf{r}} \cdot y'' + [y', y'']_{\mathbf{r}} \cdot y + [y'', y]_{\mathbf{r}} \cdot y'.$$

Let us choose an"orthonormal" basis $\{x_i\},\ i = 1, \ldots, r$ of $\mathbf{r}$ and write

$$[y, y']_{\mathbf{r}} = \sum_i \epsilon_i([y, y'], x_i)x_i, \quad \epsilon_i := (x_i, x_i)_{\mathbf{r}} = \pm 1$$

so

$$[y, y']_{\mathbf{r}} \cdot y'' = \sum_i \epsilon_i([y, y']_{\mathbf{r}}, x_i)x_i \cdot y''.$$

Then by (9.11) and (10.4) we see that the Jacobi identity is

$$\iota(y)\iota(y')\iota(y'')\left(v^2 + \nu(\mathrm{Cas}_{\mathbf{r}})\right) = 0$$

where

$$\mathrm{Cas}_{\mathbf{r}} := \sum_i \epsilon_i x_i^2 \in U(\mathbf{r})$$

does not depend on the choice of basis, and $\nu : U(\mathbf{r}) \to C(\mathbf{p})$ is the extension of the homomorphism $\nu : \mathbf{r} \to C(\mathbf{p})$. In particular, we have proved that v defines an extension of the Lie algebra structure satisfying our condition if and only if

$$v^2 + \nu(\mathrm{Cas}_{\mathbf{r}}) \in \mathbf{C} \tag{10.8}$$

i.e. has no component of degree four.

Suppose that this condition holds. We then have defined a Lie algebra structure on

$$\mathbf{g} = \mathbf{r} \oplus \mathbf{p}.$$

We let $P_{\mathbf{r}}$ and $P_{\mathbf{p}}$ denote projections onto the first and second components of our decomposition. Our Lie bracket on $\mathbf{g}$, denoted simply by [,] satisfies

$$\begin{aligned}
[x, x'] &= [x, x']_{\mathbf{r}}, \quad x, x' \in \mathbf{r} && (10.9)\\
[x, y] &= x \cdot y, \quad x \in \mathbf{r}, y \in \mathbf{p} && (10.10)\\
P_{\mathbf{r}}[y, y'] &= [y, y']_{\mathbf{r}} = -2\nu^\dagger(y \wedge y') \quad y, y' \in \mathbf{p} && (10.11)\\
P_{\mathbf{p}}[y, y'] &= b(y, y') = 2\iota(y)\iota(y')v, \quad y, y' \in \mathbf{p}. && (10.12)
\end{aligned}$$

From now on we will assume that we are over the complex numbers or that we are over the reals and the symmetric bilinear forms are positive definite. This is not for any mathematical reasons but because the formulas become a bit complicated if we put in all the signs. We leave the general case to the reader.

10.4 The value of $[v^2 + \nu(\mathrm{Cas}_{\mathbf{r}})]_0$.

Condition (10.8) says that the degree four component of $v^2 + \nu(\mathrm{Cas}_{\mathbf{r}})$ vanishes. Assume that this holds, so we have constructed a Lie algebra. We will now compute the degree zero component of $v^2 + \nu(\mathrm{Cas}_{\mathbf{r}})$. The answer will be given in equation (10.13) below.

We can write (10.5) as

$$\begin{aligned}(v^2)_0 &= \frac{1}{24}\operatorname{tr} P_{\mathbf{p}}\sum_{j=1}^{n}\operatorname{ad}_{\mathbf{g}}(y_j)P_{\mathbf{p}}\operatorname{ad}_{\mathbf{g}}(y_j)P_{\mathbf{p}}\\ &= \frac{1}{24}\operatorname{tr} P_{\mathbf{p}}\sum_{j=1}^{n}\operatorname{ad}_{\mathbf{g}}(y_j)P_{\mathbf{p}}\operatorname{ad}_{\mathbf{g}}(y_j)\end{aligned}$$

in view of (10.12) where $\operatorname{ad}_{\mathbf{g}}$ denotes the adjoint action on all of $\mathbf{g}$. On the other hand, we have from (9.12) that

$$\nu(\mathrm{Cas}_{\mathbf{r}})_0 = \frac{1}{8}\operatorname{tr}\sum_i(\operatorname{ad} x_i)^2 P_{\mathbf{p}} = \frac{1}{8}\sum_{i=1}^{r}\sum_{j=1}^{n}([x_i,[x_i,y_j]],y_j).$$

We can rewrite this sum as

$$\frac{1}{8}\sum_{i=1,j=1}^{r,n}([x_i,y_j],[y_j,x_i])$$

which equals $\frac{1}{8}\sum_{i=1,j=1,k=1}^{r,n,n}([x_i,y_j],y_k)(y_k,[y_j,x_i])$. But this equals

$$\begin{aligned}&\frac{1}{8}\sum_{i=1,j=1,k=1}^{r,n,n}(x_i,[y_j,y_k])([y_k,y_j],x_i)\\ =\ &\frac{1}{8}\sum_{j=1,k=1}(P_{\mathbf{r}}[y_j,y_k],P_{\mathbf{r}}[y_k,y_j])\\ =\ &\frac{1}{8}\sum_{j=1,k=1}^{n,n}(P_{\mathbf{r}}[y_j,y_k],[y_k,y_j])\\ =\ &\frac{1}{8}\sum_{j=1,k=1}^{n,n}([y_j,P_{\mathbf{r}}[y_j,y_k]],y_k)\\ =\ &\frac{1}{8}\operatorname{tr} P_{\mathbf{p}}\sum_{j=1}^{n}\operatorname{ad}_{\mathbf{g}}(y_j)P_{\mathbf{r}}\operatorname{ad}_{\mathbf{g}}(y_j)P_{\mathbf{p}}.\end{aligned}$$

In other words

$$\nu(\mathrm{Cas}_{\mathbf{r}})_0 = \frac{1}{8}\operatorname{tr} P_{\mathbf{p}}\sum_{j=1}^{n}\operatorname{ad}_{\mathbf{g}}(y_j)P_{\mathbf{r}}\operatorname{ad}_{\mathbf{g}}(y_j)P_{\mathbf{p}} = \frac{1}{8}\operatorname{tr}\sum_{j=1}^{n}\operatorname{ad}_{\mathbf{g}}(y_j)P_{\mathbf{r}}\operatorname{ad}_{\mathbf{g}}(y_j)P_{\mathbf{p}}.$$

Multiplying this equation by 1/3 and adding it to the above expression for $(v^2)_0$ gives

$$\frac{1}{3}\nu(\mathrm{Cas}_{\mathbf{r}})_0 + (v^2)_0 = \frac{1}{24}\operatorname{tr}\sum_{j=1}^{n}(\operatorname{ad}_{\mathbf{g}} y_j)^2 P_{\mathbf{p}}.$$

We can write

$$\nu(\mathrm{Cas}_{\mathbf{r}})_0 = \frac{1}{8}\sum_{i=1,j=1}^{r,n}([x_i,y_j],[y_j,x_i]) = \frac{1}{8}\sum_{i=1,j=1}^{r,n}(x_i,[y_j,[y_j,x_i]])$$

$$= \frac{1}{8}\operatorname{tr} P_{\mathbf{r}}\sum_{j=1}^{n}(\mathrm{ad}_{\mathbf{g}}\, y_j)^2 P_{\mathbf{r}} = \frac{1}{8}\operatorname{tr}\sum_{j=1}^{n}(\mathrm{ad}_{\mathbf{g}}\, y_j)^2 P_{\mathbf{r}}.$$

Multiplying by 1/3 and adding to the preceding equation gives

$$\frac{2}{3}\nu(\mathrm{Cas}_{\mathbf{r}})_0 + (v^2)_0 = \frac{1}{24}\operatorname{tr}\sum_{j=1}^{n}(\mathrm{ad}_{\mathbf{g}}\, y_j)^2.$$

On the other hand

$$\nu(\mathrm{Cas}_{\mathbf{r}})_0 = \frac{1}{8}\sum_{i=1,j=1}^{r,n}([x_i,[x_i,y_j]],y_j) = \frac{1}{8}\operatorname{tr}\mathrm{ad}_{\mathbf{p}}(\mathrm{Cas}_{\mathbf{r}})$$

$$= \frac{1}{8}\left(\operatorname{tr}\mathrm{ad}_{\mathbf{g}}(\mathrm{Cas}_{\mathbf{r}}) - \operatorname{tr}\mathrm{ad}_{\mathbf{r}}(\mathrm{Cas}_{\mathbf{r}})\right).$$

Multiplying by 1/3 and adding to the preceding equation, and using the fact that $\mathrm{Cas}_{\mathbf{g}} = \mathrm{Cas}_{\mathbf{r}} + \sum y_j^2$ gives

$$\nu(\mathrm{Cas}_{\mathbf{r}}) + v^2 = \frac{1}{24}\left(\operatorname{tr}\mathrm{ad}_{\mathbf{g}}(\mathrm{Cas}_{\mathbf{g}}) - \operatorname{tr}\mathrm{ad}_{\mathbf{r}}(\mathrm{Cas}_{\mathbf{r}})\right) \tag{10.13}$$

when (10.8) holds.

Suppose now that the Lie algebra $\mathbf{r}$ is reductive and that the Lie algebra $\mathbf{g}$ we created out of $\mathbf{r}$ and $\mathbf{p}$ using a $v \in \wedge^3\mathbf{p}$ satisfying (10.8) is also reductive. Using (7.29) for $\mathbf{g}$ and for $\mathbf{r}$ in the right hand side of (10.13) yields

$$\nu(\mathrm{Cas}_{\mathbf{r}}) + v^2 = \left((\rho_{\mathbf{g}},\rho_{\mathbf{g}}) - (\rho_{\mathbf{r}},\rho_{\mathbf{r}})\right). \tag{10.14}$$

10.5 Kostant's Dirac Operator.

Suppose that we have constructed our Lie algebra $\mathbf{g} = \mathbf{r} + \mathbf{p}$ from a $v \in \wedge^3\mathbf{p}$ satisfying (10.8). We are going to define

$$\not{K} \in U(\mathbf{g}) \otimes C(\mathbf{p})$$

as follows: Let $y_1,\dots,y_n$ be an orthonormal basis of $\mathbf{p}$. Then

$$\not{K} := \sum_i y_i \otimes y_i + 1 \otimes v. \tag{10.15}$$

(On the left of the tensor product sign the $y_i \in \mathbf{p}$ is considered as an element of $U(\mathbf{g})$ via the canonical injection of $\mathbf{p}$ in $U(\mathbf{p}) \subset U(\mathbf{g})$ and on the right of the tensor product sign it lies in $C(\mathbf{p})$ via the canonical injection of $\mathbf{p}$ into $C(\mathbf{p})$.)

