261A Lie Groups

Vera Serganova
Notes by Qiaochu Yuan

Fall 2013

1 Introduction

We will first begin with Lie groups and some differential geometry. Next we will discuss some generalities about Lie algebras. We will discuss the classification of semisimple Lie algebras, root systems, the Weyl group, and Dynkin diagrams. This will lead into finite-dimensional representations and the Weyl character formula. Finally we will apply this to the classification of compact Lie groups.

To specify a group we specify three maps $m : G \times G \to G, i : G \to G^{-1}, e : \bullet \to G$ which satisfy various axioms. This definition can be internalized to various categories:

1. If G is a topological space and the maps are continuous, we get a topological group.
2. If G is a smooth manifold and the maps are smooth, we get a Lie group.
3. If G is a complex manifold and the maps are complex analytic, we get a complex Lie group.
4. If G is an algebraic variety and the maps are algebraic, we get an algebraic group.

The last three categories are surprisingly close to equivalent (the morphisms are given by smooth resp. complex analytic resp. algebraic maps which respect group structure).

Example $(\mathbb{R}, +)$ is a simple example of a Lie group. $(\mathbb{R}_{>0}, \times)$ is another. These two Lie groups are isomorphic with the isomorphism given by the exponential map. These groups are also (real) algebraic groups, but this isomorphism is not algebraic.

Example For $F = \mathbb{R}, \mathbb{C}$ the general linear group $\text{GL}_n(F)$ is a Lie group. $\text{GL}_n(\mathbb{C})$ is even a complex Lie group and a complex algebraic group. In particular, $\text{GL}_1(\mathbb{C}) \cong (\mathbb{C} \setminus \{0\}, \times)$.

$\text{GL}_n(\mathbb{R})$ is the smooth manifold $\mathbb{R}^{n^2}$ minus the closed subspace on which the determinant vanishes, so it is a smooth manifold. It has two connected components, one where $\det > 0$ and one where $\det < 0$. The connected component containing the identity is the normal subgroup $\text{GL}_n^+(\mathbb{R})$ of invertible matrices of positive determinant, and the quotient is $\mathbb{Z}/2\mathbb{Z}$. $\text{GL}_n(\mathbb{C})$ is similar except that it is connected.

$\text{GL}_n(F)$ also has as a normal subgroup the special linear group $\text{SL}_n(F)$ of elements of determinant 1. There is a short exact sequence

$$0 \to \text{SL}_n(F) \to \text{GL}_n(F) \to F^\times \to 0 \tag{1}$$

where the last map is given by the determinant.

Example S^1 is a Lie subgroup of $\text{GL}_1(\mathbb{C})$. Unlike the examples above, as a topological space S^1 is compact. There is a short exact sequence

$$0 \to \mathbb{Z} \to \mathbb{R} \to S^1 \to 0 \tag{2}$$

where the map $\mathbb{R} \to S^1$ sends x to e^{ix}.

Example The torus $T^2 \cong S^1 \times S^1$ is a Lie group. We can also think of T^2 as a smooth projective algebraic curve over $\mathbb{C}$ of genus 1 (an elliptic curve), in which case there are many distinct isomorphism classes of such curves.

Example The orthogonal group $\mathrm{O}(n)$ is the subgroup of $\mathrm{GL}_n(\mathbb{R})$ of elements X such that $X^T X = \mathrm{id}$, where X^T denotes the transpose. Equivalently, $\mathrm{O}(n)$ is the group of linear operators preserving the standard inner product on $\mathbb{R}^n$. This group has two connected components. The connected component containing the identity is the special orthogonal group $\mathrm{SO}(n)$ of elements of $\mathrm{O}(n)$ with determinant 1, and the quotient is $\mathbb{Z}/2\mathbb{Z}$.

This group is compact because it is closed and bounded with respect to the Hilbert-Schmidt norm $\left|\mathrm{tr}(A^T A)\right|$.

Example Similarly, the unitary group $\mathrm{U}(n)$ is the subgroup of $\mathrm{GL}_n(\mathbb{C})$ of elements such that $X^\dagger X = \mathrm{id}$ where $X^\dagger$ denotes the adjoint or conjugate transpose. Equivalently, it is the subgroup preserving an inner product or Hermitian form on $\mathbb{C}^n$. It is connected. As above, this group is compact because it is closed and bounded with respect to the Hilbert-Schmidt norm. $\mathrm{U}(n)$ is a Lie group but not a complex Lie group because the adjoint is not algebraic.