We have a homomorphism $U(\mathbf{r}) \to U(\mathbf{g})$, in particular a Lie algebra injection $\mathbf{r} \to U(\mathbf{g})$. We also have a Lie algebra homomorphism $\nu : \mathbf{r} \to C(\mathbf{p})$. In particular, we have the diagonal Lie algebra map

$$\mathrm{diag} : \mathbf{r} \to U(\mathbf{g}) \otimes C(\mathbf{p}), \quad \mathrm{diag}(x) = x \otimes 1 + 1 \otimes \nu(x)$$

and this extends to an algebra map

$$\mathrm{diag} : U(\mathbf{r}) \to U(\mathbf{g}) \otimes C(\mathbf{p}).$$

For example,

$$\mathrm{diag}(\mathrm{Cas}_{\mathbf{r}}) = \sum_i \left(x_i \otimes 1 + 1 \otimes \nu(x_i)\right)^2$$

where $x_1, \ldots, x_r$ is an orthonormal basis of $\mathbf{r}$. In other words

$$\mathrm{diag}(\mathrm{Cas}_{\mathbf{r}}) = \sum_{i=1}^r x_i^2 \otimes 1 + 2\sum_{i=1}^r x_i \otimes \nu(x_i) + \sum_{i=1}^r 1 \otimes \nu(x_i)^2. \tag{10.16}$$

We claim that

$$\not{K}^2 = \mathrm{Cas}_{\mathbf{g}} \otimes 1 - \mathrm{diag}(\mathrm{Cas}_{\mathbf{r}}) + \frac{1}{24}\left(\mathrm{tr}\,\mathrm{ad}_{\mathbf{g}}(\mathrm{Cas}_{\mathbf{g}}) - \mathrm{tr}\,\mathrm{ad}_{\mathbf{r}}(\mathrm{Cas}_{\mathbf{r}})\right) 1 \otimes 1. \tag{10.17}$$

To prove this, let us write (10.15) as

$$\not{K} = \not{K}' + \not{K}''.$$

So

$$(\not{K}'')^2 = 1 \otimes v^2$$

and hence

$$(\not{K}'')^2 + \sum_{j=1}^r 1 \otimes \nu(x_i)^2 = \frac{1}{24}\left(\mathrm{tr}\,\mathrm{ad}_{\mathbf{g}}(\mathrm{Cas}_{\mathbf{g}}) - \mathrm{tr}\,\mathrm{ad}_{\mathbf{r}}(\mathrm{Cas}_{\mathbf{r}})\right) 1 \otimes 1$$

by (10.13). We have

$$\begin{aligned}
(\not{K}')^2 &= \sum_{ij} y_i y_j \otimes y_i y_j \\
&= \sum_i y_i^2 \otimes 1 + \sum_{i \neq j} y_i y_j \otimes y_i y_j \\
&= \sum_i y_i^2 \otimes 1 + \sum_{i<j} (y_i y_j - y_j y_i) \otimes y_i y_j \\
&= \sum_i y_i^2 \otimes 1 + \sum_{i<j} [y_i, y_j] \otimes y_i y_j \\
&= \sum_i y_i^2 \otimes 1 - 2\sum_{i<j} \nu^\dagger(y_i \wedge y_j) \otimes y_i \wedge y_j + 2\sum_{i<j} \iota(y_i)\iota(y_j)v \otimes y_i \wedge y_j,
\end{aligned}$$

where we have used the decomposition of $[y_i, y_j]$ into its $\mathbf{r}$ and $\mathbf{p}$ components to get to the last expression. We can write the middle term in the last expression as

$$\begin{aligned}
-2\sum_{i<j}\nu^\dagger(y_i\wedge y_j)\otimes y_i\wedge y_j &= -2\sum_{k=1}^r\sum_{i<j}(\nu^\dagger(y_i\wedge y_j),x_k)x_k\otimes y_i\wedge y_j\\
&= -2\sum_{k=1}^r\sum_{i<j}(y_i\wedge y_j,\nu(x_k))x_k\otimes y_i\wedge y_j\\
&= -2\sum_{k=1}^r\sum_{i<j}x_k\otimes(y_i\wedge y_j,\nu(x_k))y_i\wedge y_j\\
&= -2\sum_i x_i\otimes\nu(x_i).
\end{aligned}$$

Since $\sum x_i^2\otimes 1+\sum y_j^2\otimes 1 = \mathrm{Cas}_{\mathbf{g}}\otimes 1$ we conclude that

$$(\not K')^2+(\not K'')^2+\mathrm{diag}(\mathrm{Cas}_{\mathbf{r}})$$

$$= \mathrm{Cas}_{\mathbf{g}}\otimes 1+\frac{1}{24}\left(\mathrm{tr}\,\mathrm{ad}_{\mathbf{g}}(\mathrm{Cas}_{\mathbf{g}})-\mathrm{tr}\,\mathrm{ad}_{\mathbf{r}}(\mathrm{Cas}_{\mathbf{r}})\right)1\otimes 1+2\sum_{i<j}\iota(y_i)\iota(y_j)v\otimes y_i\wedge y_j.$$

To complete the proof of (10.17) we must show that

$$\not K'\not K''+\not K''\not K' = -2\sum_{i<j}\iota(y_i)\iota(y_j)v\otimes y_i\wedge y_j.$$

But

$$\begin{aligned}
\not K'\not K''+\not K''\not K' &= \sum y_j\otimes[y_j,v]\\
&= 2\sum y_j\otimes\iota(y_j)v\\
&= 2\sum_{i<k}\sum_j y_j\otimes(\iota(y_j)v,y_i\wedge y_k)y_i\wedge y_k\\
&= \sum_{i<k}\sum_j y_j\otimes(2v,y_i\wedge y_k\wedge y_j)y_i\wedge y_k\\
&= \sum_{i<k}\sum_j y_j\otimes(2\iota(y_k)\iota(y_i)v,y_j)y_i\wedge y_j\\
&= \sum_{i<k}\sum_j(2\iota(y_k)\iota(y_i)v,y_j)y_j\otimes y_i\wedge y_j\\
&= \sum_{i<k}2\iota(y_k)\iota(y_i)v\otimes y_i\wedge y_k,
\end{aligned}$$

completing the proof.

In the case where $\mathbf{r}$ and $\mathbf{g}$ are reductive we have the alternative formula

$$\not K^2 = \mathrm{Cas}_{\mathbf{g}}\otimes 1-\mathrm{diag}(\mathrm{Cas}_{\mathbf{r}})+((\rho_{\mathbf{g}},\rho_{\mathbf{g}})-(\rho_{\mathbf{r}},\rho_{\mathbf{r}}))\,1\otimes 1. \tag{10.18}$$

Suppose that λ is the highest weight of a finite dimensional irreducible representation V_λ of $\mathbf{g}$ so that we get a surjective homomorphism

$$U(\mathbf{g}) \to \mathrm{End}(V_\lambda)$$

and hence a corresponding homomorphism

$$U(\mathbf{g}) \otimes C(\mathbf{p}) \to \mathrm{End}(V_\lambda) \otimes C(\mathbf{p}).$$

Also let diag_λ denote the composition of this homomorphism with

$$\mathrm{diag} : U(\mathbf{r}) \to \mathbf{U}(\mathbf{g}) \otimes \mathbf{C}(\mathbf{p}).$$

Then from the value of the Casimir (7.8) and (10.18) we get

$$\not{K}_\lambda^2 = ((\lambda + \rho_{\mathbf{g}}, \lambda + \rho_{\mathbf{g}}) - (\rho_{\mathbf{r}}, \rho_{\mathbf{r}}))\, 1 \otimes 1 - \mathrm{diag}_\lambda(\mathrm{Cas}_{\mathbf{r}}). \tag{10.19}$$

10.6 Eigenvalues of the Dirac operator.

We consider the situation where $\mathbf{g} = \mathbf{r} \oplus \mathbf{p}$ is a Lie algebra with invariant symmetric bilinear form, where $\mathbf{r}$ has the same rank as $\mathbf{g}$, and where we have chosen a common Cartan subalgebra

$$\mathbf{h} \subset \mathbf{r} \subset \mathbf{g}.$$

We let ℓ denote the dimension of $\mathbf{h}$, i.e. the common rank of $\mathbf{r}$ and $\mathbf{g}$. We let $\Phi = \Phi_{\mathbf{g}}$ denote the set of roots of $\mathbf{g}$, let $W = W_{\mathbf{g}}$ denote the Weyl group of $\mathbf{g}$, and let $W_{\mathbf{r}}$ denote the Weyl group of $\mathbf{r}$ so that

$$W_{\mathbf{r}} \subset W$$

and we let c denote the index of $W_{\mathbf{r}}$ in W.

A choice of positive roots Φ^+ for $\mathbf{g}$ amounts to a choice of a Borel subalgebra $\mathbf{b}$ of $\mathbf{g}$ and then $\mathbf{b} \cap \mathbf{r}$ is a Borel subalgebra of $\mathbf{r}$, which picks out a system of positive roots $\Phi_{\mathbf{r}}^+$ for $\mathbf{r}$ and then

$$\Phi_{\mathbf{r}}^+ \subset \Phi^+.$$

The corresponding Weyl chambers are

$$D = D_{\mathbf{g}} = \{\lambda \in \mathbf{h}_{\mathbf{R}}^* | (\lambda, \phi) \geq 0 \ \ \forall \phi \in \Phi^+\}$$

and

$$D_{\mathbf{r}} = \{\lambda \in \mathbf{h}_{\mathbf{R}}^* | (\lambda, \phi) \geq 0 \ \ \forall \phi \in \Phi_{\mathbf{r}}^+\}$$

so

$$D \subset D_{\mathbf{r}}$$

and we have chosen a cross-section C of $W_{\mathbf{r}}$ in W as

$$C = \{w \in W | wD \subset D_{\mathbf{r}}\},$$

so

$$W = W_{\mathbf{r}} \cdot C, \quad D_{\mathbf{r}} = \bigcup_{w \in C} wD.$$

We let $\mathbf{L} = \mathbf{L}_{\mathbf{g}} \subset \mathbf{h}^*_{\mathbf{R}}$ denote the lattice of $\mathbf{g}$ integral linear forms on $\mathbf{h}$, i.e.

$$\mathbf{L} = \{\mu \in \mathbf{h}^* | 2\frac{(\mu, \phi)}{(\phi, \phi)} \in \mathbf{Z} \ \forall \phi \in \Delta\}.$$

We let

$$\rho = \rho_{\mathbf{g}} = \frac{1}{2} \sum_{\phi \in \Delta^+} \phi$$

and

$$\rho_{\mathbf{r}} = \frac{1}{2} \sum_{\phi \in \Delta^+_{\mathbf{r}}} \phi.$$

We set

$$\mathbf{L}_{\mathbf{r}} = \text{ the lattice spanned by } \mathbf{L} \text{ and } \rho_{\mathbf{r}},$$

and

$$\Lambda := \mathbf{L} \cap D, \qquad \Lambda_{\mathbf{r}} := \mathbf{L}_{\mathbf{r}} \cap D_{\mathbf{r}}.$$

For any $\mathbf{r}$ module Z we let $\Gamma(Z)$ denote its set of weights, and we shall assume that

$$\Gamma(Z) \subset \mathbf{L}_{\mathbf{r}}.$$

For such a representation define

$$m_Z := \max_{\gamma \in \Gamma(Z)} (\gamma + \rho_{\mathbf{r}}, \gamma + \rho_{\mathbf{r}}). \tag{10.20}$$

For any $\mu \in \Lambda_{\mathbf{r}}$ we let Z_μ denote the irreducible module with highest weight μ.

Proposition 30 *Let*

$$\Gamma_{\max}(Z) := \{\mu \in \Gamma(Z) | (\mu + \rho_{\mathbf{r}}, \mu + \rho_{\mathbf{r}}) = m_Z\}.$$

Let $\mu \in \Gamma_{\max}(Z)$. *Then*

1. $\mu \in \Lambda_{\mathbf{r}}$.
2. *If* $z \neq 0$ *is a weight vector with weight* μ *then* z *is a highest weight vector, and hence the submodule* $U(\mathbf{r})z$ *is irreducible and equivalent to* Z_μ.
3. *Let*

$$Y_{\max} := \sum_{\mu \in \Gamma_{\max}(Z)} Z_\mu$$

and

$$Y := U(\mathbf{r}) Y_{\max}.$$

Then $m_Z - (\rho_{\mathbf{r}}, \rho_{\mathbf{r}})$ *is the maximal eigenvalue of* $\mathrm{Cas}_{\mathbf{r}}$ *on* Z *and* Y *is the corresponding eigenspace.*

Proof. We first show that

$$\mu \in \Gamma_{\max} \Rightarrow \ \mu + \rho_{\mathbf{r}} \in \Lambda_{\mathbf{r}}.$$

Suppose not, so there exists a $w \neq 1,\ w \in W_{\mathbf{r}}$ such that

$$w\mu + w\rho_{\mathbf{r}} \in \Lambda_{\mathbf{r}}.$$

But w changes the sign of some of the positive roots (the number of such changes being equal the length of w in terms of the generating reflections), and so $\rho_{\mathbf{r}} - w\rho_{\mathbf{r}}$ is a non-trivial sum of positive roots. Therefore

$$(w\mu + w\rho_{\mathbf{r}}, \rho_{\mathbf{r}} - w\rho_{\mathbf{r}}) \geq 0, \quad (\rho_{\mathbf{r}} - w\rho_{\mathbf{r}}, \rho_{\mathbf{r}} - w\rho_{\mathbf{r}}) > 0$$

and

$$w\mu + \rho_{\mathbf{r}} = (w\mu + w\rho_{\mathbf{r}}) + (\rho_{\mathbf{r}} - w\rho_{\mathbf{r}})$$

satisfies

$$(w\mu + \rho_{\mathbf{r}}, w\mu + \rho_{\mathbf{r}}) > (w\mu + w\rho_{\mathbf{r}}, w\mu + w\rho_{\mathbf{r}}) = (\mu + \rho_{\mathbf{r}}, \mu + \rho) = m_Z$$

contradicting the definition of m_Z. Now suppose that z is a weight vector with weight μ which is not a highest weight vector. Then there will be some irreducible component of Z containing z and having some weight μ' such that $\mu' - \mu$ is a non trivial sum of positive roots. We have

$$\mu' + \rho_{\mathbf{r}} = (\mu' - \mu) + (\mu + \rho_{\mathbf{r}})$$

so by the same argument we conclude that

$$(\mu' + \rho_{\mathbf{r}}, \mu' + \rho_{\mathbf{r}}) > m_Z$$

since $\mu + \rho_{\mathbf{r}} \in \Lambda_{\mathbf{r}}$, and again this is impossible. Hence z is a highest weight vector implying that $\mu \in \Lambda_{\mathbf{r}}$. This proves 1) and 2).

We have already verified that the eigenvalue of the Casimir $\mathrm{Cas}_{\mathbf{r}}$ on any Z_γ is $(\gamma + \rho_{\mathbf{r}}, \gamma + \rho_{\mathbf{r}}) - (\rho_{\mathbf{r}}, \rho_{\mathbf{r}})$. This proves 3).

Consider the irreducible representation V_ρ of $\mathbf{g}$ corresponding to $\rho = \rho_{\mathbf{g}}$. By the same arguments, any weight $\gamma \neq \rho$ of V_ρ lying in D must satisfy $(\gamma, \gamma) < (\rho, \rho)$ and hence any weight γ of V_ρ satisfying $(\gamma, \gamma) = (\rho, \rho)$ must be of the form

$$\gamma = w\rho$$

for a unique $w \in W$. But

$$w\rho = \rho - \sum_{\phi \in J_w} \phi = \rho - \phi_J$$

where

$$J_w := w(-\Phi^+) \cap \Phi^+.$$

We know that all the weights of V_ρ are of the form $\rho - \phi_J$ as J ranges over all subsets of Φ^+. So

$$(\rho, \rho) \geq (\rho - \phi_J, \rho - \phi_J) \tag{10.21}$$

where we have strict inequality unless $J = J_w$ for some $w \in W$.

Now let $\lambda \in \Lambda$, let V_λ be the corresponding irreducible module with highest weight λ and let γ be a weight of V_λ. As usual, let J denote a subset of the positive roots, $J \subset \Phi^+$. We claim that

Proposition 31 *We have*

$$(\lambda + \rho, \lambda + \rho) \geq (\gamma + \rho - \phi_J, \gamma + \rho - \phi_J) \tag{10.22}$$

with strict inequality unless there exists a $w \in W$ such that

$$\gamma = w\lambda, \quad \textit{and} \;\; J = J_w$$

in which case the w is unique.

Proof. Choose w such that

$$w^{-1}(\gamma + \rho - \phi_J) \in \Lambda.$$

Since $w^{-1}(\gamma)$ is a weight of V_λ, $\lambda - w^{-1}(\gamma)$ is a sum (possibly empty) of positive roots. Also $w^{-1}(\rho - \phi_J)$ is a weight of V_ρ and hence $\rho - w^{-1}(\rho - \phi_J)$ is a sum (possibly empty) of positive roots. Since

$$\lambda + \rho = (\lambda - w^{-1}\gamma) + (\rho - w^{-1}(\rho - \phi_J) + w^{-1}(\gamma + \rho - \phi_J)),$$

we conclude that

$$(\lambda + \rho, \lambda + \rho) \geq (w^{-1}(\gamma + \rho - \phi_J), w^{-1}(\gamma + \rho - \phi_J)) = (\gamma + \rho - \phi_J, \gamma + \rho - \phi_J)$$

with strict inequality unless $\lambda - w^{-1}(\gamma) = 0 = \rho - w^{-1}(\rho - \phi_J)$, and this last equality implies that $J = J_w$. QED

We have the spin representation Spin ν where $\nu : \mathbf{r} \to C(\mathbf{p})$. Call this module S. Consider

$$V_\lambda \otimes S$$

as a $\mathbf{r}$ module. Then, letting γ denote a weight of V_λ, we have

$$\Gamma(V_\lambda \otimes S) = \{\mu = \gamma + \rho_{\mathbf{p}} - \phi_J\} \tag{10.23}$$

where

$$\rho_{\mathbf{p}} = \frac{1}{2} \sum_{J \in \Phi^+_{\mathbf{p}}} \phi, \quad \Phi^+_{\mathbf{p}} := \Phi^+ / \Phi^+_{\mathbf{r}}.$$

In other words, $\Phi_{\mathbf{p}}$ are the roots of $\mathbf{g}$ which are not roots of $\mathbf{r}$, or, put another way, they are the weights of $\mathbf{p}$ considered as a $\mathbf{r}$ module. (Our equal rank assumption says that 0 does not occur as one of these weights.) For the weights μ of $V_\lambda \otimes S$ the form (10.23) gives

$$\mu + \rho_{\mathbf{r}} = \gamma + \rho - \phi_J, \quad J \subset \Delta^+_{\mathbf{p}}.$$

So if we set $Z = V_\lambda \otimes S$ as a $\mathbf{r}$ module, (10.22) says that

$$(\lambda + \rho, \lambda + \rho) \geq m_Z.$$

But we may take $J = \emptyset$ as one of our weights showing that

$$m_Z = (\lambda + \rho_{\mathbf{g}}, \lambda + \rho_{\mathbf{g}}). \tag{10.24}$$

To determine $\Gamma_{\max}(Z)$ as in Prop. 30 we again use Prop.31 and (10.23): A $\mu = \gamma + \rho_{\mathbf{p}} - \phi_J$ belongs to $\Gamma_{\max}(Z)$ if and only if $\gamma = w\lambda$ and $J = J_w$. But then

$$\rho_{\mathbf{g}} - \phi_J = w\rho_{\mathbf{g}}.$$

Since $\rho_{\mathbf{g}} = \rho_{\mathbf{r}} + \rho_{\mathbf{p}}$ we see from the form (10.23) that

$$\mu + \rho_{\mathbf{r}} = w(\lambda + \rho_{\mathbf{g}}) \tag{10.25}$$

where w is unique, and

$$J_w \subset \Phi^+_{\mathbf{p}}.$$

We claim that this condition is the same as the condition $w(D) \subset D_{\mathbf{r}}$ defining our cross-section, C. Indeed, $w \in C$ if and only if $(\phi, w\rho_{\mathbf{g}}) > 0, \ \forall \phi \in \Phi^+_{\mathbf{r}}$. But $(\phi, w\rho) = (w^{-1}\phi, \rho) > 0$ if and only if $\phi \in w(\Phi^+)$. Since $J_w = w(-\Phi^+) \cap \Phi^+$, we see that $J_w \subset \Phi^+_{\mathbf{p}}$ is equivalent to the condition $w \in C$.

Now for $\mu \in \Gamma_{\max}(Z)$ we have

$$\mu = w(\lambda + \rho) - \rho_{\mathbf{r}} =: w \bullet \lambda \tag{10.26}$$

where $\gamma = w(\lambda)$ and so has multiplicity one in V_λ.

Furthermore, we claim that the weight $\rho_{\mathbf{p}} - \phi_{J_w}$ has multiplicity one in S. Indeed, consider the representation

$$Z_{\rho_{\mathbf{r}}} \otimes S$$

of $\mathbf{r}$. It has the weight $\rho = \rho_{\mathbf{r}} + \rho_{\mathbf{p}}$ as a highest weight, and in fact, all of the weights of $V_{\rho_{\mathbf{g}}}$ occur among its weights. Hence, on dimensional grounds, say from the Weyl character formula, we conclude that it coincides, as a representation of $\mathbf{r}$, with the restriction of the representation $V_{\rho_{\mathbf{g}}}$ to $\mathbf{r}$. But since $\rho_{\mathbf{g}} - \phi_{J_w} = w\rho_{\mathbf{g}}$ has multiplicity one in $V_{\rho_{\mathbf{g}}}$, we conclude that $\rho_{\mathbf{p}} - \phi_{J_w}$ has multiplicity one in S.