The determinant gives a map $\mathrm{U}(n) \to \mathrm{U}(1) \cong S^1$ whose kernel is the special unitary group $\mathrm{SU}(n)$, giving a short exact sequence

$$0 \to \mathrm{SU}(n) \to \mathrm{U}(n) \to S^1 \to 0. \tag{3}$$

Consider the example of $\mathrm{SU}(2)$ in particular. This consists of 2×2 matrices

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{4}$$

satisfying $|a|^2 + |b|^2 = 1$, $|c|^2 + |d|^2 = 1$, $a\bar{c} + b\bar{d} = 1$, and $ad - bc = 1$. All such matrices can be written

$$\begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix} \tag{5}$$

where $a = x + yi, b = z + wi$ satisfy $x^2 + y^2 + z^2 + w^2 = 1$, and there are no such conditions. So $\mathrm{SU}(2)$ can be identified with the 3-sphere S^3 as a smooth manifold. Without the last condition, the above is a representation of the quaternions $\mathbb{H}$ as 2×2 matrices over $\mathbb{C}$. The quaternions have a multiplicative norm $x^2 + y^2 + z^2 + w^2$, and $\mathrm{SU}(2)$ can also be thought of as the group $\mathrm{Sp}(1)$ of quaternions of norm 1.

$\mathrm{Sp}(1)$ acts by conjugation $v \mapsto qvq^{-1}$ on the quaternions. This action preserves the subspace of quaternions with zero imaginary part (a 3-dimensional real representation of $\mathrm{SU}(2)$), giving a homomorphism $\mathrm{SU}(2) \to \mathrm{GL}_3(\mathbb{R})$. This homomorphism is not surjective since $\mathrm{SU}(2)$ also preserves the norm, giving an action of $\mathrm{SU}(2)$ on the 2-sphere S^2 which turns out to furnish a surjective homomorphism

$$\mathrm{SU}(2) \to \mathrm{SO}(3). \tag{6}$$

This is a double cover with kernel -1. Since $\mathrm{SU}(2) \cong S^3$ is simply connected, it follows that $\pi_1(\mathrm{SO}(3)) \cong \mathbb{Z}/2\mathbb{Z}$.

2 Generalities

Definition If G is a Lie group and X is a manifold, an *action* of G on X is a smooth map $G \times X \to X$ such that $g_1(g_2 x) = (g_1 g_2)x$ and $ex = x$, where $g_1, g_2 \in G, e$ is the identit, and $x \in X$.

The above definition can also be internalized into other categories, e.g. complex manifolds or algebraic varieties.

Definition With an action as above, the *stabilizer* $\mathrm{Stab}(x)$ of a point $x \in X$ is the subgroup $\{g \in G : gx = x\}$.

The stabilizer is always a closed subgroup.

Example $\mathrm{SL}_n(\mathbb{R})$ naturally acts on $\mathbb{R}^n$. When $n > 1$, this action has two orbits: 0 and the nonzero vectors. The stabilizer of any two nonzero vectors is conjugate, so to compute the stabilizer of any nonzero vector it suffices to compute the stabilizer of one such vector. We will take $x_0 = (1, 0, ...0)$. The resulting stabilizer consists of block matrices of the form

$$\begin{bmatrix} 1 & * \\ 0 & M \end{bmatrix} \tag{7}$$

where $M \in \mathrm{SL}_{n-1}(\mathbb{R})$. This is the semidirect product of $\mathrm{SL}_{n-1}(\mathbb{R})$ and $\mathbb{R}^{n-1}$; topologically, it is homeomorphic to their product.

Consider the map

$$\gamma : \mathrm{SL}_n(\mathbb{R}) \ni g \mapsto gx_0 \in \mathbb{R}^n \setminus \{0\}. \tag{8}$$

This map sends a matrix g to its first column. It is surjective, smooth, and open. The fiber of any point is homeomorphic to the stabilizer subgroup above, and in fact γ is a fiber bundle.

Proposition 2.1. *$SL_n(\mathbb{R})$ is connected.*

Proof. We proceed by induction on n. When $n = 1$ the statement is clear. Inductively, suppose we have shown that $\mathrm{SL}_{n-1}(\mathbb{R})$ is connected. Then γ is a surjective open map with connected fibers and a connected codomain.

Proposition 2.2. *Let $\gamma : Y \to Z$ be a surjective open map with Z connected and connected fibers. Then Y is connected.*

Proof. We prove the contrapositive. Suppose $Y = Y_1 \sqcup Y_2$ is a disconnection of Y. Since the fibers are connected, Y_1, Y_2 cannot both intersect the same fiber, so $\gamma(Y_1) \cap \gamma(Y_2) = \emptyset$. But then Z cannot be connected. □

The conclusion follows. □

If G is a topological group, by convention G_0 will denote the connected component of the identity. For example, if $G = \mathbb{R} \setminus \{0\}$, then $G_0 = \mathbb{R}^+$.