We have proved that each of the $w \bullet \lambda$ have multiplicity one in $V_\lambda \otimes S$ with corresponding weight vectors

$$z_{w\bullet\lambda} := v_{w\lambda} \otimes e_-^{J_w} e^+.$$

So each of the submodules

$$Z_{w\bullet\lambda} := U(\mathbf{r})z_{w\bullet\lambda} \tag{10.27}$$

occurs with multiplicity one in $V_\lambda \otimes S$. The length of $w \in C$ (in terms of the simple reflections of W determined by Δ) is the number of positive roots changed into negative roots, i.e. the cardinality of J_w. This cardinality is the sign of $\det w$ and also determines whether $e_-^J e_+$ belongs to S_+ or to S_-. From Prop.31 and equation (10.24) we know that the maximum eigenvalue of $\mathrm{Cas}_\mathbf{r}$ on $V_\lambda \otimes S$ is

$$(\lambda + \rho_\mathbf{g}, \lambda + \rho_\mathbf{g}) - (\rho_\mathbf{r}, \rho_\mathbf{r}).$$

Now $\not{K}_\lambda \in \mathrm{End}(V_\lambda \otimes S)$ commutes with the action of $\mathbf{r}$ with

$$\not{K}_\lambda : \begin{array}{c} V_\lambda \otimes S_+ \to V_\lambda \otimes S_- \\ \\ V_\lambda \otimes S_- \to V_\lambda \otimes S_+. \end{array}$$

Furthermore, by (10.19), the kernel of $\not{K}_\lambda^2$ is the eigenspace of $\mathrm{Cas}_\mathbf{r}$ corresponding to the eigenvalue $(\lambda + \rho, \lambda + \rho) - (\rho_\mathbf{r}, \rho_\mathbf{r})$. Thus

$$\mathrm{Ker}(\not{K}_\lambda^2) = \sum_{w \in C} Z_{w\bullet\lambda}.$$

Each of these modules lies either in $V \otimes S_+$ or $V \otimes S_-$, one or the other but not both. Hence

$$\mathrm{Ker}(\not{K}_\lambda^2) = \mathrm{Ker}(\not{K}_\lambda)$$

and so

$$\mathrm{Ker}(\not{K}_\lambda)|_{V_\lambda \otimes S_+} = \sum_{w \in C,\ \det w = 1} Z_{w\bullet\lambda} \tag{10.28}$$

and

$$\mathrm{Ker}(\not{K}_\lambda)|_{V_\lambda \otimes S_-} = \sum_{w \in C,\ \det w = -1} Z_{w\bullet\lambda} \tag{10.29}$$

Let

$$K_\pm := \sum_{w \in C,\ \det w = \pm 1} Z_{w\bullet\lambda}. \tag{10.30}$$

It follows from (10.28) that $\not{K}_\lambda$ induces an injection of

$$(V_\lambda \otimes S_+)/K_+ \to V \otimes S_-$$

which we can follow by the projection

$$V_\lambda \otimes S_- \to (V_\lambda \otimes S_-)/K_-.$$

Hence $\not{K}_\lambda$ induces a bijection

$$\tilde{\not{K}}_\lambda : (V \otimes S_+)/K_+ \to (V_\lambda \otimes S_-)/K_-. \tag{10.31}$$

In short, we have proved that the sequence

$$0 \to K_+ \to V_\lambda \otimes S_+ \to V_\lambda \otimes S_- \to K_- \to 0 \tag{10.32}$$

is exact in a very precise sense, where the middle map is the Kostant Dirac operator: each summand of K_+ occurs exactly once in $V_\lambda \otimes S_+$ and similarly for K_-. This gives a much more precise statement of Theorem 16 and a completely different proof.

10.7 The geometric index theorem.

Let r be the representation of G on the space $\mathcal{F}(G)$ of smooth or on $L^2(G)$ of L^2 functions on G coming from right multiplication. Thus

$$[r(g)f](a) = f(ag).$$

Then $\not{K}$ acts on $\mathcal{F}(G) \otimes S$ or on $L^2(G) \otimes S$ and centralizes the action of $\operatorname{diag} \mathbf{r}$. If U is a module for R, we may consider $\mathcal{F}(G) \otimes S \otimes U$ or $L^2(G) \otimes S \otimes U$, and $\not{K} \otimes 1$ commutes with $\operatorname{diag} \mathbf{r} \otimes 1$ and with the action ρ of R on U, i.e with $1 \otimes 1 \otimes \rho$. If R is connected, this implies that $\not{K}$ commutes with the diagonal action of $\tilde{R}$, the universal cover of R, on $\mathcal{F} \otimes S \otimes U$ or $L^2(G) \otimes S \otimes U$ given by

$$k \mapsto r(k) \otimes \operatorname{Spin}(k) \otimes \rho(k), \qquad k \in R$$

where $\operatorname{Spin} : \tilde{R} \to \operatorname{Spin}(\mathbf{p})$ is the group homomorphism corresponding to the Lie algebra homomorphism ν. If G/R is a spin manifold, the invariants under this R action correspond to smooth or L^2 sections of $\mathbf{S} \otimes \mathcal{U}$ where $\mathbf{S}$ is the spin bundle of G/R and $\mathcal{U}$ is the vector bundle on G/R corresponding to U. Thus $\not{K}$ descends (by restriction) to a differential operator $\not{\partial}$ on G/R and we shall compute its G-index for irreducible U. The key result, due to Landweber, asserts that if U belongs to a multiplet coming from an irreducible V of G, then this index is, up to a sign, equal to V. If U does not belong to a multiplet, then this index is zero. We begin with some preliminary results due to Bott.

10.7.1 The index of equivariant Fredholm maps.

Let E and F be Hilbert spaces which are unitary modules for the compact Lie group G. Suppose that

$$E = \widehat{\bigoplus_n} E_n, \quad F = \widehat{\bigoplus_n} F_n$$

are completed direct sum decompositions into subspaces which are G-invariant and finite dimensional, and that

$$T : E \to F$$

is a Fredholm map (finite dimensional kernel and cokernel) such that

$$T(E_n) \subset F_n.$$

We write

$$\mathrm{Index}_G\, T = \mathrm{Ker}\, T - \mathrm{Coker}\, T$$

as an element of R(G), the ring of virtual representations of G. Thus $\mathrm{R}(G)$ is the space of finite linear combinations $\sum_\lambda a_\lambda V_\lambda, \quad a_\lambda \in \mathbf{Z}$ as V_λ ranges over the irreducible representations of G. (Here, and in what follows, we are regarding any finite dimensional representation of G as an element of $\mathrm{R}(G)$ by its decomposition into irreducibles, and similarly the difference of any two finite dimensional representations is an element of $\mathrm{R}(G)$.)

If we denote the restriction of T to E_n by T_n, then

$$\mathrm{Index}_G\, T = \sum \mathrm{Index}_G\, T_n$$

where all but a finite number of terms on the right vanish. For each n we have the exact sequence

$$0 \to \mathrm{Ker}\, T_n \to E_n \to F_n \to \mathrm{Coker}\, T_n \to 0.$$

Thus

$$\mathrm{Index}_G\, T_n = E_n - F_n$$

as elements of $\mathrm{R}(G)$. Therefore we can write

$$\mathrm{Index}_G\, T = \sum (E_n - F_n) \tag{10.33}$$

in $\mathrm{R}(G)$, where all but a finite number of terms on the right vanish. We shall refer to this as Bott's equation.

10.7.2 Induced representations and Bott's theorem.

Let R be a closed subgroup of G. Given any R-action ρ on a vector space U, we consider the associated vector bundle $G \times_R V$ over the homogeneous space G/R. The sections of this bundle are then equivariant U-valued functions on G satisfying $s(gk) = \rho(k)^{-1}s(g)$ for all $k \in R$. Applying the Peter-Weyl theorem, we can decompose the space of L^2 maps from G to U into a sum over the irreducible representations V_λ of G,

$$L^2(G) \otimes U \cong \widehat{\bigoplus}_\lambda V_\lambda \otimes V_\lambda^* \otimes U,$$

with respect to the $G \times G \times R$ action $l \otimes r \otimes \rho$. The R-equivariance condition is equivalent to requiring that the functions be invariant under the diagonal R-action $k \mapsto r(k) \otimes \rho(k)$. Restricting the Peter-Weyl decomposition above to the R invariant subspace, we obtain

$$\begin{aligned} L^2(G \times_R U) \quad &\cong \widehat{\bigoplus}_\lambda V_\lambda \otimes (V_\lambda^* \otimes U)^R \\ &\cong \widehat{\bigoplus}_\lambda V_\lambda \otimes \mathrm{Hom}_R(V_\lambda, U). \end{aligned} \tag{10.34}$$

The Lie group G acts on the space of sections by $l(g)$, the left action of G on functions, which is preserved by this construction. The space $L^2(G \times_H U)$ is thus an infinite dimensional representation of G.

The intertwining number of two representations gives us an inner product

$$\langle V, W \rangle_G = \dim_{\mathbf{C}} \mathrm{Hom}_G(V, W)$$

on $\mathrm{R}(G)$, with respect to which the irreducible representations of G form an orthonormal basis. Taking the formal completion of $\mathrm{R}(G)$, we define $\hat{\mathrm{R}}(G)$ to be the space of possibly infinite formal sums $\sum_\lambda a_\lambda V_\lambda$. The intertwining number then extends to a pairing $\mathrm{R}(G) \times \hat{\mathrm{R}}(G) \to \mathbf{Z}$.

If R is a subgroup of G, every representation of G automatically restricts to a representation of R. This gives us a pullback map $i^* : \mathrm{R}(G) \to \mathrm{R}(R)$, corresponding to the inclusion $i : R \hookrightarrow G$. The map $U \mapsto L^2(G \times_H U)$ discussed above assigns to each R-representation an induced infinite dimensional G-representation. Expressed in terms of our representation ring notation, this induction map becomes the homomorphism $i_* : \mathrm{R}(R) \to \hat{\mathrm{R}}(G)$ given by

$$i_* U = \sum_\lambda \langle i^* V_\lambda, U \rangle_R V_\lambda,$$

the formal adjoint to the pullback i^*. This is the content of the Frobenius reciprocity theorem.

A homogeneous differential operator on G/R is a differential operator $D : \Gamma(\mathcal{E}) \to \Gamma(\mathcal{F})$ between two homogeneous vector bundles $\mathcal{E}$ and $\mathcal{F}$ that commutes with the left action of G on sections. If the operator is elliptic, then its kernel and cokernel are both finite dimensional representations of G, and thus its G-index is a virtual representation in $R(G)$. In this case, the index takes a particularly elegant form.

Theorem 18 (Bott) *If $D : \Gamma(G \times_H U_0) \to \Gamma(G \times_H U_1)$ is an elliptic homogeneous differential operator, then the G-equivariant index of D is given by*

$$\mathrm{Index}_G\, D = i_*(U_0 - U_1),$$

where $i_*(U_0 - U_1)$ is a finite element in $\hat{\mathrm{R}}(G)$, i.e. belongs to $\mathrm{R}(G)$.

In particular, note that the index of a homogeneous differential operator depends only on the vector bundles involved and not on the operator itself! To prove the theorem, just use Bott's formula (10.33), where the subscript n is replaced by λ labeling the the G-irreducibles.

10.7.3 Landweber's index theorem.

Suppose that G is semi-simple and simply connected and R is a reductive subgroup of maximal rank. Suppose further that G/R is a spin manifold, then we can compose the spin representation $S = S_+ \oplus S_-$ of $\mathrm{Spin}(\mathbf{p})$ with the lifted map Spin : $\tilde{R} \to \mathrm{Spin}(\mathbf{p})$ to obtain a homogeneous vector bundle, the spin bundle $\mathbf{S}$ over G/R. For any representation of R on U the Kostant Dirac operator descends to an operator

$$\partial\!\!\!/_U : \Gamma(\mathbf{S}_\pm \otimes \mathcal{U}) \to \Gamma(\mathbf{S}_\mp \otimes \mathcal{U}).$$

(This operator has the same symbol as the Dirac operator arising from the Levi-Civita connection on G/R twisted by $\mathcal{U}$, and has the same index by Bott's theorem. For the precise relation between this Dirac operator coming from $K\!\!\!/$ and the dirac operator coming from the Levi- civita connection we refer to Landweber's thesis.)

The following theorem of Landweber gives an expression for the index of this Kostant Dirac operator. In particular, if we consider G/T, where T is a maximal torus (which is always a spin manifold), this theorem becomes a version of the Borel-Weil-Bott theorem expressed in terms of spinors and the Dirac operator, instead of in its customary form involving holomorphic sections and Dolbeault cohomology.

Theorem 19 (Landweber) *Let G/R be a spin manifold, and let U_μ be an irreducible representation U_μ of R with highest weight μ. The G-equivariant index of the Dirac operator $\partial\!\!\!/_U$ is the virtual G-representation*

$$\mathrm{Index}_G\, \partial\!\!\!/_{U_\mu} = (-1)^{\dim \mathbf{p}/2}\ (-1)^w V_{w(\mu+\rho_H)-\rho_G} \tag{10.35}$$

if there exists an element $w \in W_G$ in the Weyl group of G such that the weight $w(\mu+\rho_H)-\rho_G$ is dominant for G. If no such w exists, then $\mathrm{Index}_G\, \partial\!\!\!/_{U_\mu} = 0$.

Proof. For any irreducible representation V_λ of G with highest weight λ we have

$$V_\lambda \otimes (S_+ - S_-) = \sum_{w\in C} (-1)^w U_{w\bullet\lambda}$$

by [GKRS]. Hence

$$\mathrm{Hom}_R(V_\lambda \otimes (S_+ - S_-), U_\mu) = 0$$

if $\mu \neq w \bullet \lambda$ for some $w \in C$ while

$$\mathrm{Hom}_R(V_\lambda \otimes (S_+ - S_-), U_\mu) = (-1)^w$$

if $\mu = w\bullet\lambda$. But, by (10.33) and Theorem 18 we have

$$\begin{aligned} \mathrm{Index}_G\, \partial\!\!\!/_U &= \widehat{\bigoplus}_\lambda V_\lambda \otimes (V_\lambda^* \otimes (S_+ - S_-) \otimes U_\mu)^R \\ &= \widehat{\bigoplus} \mathrm{Hom}_R(V_\lambda \otimes (S_+ - S_-)^*, U_\mu). \end{aligned}$$

Now $(S_+ - S_-)^* = S_+ - S_-$ if $\dim \mathbf{p} \cong 0 \mod(4)$ while $(S_+ - S_-)^* = S_- - S_+$ if $\dim \mathbf{p} \cong 2 \mod(4)$. Hence

$$\mathrm{Index}_G\, \partial\!\!\!/_{U_\mu} = (-1)^{\dim \mathbf{p}/2} \widehat{\bigoplus} \mathrm{Hom}_R(V_\lambda \otimes (S_+ - S_-), U_\mu). \tag{10.36}$$

The right hand side of (10.36) vanishes if μ does not belong to a multiplet, i.e is not of the form

$$w \bullet \lambda = w(\lambda + \rho_{\mathbf{g}}) - \rho_{\mathbf{r}}$$

for some λ. The condition $w \bullet \lambda = \mu$ can thus be written as

$$w^{-1}(\mu + \rho_{\mathbf{r}}) - \rho_{\mathbf{g}} = \lambda.$$

If this equation does hold, then we get the formula in the theorem (with w replaced by w^{-1} which has the same determinant). QED

In general, if G/R is not a spin manifold, then in order to obtain a similar result we must instead consider the operator

$$\bar{\partial}_{U_\mu} : \left(L^2(G) \otimes (S_\pm) \otimes U_\mu\right)^{\mathbf{r}} \to \left(L^2(G) \otimes (S_\mp) \otimes U_\mu\right)^{\mathbf{r}}$$

viewed as an operator on G, restricted to the space of $(S \otimes U_\mu)$-valued functions on G that are invariant under the diagonal $\mathbf{r}$-action $\varrho(Z) = \operatorname{diag}(Z) + \sigma(Z)$, where σ is the $\mathbf{r}$-action on U_μ. Note that if $S \otimes U_\mu$ is induced by a representation of the Lie group R, then this operator descends to a well-defined operator on G/R as before. In general, the G-equivariant index of this operator $\bar{\partial}_{U_\mu}$ is once again given by (10.35). To prove this, we note that Bott's identity (10.33) and his theorem continue to hold for the induction map $i_* : \mathrm{R}(\mathbf{r}) \to \hat{\mathrm{R}}(\mathbf{g})$ using the representation rings for the Lie algebras instead of the Lie groups. Working in the Lie algebra context, we no longer need concern ourselves with the topological obstructions occurring in the global Lie group picture. The rest of the proof of Theorem 19 continues unchanged.

Chapter 11

The center of $U(\mathbf{g})$.

The purpose of this chapter is to study the center of the universal enveloping algebra of a semi-simple Lie algebra $\mathbf{g}$. We have already made use of the (second order) Casimir element.

11.1 The Harish-Chandra isomorphism.

Let us return to the situation and notation of Section 7.3. We have the monomial basis

$$f_1^{i_1} \cdots f_m^{i_m} h_1^{j_1} \cdots h_\ell^{j_\ell} e_1^{k_1} \cdots e_m^{k_m}$$

of $U(\mathbf{g})$, the decomposition

$$U(\mathbf{g}) = U(\mathbf{h}) \oplus (U(\mathbf{g})n_+ + n_- U(\mathbf{g}))$$

and the projection

$$\gamma : U(\mathbf{g}) \to U(\mathbf{h})$$

onto the first factor of this decomposition. This projection restricts to a projection, also denoted by γ

$$\gamma : Z(\mathbf{g}) \to U(\mathbf{h}).$$

The projection $\gamma : Z(\mathbf{g}) \to U(\mathbf{h})$ is a bit awkward. However Harish-Chandra showed that by making a slight modification in γ we get an isomorphism of $Z(\mathbf{g})$ onto the ring of Weyl group invariants of $U(\mathbf{h}) = S(\mathbf{h})$. Harish-Chandra's modification is as follows: As usual, let

$$\rho := \frac{1}{2} \sum_{\alpha > 0} \alpha.$$

Recall that for each i, the reflection s_i sends $\alpha_i \mapsto -\alpha_i$ and permutes the remaining positive roots. Hence

$$s_i \rho = \rho - \alpha_i$$

But by definition,

$$s_i\rho = \rho - \langle \rho, \alpha_i \rangle \alpha_i$$

and so

$$\langle \rho, \alpha_i \rangle = 1$$

for all $i = 1, \dots, m$. So

$$\rho = \omega_1 + \cdots + \omega_m,$$

i.e. ρ is the sum of the fundamental weights.

11.1.1 Statement

Define

$$\sigma : \mathbf{h} \to U(\mathbf{h}), \quad \sigma(h) = h - \rho(h)1. \tag{11.1}$$

This is a linear map from $\mathbf{h}$ to the commutative algebra $U(\mathbf{h})$ and hence, by the universal property of $U(\mathbf{h})$, extends to an algebra homomorphism of $U(\mathbf{h})$ to itself which is clearly an automorphism. We will continue to denote this automorphism by σ. Set

$$\gamma_H := \sigma \circ \gamma.$$

Then Harish-Chandra's theorem asserts that

$$\gamma_H : Z(\mathbf{g}) \to U(\mathbf{h})^W$$

and is an isomorphism of algebras.

11.1.2 Example of $sl(2)$.

To see what is going on let's look at this simplest case. The Casimir of degree two is

$$\frac{1}{2}h^2 + ef + fe,$$

as can be seen from the definition. Or it can be checked directly that this element is in the center. It is not written in our standard form which requires that the f be on the left. But $ef = fe + [e, f] = fe + h$. So the way of writing this element in terms of the above basis is

$$\frac{1}{2}h^2 + h + 2fe,$$

and applying γ to it yields

$$\frac{1}{2}h^2 + h.$$

There is only one positive root and its value on h is 2, so $\rho(h) = 1$. Thus σ sends $\frac{1}{2}h^2 + h$ into

$$\frac{1}{2}(h-1)^2 + h - 1 = \frac{1}{2}h^2 - h + \frac{1}{2} + h - 1 = \frac{1}{2}(h^2 - 1).$$

The Weyl group in this case is just the identity together with the reflection $h \mapsto -h$, and the expression on the right is clearly Weyl group invariant.

11.1.3 Using Verma modules to prove that $\gamma_H : Z(\mathbf{g}) \to U(\mathbf{h})^W$.

Any $\mu \in \mathbf{h}^*$ can be thought of as a linear map of $\mathbf{h}$ into the commutative algebra, $\mathbf{C}$ and hence extends to an algebra homomorphism of $U(\mathbf{h})$ to $\mathbf{C}$. If we regard an element of $U(\mathbf{h}) = S(\mathbf{h})$ as a polynomial function on $\mathbf{h}^*$, then this homomorphism is just evaluation at μ.

From our definitions,

$$(\lambda - \rho)\gamma(z) = \gamma_H(z)(\lambda) \quad \forall z \in Z(\mathbf{g}).$$

Let us consider the Verma module $\mathrm{Verm}(\lambda - \rho)$ where we denote its highest weight vector by v_+. For any $z \in Z(\mathbf{g})$, we have $hzv_+ = zhv_+ = (\lambda - \rho)(h)zv_+$ and $e_i z v_+ = z e_i v_+ = 0$. So zv_+ is a highest weight vector with weight $\lambda - \rho$ and hence must be some multiple of v_+. Call this multiple $\varphi_\lambda(z)$. Then

$$z f_1^{i_1} \cdots f_m^{i_m} v_+ = f_1^{i_1} \cdots f_m^{i_m} \varphi_\lambda(z) v_+,$$

so z acts as scalar multiplication by $\varphi_\lambda(z)$ on all of $\mathrm{Verm}(\lambda - \rho)$. To see what this scalar is, observe that since $z - \gamma(z) \in U(\mathbf{g})n_+$, we see that z has the same action on v_+ as does $\gamma(z)$ which is multiplication by $(\lambda - \rho)(\gamma(z)) = \lambda(\gamma_H(z))$. In other words,

$$\varphi_\lambda(z) = \lambda(\gamma_H(z)) = \gamma_H(z)(\lambda).$$

Notice that in this argument we only used the fact that $\mathrm{Verm}(\lambda - \rho)$ is a cyclic highest weight module: If V is any cyclic highest weight module with highest weight $\mu - \rho$ then z acts as multiplication by $\varphi_\mu(z) = \mu(\gamma_H(z)) = \gamma_H(z)(\mu)$. We will use this observation in a moment.

Getting back to the case of $\mathrm{Verm}(\lambda - \rho)$, for a simple root $\alpha = \alpha_i$ let $m = m_i := \langle \lambda, \alpha_i \rangle$ and suppose that m is an integer. The element

$$f_i^m v_+ \in \mathrm{Verm}(\lambda - \rho)_\mu$$

where

$$\mu = \lambda - \rho - m\alpha = s_i\lambda - \rho.$$

Now from the point of view of the $sl(2)$ generated by e_i, f_i, the vector v_+ is a maximal weight vector with weight $m - 1$. Hence $e_i f_i^m v_+ = 0$. Since $[e_j, f_i] = 0, i \neq j$ we have $e_j f_i^m v_+ = 0$ as well. So $f_i^m v_+ \neq 0$ is a maximal weight vector with weight $s_i\lambda - \rho$. Call the highest weight module it generates, M. Then from M we see that

$$\varphi_{s_i\lambda}(z) = \varphi_\lambda(z).$$

Hence we have proved that

$$(\gamma_H(z))(w\lambda) = (\gamma_H(z))(\lambda) \quad \forall\ w \in W$$

if λ is dominant integral. But two polynomials which agree on all dominant integral weights must agree everywhere. We have shown that the image of γ lies in $S(\mathbf{h})^W$.