Proposition 2.3. *G_0 is a normal subgroup of G.*

Proof. Multiplication restricts to a map $G_0 \times G_0 \to G$ whose image is connected and contains the identity, so must be contained in G_0. Similarly for inversion and conjugation by fixed $g \in G$. □

We can regard conjugation by fixed $g \in G$ as the composition of two maps

$$R_g : G \ni x \mapsto xg \in G \tag{9}$$

$$L_g : G \ni x \mapsto gx \in G. \tag{10}$$

These maps commute and provide families of diffeomorphisms $G \to G$.

Proposition 2.4. *Let G be a connected topological group. Then G is generated by any open set containing the identity.*

Proof. Let U be such an open set. We may assume WLOG that U is closed under inverses. Let $G' = \bigcup_{n=1}^{\infty} U^n$ be the subgroup generated by U. It is open since it is a union of open sets. G is the disjoint union of the cosets of G', which are all open, hence if G is connected then there is necessarily one such coset and $G' = G$. □

Corollary 2.5. *Any open subgroup in a topological group G contains G_0.*

Let $\varphi : G \to H$ be a continuous homomorphism of topological spaces. Then the kernel is a closed (normal) subgroup of G. The image is a subgroup, but it is not necessarily closed.

Example Let $H = \mathbb{T}^2 = S^1 \times S^1$, let $G = \mathbb{R}$, and consider the homomorphism

$$\varphi : \mathbb{R} \ni x \mapsto (e^{i\alpha x}, e^{i\beta x}) \in S^1 \times S^1 \tag{11}$$

where $\alpha, \beta \in \mathbb{R}$. This describes a geodesic on the torus (with respect to the flat metric). If $\frac{\alpha}{\beta}$ is rational, the geodesic is closed and the image is closed. Otherwise, the image is dense (in particular, not closed).

Theorem 2.6. *Let G be a connected topological group and let Z be a discrete normal subgroup. Then Z is central.*

Proof. Choose $z \in Z$. The map

$$G \ni g \mapsto gzg^{-1} \in Z \tag{12}$$

has connected image, which must therefore be a point, which must therefore be z. □

3 Lie algebras

Definition A *closed linear group* is a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$.

All of the examples we gave previously have this form.

Example $B \subset \mathrm{GL}_n(\mathbb{R})$, the subgroup of upper-triangular matrices (the *Borel subgroup*), is also a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$.

Consider the exponential map

$$\exp : \mathcal{M}_n(\mathbb{R}) \ni X \mapsto \sum_{n=0}^{\infty} \frac{X^n}{n!} \in \mathcal{M}_n(\mathbb{R}). \tag{13}$$

It is straightforward to check that this series converges; for example, we can use the operator norm with respect to the inner product on $\mathbb{R}^n$. It is also straightforward to check that $\exp(X+Y) = \exp(X)\exp(Y)$ when X, Y commute; in particular, $\exp(0) = 1 = \exp(X)\exp(-X)$, so $\exp(X)$ is always invertible and hence its image lies in $\mathrm{GL}_n(\mathbb{R})$. For fixed $A \in \mathcal{M}_n(\mathbb{R})$, it also follows that the map

$$\varphi : \mathbb{R} \ni t \mapsto \exp(At) \in \mathrm{GL}_n(\mathbb{R}) \tag{14}$$

is a homomorphism (a *one-parameter subgroup*).

exp is smooth. The Taylor series for the logarithm converges for matrices of norm less than 1, so exp is a diffeomorphism in some neighborhood of zero.

exp satisfies $\exp(XAX^{-1}) = X\exp(A)X^{-1}$ by inspection, so we can compute exponentials of matrices by computing their Jordan normal form. In particular, if $J = \lambda I + N$ is a Jordan block with N nilpotent, then $\exp(J) = e^{\lambda}\left(1 + N + \frac{N^2}{2!} + ...\right)$ where the series is finite.