Furthermore, we have

$$z_1 z_2 - \gamma(z_1)\gamma(z_2) = z_1(z_2 - \gamma(z_2)) + \gamma(z_2)(z_1 - \gamma(z_2)) \in U(\mathbf{g})n_+.$$

So

$$\gamma(z_1 z_2) = \gamma(z_1)\gamma(z_2).$$

This says that γ is an algebra homomorphism, and since $\gamma_H = \sigma \circ \gamma$ where σ is an automorphism, we conclude that γ_H is a homomorphism of algebras.

Equally well, we can argue directly from the fact that $z \in Z(\mathbf{g})$ acts as multiplication by $\varphi_\lambda(z) = \gamma_H(z)(\lambda)$ on $\mathrm{Verm}(\lambda - \rho)$ that γ_H is an algebra homomorphism.

11.1.4 Outline of proof of bijectivity.

To complete the proof of Harish-Chandra's theorem we must prove that γ_H is a bijection. For this we will introduce some intermediate spaces and homomorphisms. Let $Y(\mathbf{g}) := S(\mathbf{g})^{\mathbf{g}}$ denote the subspace fixed by the adjoint representation (extended to the symmetric algebra by derivations). This is a subalgebra, and the filtration on $S(\mathbf{g})$ induces a filtration on $Y(\mathbf{g})$. We shall produce an isomorphism

$$f : Y(\mathbf{g}) \to S(\mathbf{h})^W.$$

We also have a linear space isomorphism of $U(\mathbf{g}) \to S(\mathbf{g})$ given by the symmetric embedding of elements of $S^k(\mathbf{g})$ into $U(\mathbf{g})$, and let s be the restriction of this to $Z(\mathbf{g})$. We shall see that $s : Z(\mathbf{g}) \to Y(\mathbf{g})$ is an isomorphism. Finally, define

$$S_k(\mathbf{g}) = S^0(\mathbf{g}) \oplus S^1(\mathbf{g}) \oplus \cdots \oplus S^k(\mathbf{g})$$

so as to get a filtration on $S(\mathbf{g})$. This induces a filtration on $S(\mathbf{h}) \subset S(\mathbf{g})$. We shall show that for any $z \in U_k(\mathbf{g}) \cap Z(\mathbf{g})$ we have

$$(f \circ s)(z) \equiv \gamma_H(z) \mod S_{k-1}(\mathbf{g}).$$

This proves inductively that γ_H is an isomorphism since s and f are. Also, since σ does not change the highest order component of an element in $S(\mathbf{h})$, it will be enough to prove that for $z \in U_k(\mathbf{g}) \cap Z(\mathbf{g})$ we have

$$(f \circ s)(z) \equiv \gamma(z) \mod S_{k-1}(\mathbf{g}). \tag{11.2}$$

We now proceed to the details.

11.1.5 Restriction from $S(\mathbf{g}^*)^{\mathbf{g}}$ to $S(\mathbf{h}^*)^W$.

We first discuss polynomials on $\mathbf{g}$ — that is elements of $S(\mathbf{g}^*)$. Let τ be a finite dimensional representation of of $\mathbf{g}$, and consider the symmetric function F of degree k on $\mathbf{g}$ given by

$$(X_1,\ldots,X_k)\mapsto\sum \operatorname{tr}\ (\tau(X_{\pi 1})\cdots\tau(X_{\pi k}))$$

where the sum is over all permutations. For any $Y\in\mathbf{g}$, by definition,

$$YF(X_1,\ldots,X_k)=F([Y,X_1],X_2,\ldots,X_k)+\cdots+F(X_1,\ldots,X_{k-1},[Y,X_k]).$$

Applied to the above

$$F(X_1,\ldots,X_k)=\operatorname{tr}\tau(X_1)\cdots\tau(X_k)$$

we get

$$\begin{aligned}\operatorname{tr}\tau(Y)\tau(X_1)\cdots\tau(X_k)-\operatorname{tr}\tau(X_1)\tau(Y)\cdots\tau(X_k)\\ +\operatorname{tr}\tau(X_1)\tau(X)\tau(X_2\cdots\tau(X_k)-\cdots\\ =\operatorname{tr}\tau(Y)\tau(X_1)\cdots\tau(X_k)-\operatorname{tr}\tau(X_1)\cdots\tau(X_k)\tau(Y)=0.\end{aligned}$$

In other words, the function

$$X\mapsto \operatorname{tr}\ \tau(X)^n$$

belongs to $S(\mathbf{g}^*)^{\mathbf{g}}$. Now since $\mathbf{h}$ is a subspace of $\mathbf{g}$, the restriction map induces a homomorphism,

$$r:S(\mathbf{g}^*)\to S(\mathbf{h}^*).$$

If $F\in S(\mathbf{g}^*)^{\mathbf{g}}$, then, as a function on $\mathbf{g}$ it is invariant under the automorphism $\tau_i:=(\exp\operatorname{ad}e_i)(\exp\operatorname{ad}-f_i)(\exp\operatorname{ad}e_i)$ and hence

$$r:S(\mathbf{g}^*)^{\mathbf{g}}\to S(\mathbf{h}^*)^W.$$

If $F\in S(\mathbf{g}^*)^{\mathbf{g}}$ is such that its restriction to $\mathbf{h}$ vanishes, then its value at any element which is conjugate to an element of $\mathbf{h}$ (under $\mathcal{E}(\mathbf{g})$ the subgroup of automorphisms of $\mathbf{g}$ generated by the τ_i) must also vanish. But these include a dense set in $\mathbf{g}$, so F, being continuous, must vanish everywhere. So the restriction of r to $S(\mathbf{g}^*)^{\mathbf{g}}$ is injective.

To prove that it is surjective, it is enough to prove that $S(\mathbf{h}^*)^W$ is spanned by all functions of the form $X\mapsto\operatorname{tr}\tau(X)^k$ as τ ranges over all finite dimensional representations and k ranges over all non-negative integers. Now the powers of any set of spanning elements of $\mathbf{h}^*$ span $S(\mathbf{h}^*)$. So we can write any element of $S(\mathbf{h}^*)^W$ as a linear combination of the $\mathcal{A}\lambda^k$ where $\mathcal{A}$ denotes averaging over W. So it is enough to show that for any dominant weight λ, we can express λ^k in terms of $\operatorname{tr}\tau^k$. Let E_λ denote the (finite) set of all dominant weights $\prec\lambda$. Let τ denote the finite dimensional representation with highest weight λ. Then $\operatorname{tr}\tau(X)^k-\mathcal{A}\lambda^k(X)$ is a combination of $\mathcal{A}\mu(X)^k$ where $\mu\in E_\lambda$. Hence by induction on the finite set E_λ we get the desired result. In short, we have proved that

$$r:S(\mathbf{g}^*)^{\mathbf{g}}\to S(\mathbf{h}^*)^W$$

is bijective.

11.1.6 From $S(\mathbf{g})^{\mathbf{g}}$ to $S(\mathbf{h})^W$.

Now we transfer all this information from $S(\mathbf{g}^*)$ to $S(\mathbf{g})$: Use the Killing form to identify $\mathbf{g}$ with $\mathbf{g}^*$ and hence get an isomorphism

$$\alpha : S(\mathbf{g}) \to S(\mathbf{g}^*).$$

Similarly, let

$$\beta : S(\mathbf{h}) \to S(\mathbf{h}^*)$$

be the isomorphism induced by the restriction of the Killing form to $\mathbf{h}$, which we know to be non-degenerate. Notice that β commutes with the action of the Weyl group. We can write

$$S(\mathbf{g}) = S(\mathbf{h}) + J$$

where J is the ideal in $S(\mathbf{g})$ generated by n_+ and n_-. Let

$$j : S(\mathbf{g}) \to S(\mathbf{h})$$

denote the homomorphism obtained by quotienting out by this ideal. We claim that the diagram

$$\begin{array}{ccc} S(\mathbf{g}) & \xrightarrow{\alpha} & S(\mathbf{g}^*) \\ j\big\downarrow & & \big\downarrow r \\ S(\mathbf{h})' & \xrightarrow[\beta]{} & S(\mathbf{h}^*) \end{array}$$

commutes. Indeed, since all maps are algebra homomorphisms, it is enough to check this on generators, that is on elements of $\mathbf{g}$. If $X \in \mathbf{g}$, then

$$\langle h, r\alpha(X)\rangle = \langle h, \alpha(X)\rangle = (h, X)$$

where the scalar product on the right is the Killing form. But since h is orthogonal under the Killing form to $n_+ + n_-$, we have

$$(h, X) = (h, jX) = \langle h, \beta(jX)\rangle. \quad QED$$

Upon restriction to the $\mathbf{g}$ and W-invariants, we have proved that the right hand column is a bijection, and hence so is the left hand column, since β is a W-module morphism. Recalling that we have defined $Y(\mathbf{g}) := S(\mathbf{g})^{\mathbf{g}}$, we have shown that the restriction of j to $Y(\mathbf{g})$ is an isomorphism, call it f:

$$f : Y(\mathbf{g}) \to S(\mathbf{h})^W.$$

11.1.7 Completion of the proof.

Now we have a canonical linear bijection of $S(\mathbf{g})$ with $U(\mathbf{g})$ which maps

$$S(\mathbf{g}) \ni X_1 \cdots X_r \mapsto \frac{1}{r!} \sum_{\pi \in \Sigma_r} X_{\pi 1} \cdots X_{\pi r},$$

where the multiplication on the left is in $S(\mathbf{g})$ and the multiplication on the right is in $U(\mathbf{g})$ and where Σ_r denotes the permutation group on r letters. This map is a $\mathbf{g}$ module morphism. In particular this map induces a bijection

$$s : Z(\mathbf{g}) \to Y(\mathbf{g}).$$

Our proof will be complete once we prove (11.2). This is a calculation: write

$$u_{AJB} := f^A h^J e^B$$

for our usual monomial basis, where the multiplication on the right is in the universal enveloping algebra. Let us also write

$$p_{AJB} := f^A h^J e^B = f_\alpha^{i_\alpha} \cdots h_1^{j_1} \cdots h_\ell^{j_\ell} \cdots e_\gamma^{k_\gamma} \in S(\mathbf{g})$$

where now the powers and multiplication are in $S(\mathbf{g})$. The image of u_{AJB} under the canonical isomorphism with of $U(\mathbf{g})$ with $S(\mathbf{g})$ will not be p_{AJB} in general, but will differ from p_{AJB} by a term of lower filtration degree. Now the projection $\gamma : U(\mathbf{g}) \to U(\mathbf{h})$ coming from the decomposition

$$U(\mathbf{g}) = U(\mathbf{h}) \oplus (n_- U(\mathbf{g}) + U(\mathbf{g}) n_+)$$

sends $u_{AJB} \mapsto 0$ unless $A = 0 = B$ and is the identity on u_{0J0}. Similarly,

$$j(p_{AJB}) = 0 \quad \text{unless} \quad A = 0 = B$$

and

$$j(p_{0J0}) = p_{0J0} = h^J.$$

These two facts complete the proof of (11.2). *QED*

11.2 Chevalley's theorem.

Harish Chandra's theorem says that the center of the universal enveloping algebra of a semi-simple Lie group is isomorphic to the ring of Weyl group invariants in the polyinomial algebra $S(\mathbf{h})$. Chevalley's theorem asserts that this ring is in fact a polynomial ring in ℓ generators where $\ell = \dim \mathbf{h}$. To prove Chevalley's theorem we need to call on some facts from field theory and from the representation theory of finite groups.

11.2.1 Transcendence degrees.

A field extension $L : K$ is *finitely generated* if there are elements $\alpha_1, \ldots, \alpha_n$ of L so that $L = K(\alpha_1, \ldots, \alpha_n)$. In other words, every element of L can be written as a rational expression in the $\alpha_1, \ldots, \alpha_n$.

Elements $t_1, \ldots, t_k$ of L are called (algebraically) *independent* (over K) if there is no non-trivial polynomial p with coefficients in K such that

$$p(t_1, \ldots, t_k) = 0.$$

Lemma 15 *If $L : K$ is finitely generated, then there exists an intermediate field M such that $M = K(\alpha_1, \ldots, \alpha_r)$ where the $\alpha_1, \ldots, \alpha_r$ are independent transcendental elements and $L : M$ is a finite extension (i.e. L has finite dimension over M as a vector space).*

Proof. We are assuming that $L = K(\beta_1, \ldots, \beta_q)$. If all the β_i are algebraic, then $L : K$ is a finite extension. Otherwise one of the β_i is transcendental. Call this α_1. If $L : K(\alpha_1)$ is a finite extension we are done. Otherwise one of the remaining β_i is transcendental over $K(\alpha_1)$. Call it α_2. So α_1, α_2 are independent. Proceed.

Lemma 16 *If there is another collection $\gamma_1 \ldots \gamma_s$ so that $L : K(\gamma_1, \ldots, \gamma_s)$ is finite then $r = s$. This common number is called the transcendence degree of L over K.*

Proof. If $s = 0$, then every element of L is algebraic, contradicting the assumption that the $\alpha_1, \ldots, \alpha_r$ are independent, unless $r = 0$. So we may assume that $s > 0$. Since $L : M$ is finite, there is a polynomial p such that

$$p(\gamma_1, \alpha_1, \ldots, \alpha_r) = 0.$$

This polynomial must contain at least one α, since γ_1 is transcendental. Renumber if necessary so that α_1 occurs in p. Then α_1 is algebraic over $K(\gamma_1, \alpha_2, \ldots, \alpha_r)$ and $L : K(\gamma_1, \alpha_2, \ldots, \alpha_r)$ is finite. Continuing this way we can successively replace αs by γs until we conclude that $L : K(\gamma_1, \ldots, \gamma_r)$ is finite. If $s > r$ then the γs are not algebraically independent. so $s \leq r$ and similarly $r \leq s$.

Notice that if $\alpha_1, \ldots, \alpha_n$ are algebraically independent then $K(\alpha_1, \ldots, \alpha_n)$ is isomorphic to the field of rational functions in n indeterminates $K(t_1, \ldots, t_n)$ since $K(\alpha_1, \ldots, \alpha_n) = K(\alpha_1, \ldots, \alpha_{n-1})(\alpha_n)$ by clearing denominators.

11.2.2 Symmetric polynomials.

The symmetric group S_n acts on $K(t_1, \ldots, t_n)$ by permuting the variables. The fixed field, F, contains all the symmetric polynomials, in particular the elementary symmetric polynomials $s_1, \ldots, s_n$ where s_r is the sum of all possible distinct products taken r at a time. Using the general theory of field extensions, we can conclude that

Proposition 32 $F = K(s_1, \ldots, s_n)$.

The strategy of the proof is to first show that the dimension of the extension $K(t_1, \ldots, t_n) : K(s_1, \ldots, s_n)$ is $\leq n!$ and then a basic theorem in Galois theory (which we shall recall) says that the dimension of $K : F$ equals the order of the group $G = S_n$ which is $n!$. Since $K(s_1, \ldots, s_n) \subset F$ this will imply the proposition. Consider the extensions

$$K(t_1, \ldots, t_n) \supset K(s_1, \ldots, s_n, t_n) \supset K(s_1, \ldots, s_n).$$

We have the equation $f(t_n) = 0$ where

$$f(t) := t^n - s_1 t^{n-1} + t^{n-2} s_2 \cdots + (-1)^n s_n.$$

This shows that the dimension of the field extension $K(s_1, \ldots, s_n, t_n) : K(s_1, \ldots, s_n)$ is $\leq n$. If we let $s'_1 \ldots s'_{n-1}$ denote the elementary symmetric functions in $n-1$ variables, we have

$$s_j = t_n s'_{j-1} + s'_j$$

so

$$K(s_1, \ldots, s_n, t_n) = K(t_n, s'_1, \ldots, s'_{n-1}).$$

By induction we may assume that

$$\begin{aligned} \dim K(t_1, \ldots, t_n) : K(s_1, \ldots, s_n, t_n) &= \dim K(t_n)(t_1, \ldots, t_{n-1}) : K(t_n)(s'_1, \ldots, s'_{n-1}) \\ &\leq (n-1)! \end{aligned}$$

proving that

$$\dim K(t_1, \ldots, t_n) : K(s_1, \ldots, s_n) \leq n!.$$

A fundamental theorem of Galois theory says

Theorem 20 *Let G be a finite subgroup of the group of automorphisms of the field L over the field K, and let F be the fixed field. Then*

$$dim[L : F] = \#G.$$

This theorem, whose proof we will recall in the next section, then completes the proof of the proposition. The proposition implies that every symmetric polynomial is a rational function of the elementary symmetric functions.

In fact, every symmetric polynomial is a *polynomial* in the elementary symmetric functions, giving a stronger result. This is proved as follows: put the lexicographic order on the set of $n-$tuples of integers, and therefor on the set of monomials; so $x_1^{i_1} \cdots x_n^{i_n}$ is greater than $x_1^{j_1} \cdots x_n^{j_n}$ in this ordering if $i_1 > j_1$ of $i_1 = j_1$ and $i_2 > j_2$ or etc. Any polynomial has a "leading monomial" the greatest monomial relative to this lexicographic order. The leading monomial of the product of polynomials is the product of their leading monomials. We shall prove our contention by induction on the order of the leading monomial. Notice that if p is a symmetric polynomial, then the exponents $i_1, \ldots, i_n$ of its leading term must satisfy

$$i_1 \geq i_2 \geq \cdots \geq i_n,$$

for otherwise the monomial obtained by switching two adjacent exponents (which occurs with the same coefficient in the symmetric polynomial, p) would be strictly higher in our lexicographic order. Suppose that the coefficient of this leading monomial is a. Then

$$q = a s_1^{i_1 - i_2} s_2^{i_2 - i_3} \cdots s_{n-1}^{i_{n-1} - i_n} s_n^{i_n}$$

has the same leading monomial with the same coefficient. Hence $p - q$ has a smaller leading monomial. QED

11.2.3 Fixed fields.

We now turn to the proof of the theorem of the previous section.

Lemma 17 *Every distinct set of monomorphisms of a field K into a field L are linearly independent over L.*

Let $\lambda_1, \dots, \lambda_n$ be distinct monomorphisms of $K \to L$. The assertion is that there can not exist $a_1, \dots, a_n \in L$ such that

$$a_1\lambda_1(x) + \cdots + a_n\lambda_n(x) \equiv 0 \quad \forall x \in K$$

unless all the $a_i = 0$. Assume the contrary, so that such an equation holds, and we may assume that none of the $a_i = 0$. Looking at all such possible equations, we may pick one which involves the fewest number of terms, and we may assume that this is the equation we are studying. In other words, no such equation holds with fewer terms. Since $\lambda_1 \neq \lambda_n$, there exists a $y \in K$ such that $\lambda_1(y) \neq \lambda_n(y)$ and in particular $y \neq 0$. Substituting yx for x gives

$$a_1\lambda_1(yx) + \cdots + a_n\lambda_n(yx) = 0$$

so

$$a_1\lambda_1(y)\lambda_1(x) + \cdots + a_n\lambda_n(y)\lambda_n(x) = 0$$

and multiplying our original equation by $\lambda_1(y)$ and subtracting gives

$$a_2(\lambda_1(y) - \lambda_2(y))\lambda_2(x) + \cdots + a_n(\lambda_n(y) - \lambda_1(y))\lambda_n(x) = 0$$

which is a non-trivial equation with fewer terms. Contradiction.

Let $n = \#G$, and let the elements of G be $g_1 = 1, \dots, g_n$. Suppose that dim $L : F = m < n$. Let $x_1, \dots, x_m$ be a basis of L over F. The system of equations

$$g_1(x_j)y_1 + \cdots + g_n(x_j)y_n = 0, \quad j = 1, \dots, m$$

has more unknowns than equations, and so we can find non-zero $y_1, \dots, y_n$ solving these equations. Any $b \in L$ can be expanded as

$$b = b_1x_1 + \cdots + b_mx_m, \quad b_i \in F,$$

and so

$$\begin{aligned} g_1(b)y_1 + \cdots + g_n(b)y_n &= \sum_j b_j[g_1(x_j)y_1 + \cdots + g_n(x_j)y_n] \\ &= 0 \end{aligned}$$

showing that the monomorphisms g_i are linearly dependent. This contradicts the lemma.

Suppose that dim $L : F > n$. Let $x_1, \dots, x_n, x_{n+1}$ be linearly independent over F, and find $y_1, \dots, y_{n+1} \in L$ not all zero solving the n equations

$$g_j(x_1)y_1 + \cdots + g_j(x_{n+1})y_{n+1} = 0, \quad j = 1, \dots, n.$$

Choose a solution with fewest possible non-zero y's and relabel so that the first are the non-vanishing ones, so the equations now read

$$g_j(x_1)y_1 + \cdots + g_j(x_r)y_r = 0, \quad j = 1, \ldots, n,$$

and no such equations hold with less than r y's. Applying $g \in G$ to the preceding equation gives

$$gg_j(x_1)g(y_1) + \cdots + gg_j(x_r)g(y_r) = 0.$$

But gg_j runs over all the elements of G, and so $g(y_1), \ldots, g(y_r)$ is a solution of our original equations. In other words we have

$$\begin{aligned} g_j(x_1)y_1 + \cdots + g_j(x_r)y_r &= 0 \text{ and} \\ g_j(x_1)g(y_1) + \cdots + g_j(x_r)g(y_r) &= 0. \end{aligned}$$

Multiplying the first equations by $g(y_1)$, the second by y_1 and subtracting, gives

$$g_j(x_2)[y_2g(y_1) - g(y_2)y_1)] + \cdots + g_j(x_r)[y_rg(y_1) - g(y_r)y_1] = 0,$$

a system with fewer $y's$. This can not happen unless the coefficients vanish, i.e.

$$y_ig(y_1) = y_1g(y_i)$$

or

$$y_iy_1^{-1} = g(y_iy_1^{-1}) \quad \forall g \in G.$$

This means that

$$y_iy_1^{-1} \in F.$$

Setting $z_i = y_i/y_1$ and $k = y_1$, we get the equation

$$x_1kz_1 + \cdots + x_rkz_r = 0$$

as the first of our system of equations. Dividing by k gives a linear relation among $x_1, \ldots, x_r$ contradicting the assumption that they are independent.

11.2.4 Invariants of finite groups.

Let G be a finite group acting on a vector space. It action on the symmetric algebra $S(V^*)$ which is the same as the algebra of polynomial functions on V by

$$(gf)(v) = f(g^{-1}v).$$

Let

$$R = S(V^*)^G$$

be the ring of invariants. Let $S = S(V^*)$ and L be the field of quotients of S, so that $L = K(t_1, \ldots, t_n)$ where $n = \dim V$. From the theorem on fixed fields, we know that the dimension of L as an extension of L^G is equal to the number of elements in G, in particular finite. So L^G has transcendence degree n over the ground field.

Clearly the field of fractions of R is contained in L^G. We claim that they coincide. Indeed, suppose that $p, q \in S, \quad p/q \in L^G$. Multiply the numerator and denominator by Πgp the product taken over all $g \in G,\ g \neq 1$. The new numerator is G invariant. Therefore so is the denominator, and we have expressed p/q as the quotient of two elements of R.

If the finite group G acts on a vector space, then averaging over the group, i.e. the map

$$E \ni f \mapsto f^\sharp := \frac{1}{\#G} \sum g \cdots f$$

is a projection onto the subspace of invariant elements:

$$\mathcal{A} : f \mapsto f^\sharp \qquad E \to E^G.$$

In particular, if E is finite dimensional,

$$\dim E^G = \operatorname{tr} \mathcal{A}. \tag{11.3}$$

We may apply the averaging operator to our (infinite dimensional) situation where $S = S(V^*)$ and $R = S^G$ in which case we have the additional obvious fact that

$$(pq)^\sharp = p^\sharp q \quad \forall\ p \in S, \quad q \in R.$$

Let $R^+ \subset R$ denote the subring of R consisting of elements with constant term zero. Let

$$I := SR^+$$

so that I is an ideal in S. By the Hilbert basis theorem (whose proof we recall in the next section) the ideal I is finitely generated, and hence, from any set of generators, we may choose a finite set of generators.

Theorem 21 *Let $f_1, \ldots, f_r$ be homogeneous elements of R^+ which generate I as an ideal of S. Then $f_1, \ldots, f_r$ together with 1 generate R as a K algebra. In particular, R is a finitely generated K algebra.*

Proof. We must show that any $f \in R$ can be expressed as a polynomial in the $f_1, \ldots, f_r$, and since every f is a sum of its homogeneous components, it is enough to do this for homogeneous f and we proceed by induction on its degree. The statement is obvious for degree zero. For positive degree, $f \in R \subset I$ so

$$f = s_1 f_1 + \cdots + s_r f_r, \quad s_i \in I$$

and since $f, f_1, \ldots, f_r$ are homogeneous, we may assume the s_i are homogeneous of degree $\deg f - \deg f_i$ since all other contributions must cancel. Now apply $\mathcal{A}$ to get

$$f = s_1^\sharp f_1 + \cdots + s_r^\sharp f_r.$$

The $s_i^\sharp$ lie in R and have lower homogeneous degree than f, and hence can be expressed as polynomials in $f_1, \ldots, f_r$. Hence so can f.

11.2.5 The Hilbert basis theorem.

A commutative ring is called *Noetherian* if any of the following equivalent conditions holds:

1. If $I_1 \subset I_2 \subset \cdots$ is an ascending chain of ideals then there is a k such that $I_k = I_{k+1} = I_{k+2} = \ldots$.

2. Every non-empty set of ideals has a maximal element with respect to inclusion.

3. Every ideal is finitely generated.

The Hilbert basis theorem asserts that if R is a Noetherian ring, then so is the polynomial ring $R[X]$. In particular, all ideals in $K[X_1, \ldots, X_n]$ are finitely generated.

Let I be an ideal in RX and for any positive integer k let $L_k(I) \subset R$ be defined by

$$L_k(I) := \{a_k \in R | \exists a_{k-1}, \ldots, a_1 \in R \text{ with } \sum_0^k a_j X^j \in I\}.$$

For each k, $L_k(I)$ is an ideal in R. Multiplying by X shows that

$$L_k(I) \subset L_{k+1}(I).$$

Hence these ideals stabilize. If $I \subset J$ and $L_k(I) = L_k(J)$ for all k, we claim that this implies that $I = J$. Indeed, suppose not, and choose a polynomial of smallest degree belonging to J but not to I, say this degree is k. Its leading coefficient belongs to $L_k(J)$ and can not belong to $L_k(I)$ because otherwise we could find a polynomial of smaller degree belonging to J and not to I.

Proof of the Hilbert basis theorem. Let

$$I_0 \subset I_1 \subset \cdots$$

be an ascending chain of ideals in $R[X]$. Consider the set of ideals $L_p(I_q)$. We can choose a maximal member. So for $k \geq p$ we have

$$L_k(I_j) = L_k(I_q) \quad \forall j \geq q.$$

For each of the finitely many values $j = 1, \ldots, p-1$, the ascending chains

$$L_i(I_0) \subset L_i(I_1) \subset \cdots$$

stabilizes. So we can find a large enough r (bigger than the finitely many large values needed to stabilize the various chains) so that

$$L_i(I_j) = L_i(I_r) \quad \forall j \geq r, \quad \forall i.$$

This shows that $I_j = Ir \ \forall j \geq r$.

11.2.6 Proof of Chevalley's theorem.

This says that if $K = \mathbf{R}$ and W is a finite subgroup of $O(V)$ generated by reflections, then its ring of invariants is a polynomial ring in n- generators, where $n = \dim V$. Without loss of generality we may assume that W acts effectively, i.e. no non-zero vector is fixed by all of W.

Let $f_1, \ldots, f_r$ be a minimal set of homogeneous generators. Suppose we could prove that they are algebraically independent. Since the transcendence degree of the quotient field of R is $n = \dim V$, we conclude that $r = n$. So the whole point is to prove that a minimal set of homogeneous generators must be algebraically independent - that there can not exist a non-zero polynomial $h = h(y_1, \ldots, y_r)$ so that

$$h(f_1, \ldots, f_r) = 0. \tag{11.4}$$

So we want to get a smaller set of generators assuming that such a relation exists. Let

$$d_1 := \deg f_1, \ldots, d_r := \deg f_r.$$

For any non-zero monomial

$$ay_1^{e_1} \cdots y_r^{f_r}$$

occurring in h the term

$$af_1^{r_1} \cdots f_r^{e_r}$$

we get by substituting f's for y's has degree

$$d = e_1 d_1 + \cdots e_r d_r$$

and hence we may throw away all monomials in h which do not satisfy this equation. Now set

$$h_i := \frac{\partial h}{\partial y_i}(f_1, \cdots, f_r)$$

so that $h_i \in R$ is homogeneous of degree $d - d_i$, and let J be the ideal in R generated by the h_i. Renumber $f_1, \ldots, f_r$ so that $h_1, \ldots, h_m$ is a minimal generating set for J. This means that

$$h_i = \sum_{j=1}^{m} g_{ij} h_j, \qquad g_{ij} \in R$$

for $i > m$ (if $m < r$; if $m = r$ we have no such equations). Once again, since the h_i are homogeneous of degree $d - d_i$ we may assume that each g_{ij} is homogeneous of degree $d_i - d_j$ by throwing away extraneous terms.

Now let us differentiate the equation (11.4) with respect to x_k, $k = 1, \ldots, n$ to obtain

$$\sum_{i=1}^{r} h_i \frac{\partial f_i}{\partial x_k} \qquad k = 1, \ldots, n$$

and substitute the above expressions for $h_i,\ i > m$ to get

$$\sum_{i=1}^{m} h_i \left(\frac{\partial f_i}{\partial x_k} + \sum_{j=m+1}^{r} g_{ji} \frac{\partial f_j}{\partial x_k} \right) \qquad k = 1, \ldots, n.$$

Set

$$p_i := \frac{\partial f_i}{\partial x_k} + \sum_{j=m+1}^{r} g_{ji} \frac{\partial f_j}{\partial x_k} \quad i = 1, \ldots, m$$

so that each p_i is homogeneous with

$$\deg p_i = d_i - 1$$

and we have the equation

$$h_1 p_1 + \cdots h_m p_m = 0. \tag{11.5}$$

We will prove that this implies that

$$p_1 \in I. \tag{11.6}$$

Assuming this for the moment, this means that

$$\frac{\partial f_1}{\partial x_k} + \sum_{j=m+1}^{r} g_{j1} \frac{\partial f_j}{\partial x_k} = sum_{i=1}^{r} f_i q_i$$

where $q_i \in S$. Multiply these equations by x_k and sum over k and apply Euler's formula for homogeneous polynomials

$$\sum x_k \frac{\partial f}{\partial x_k} = (\deg f) f.$$

We get

$$d_1 f_1 + \sum d_j g_{j1} f_j = \sum f_i r_i$$

with $degr_1 > 0$ if it is not zero. Once again, the left hand side is homogeneous of degree d_1 so we can throw away all terms on the right which are not of this degree because of cancellation. But this means that we throw away the term involving f_1, and we have expressed f_1 in terms of $f_2, \ldots, f_r$, contradicting our choice of $f_1, \ldots, f_r$ as a minimal generating set.

So the proof of Chevalley's theorem reduced to proving that (11.5) implies (11.6), and for this we must use the fact that W is generated by reflections, which we have not yet used. The desired implication is a consequence of the following

Proposition 33 *Let $h_1, \ldots, h_m \in R$ be homogeneous with h_1 not in the ideal of R generated by $h_2, \ldots, h_m$. Suppose that (11.5) holds with homogeneous elements $p_i \in S$. Then (11.6) holds.*

Notice that h_1 can not lie in the ideal of S generated $h_2, \dots h_m$ because we can apply the averaging operator to the equation

$$h_1 = k_2 h_2 + \cdots + k_m h_m \quad k_i \in S$$

to arrange that the same equation holds with k_i replaced by $k_i^\sharp \in R$.

We prove the proposition by induction on the degree of p_1. This must be positive, since $p_1 \neq 0$ constant would put h_1 in the ideal generated by the remaining h_i. Let s be a reflection in W and H its hyperplane of fixed vectors. Then

$$sp_i - p_i = 0 \quad \text{on } H.$$

Let ℓ be a non-zero linear function whose zero set is this hyperplane. With no loss of generality, we may assume that the last variable, x_n, occurs with non-zero coefficient in ℓ relative to some choice of orthogonal coordinates. In fact, by rotation, we can arrange (temporarily) that $\ell = x_n$. Expanding out the polynomial $sp_i - p_i$ in powers of the (rotated) variables, we see that $sg_i - g_i$ must have no terms which are powers of $x_1, \dots, x_{n-1}$ alone. Put invariantly, we see that

$$sp_i - p_i = \ell r_i$$

where r_i is homogeneous of degree one less that that of p_i. Apply s to equation (11.5) and subtract to get

$$\ell\,(h_1 r_1 + \cdots h_m r_m) = 0.$$

Since $\ell \neq 0$ we may divide by ℓ to get an equation of the form (11.5) with p_1 replaced by r_1 of lower degree. So $r_1 \in I$ by induction. So

$$sp_1 - p_1 \in I.$$

Now W stabilizes R^+ and hence I and we have shown that each $w \in W$ acts trivially on the quotient of p_1 in this quotient space S/I. Thus $p_1^\sharp = \mathcal{A} p_1 \equiv p_1$ mod I. So $p_1 \in I$ since $p_1^\sharp \in I$. QED