The one-parameter subgroup $\varphi(t) = \exp(At)$ satisfies

$$\frac{d}{dt}\varphi(t) = A\varphi(t). \tag{15}$$

Lemma 3.1. *Let $\varphi : \mathbb{R} \to GL_n(\mathbb{R})$ be a smooth homomorphism such that $\varphi'(0) = A$. Then $\varphi(t) = \exp(At)$.*

Proof. Follows from the uniqueness of solutions to ODEs. □

Definition Let G be a closed subgroup in $\mathrm{GL}_n(\mathbb{R})$. Its *Lie algebra* is

$$\mathrm{Lie}(G) = \mathfrak{g} = \{A \in \mathcal{M}_n(\mathbb{R}) | e^{At} \in G \forall t \in \mathbb{R}\}. \tag{16}$$

Example If $G = \mathrm{SL}_n(\mathbb{R})$, then its Lie algebra $\mathfrak{sl}_n(\mathbb{R})$ consists of precisely the traceless matrices. This is a corollary of the identity $\det\exp(M) = \exp\operatorname{tr}(M)$, which follows from triangularizing M over $\mathbb{C}$. Alternatively, it suffices to verify that $\det\exp(At) = \exp\operatorname{tr}(At)$, and these are two one-parameter subgroups $\mathbb{R} \to \mathbb{R}^{\times}$ whose derivatives at 0 agree.

(Note: the following proposition did not appear in lecture. It corrects a gap in the proof of the next proposition.)

Proposition 3.2. *$A \in Lie(G)$ if and only if there exists a smooth function $\gamma : (-\epsilon, \epsilon) \to GL_n(\mathbb{R})$ such that $\gamma(0) = 1$, $\gamma(t) \in G$ for all t, and $\gamma'(0) = A$.*

Proof. One direction is clear: if $A \in \text{Lie}(G)$, we can take $\gamma(t) = e^{At}$. In the other direction, suppose we have γ as in the proposition. Observe that for fixed t and sufficiently large n we have a Taylor series expansion

$$\left(\gamma\left(\frac{t}{n}\right)\right)^n = 1 + At + O(|t|^2) \tag{17}$$

hence

$$\left(\gamma\left(\frac{t}{nm}\right)\right)^{nm} = \left(1 + \frac{At}{m} + \frac{O(|t|^2)}{m^2}\right)^m \tag{18}$$

and, taking $m \to \infty$, using the fact that $\lim_{m\to\infty}\left(1 + O\left(\frac{1}{m^2}\right)\right)^m = 1$, we have

$$\lim_{m\to\infty}\left(\gamma\left(\frac{t}{nm}\right)\right)^{nm} = \lim_{m\to\infty}\left(1 + \frac{At}{m} + \frac{O(|t|^2)}{m^2}\right)^m = e^{At}. \tag{19}$$

Since G is a closed subgroup, this limit, hence e^{At}, lies in G, which gives the desired result. $\square$

Proposition 3.3. *$Lie(G)$ is a subspace of $\mathcal{M}_n(\mathbb{R})$ closed under the* commutator bracket

$$[A, B] = AB - BA. \tag{20}$$

Proof. By definition $\text{Lie}(G)$ is closed under scalar multiplication. It is also closed: if $X \in \text{Lie}(G)$ we can choose $\epsilon > 0$ so that ϵX lies in a neighborhood of 0 such that the exponential map is a diffeomorphism, and then $\text{Lie}(G)$ is generated under scalar multiplication by the preimage of a neighborhood of the identity intersected with G, which is closed.

If $A, B \in \text{Lie}(G)$, then we compute that

$$\frac{d}{dt}\left(e^{At}e^{Bt}\right)|_{t=0} = A + B \tag{21}$$

hence that $A + B \in \text{Lie}(G)$ by the previous proposition. Similarly, we compute that

$$\frac{d}{dt}\left(e^{At}e^{Bs}e^{-At}e^{-Bs}\right)|_{t=0} = A - e^{Bs}Ae^{-Bs} \tag{22}$$

hence that $A - e^{Bs}Ae^{-Bs} \in \text{Lie}(G)$ by the previous proposition. We have

$$A - e^{Bs}Ae^{-Bs} = [A, B]s + O(|s|^2) \tag{23}$$

hence, dividing by s and taking $s \to 0$, using the fact that $\text{Lie}(G)$ is closed and closed under scalar multiplication, we conclude that $[A, B] \in \text{Lie}(G)$ as desired. $\square$

The above proposition suggests the following definition.

Definition A *Lie algebra* is a vector space $\mathfrak{g}$ together with a bilinear map $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying the following axioms:

1. (Alternativity) $\forall X \in \mathfrak{g} : [X, X] = 0$
2. (Jacobi identity) $\forall X, Y, Z \in \mathfrak{g} : [X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$.

Eventually we will write down an intrinsic definition of the Lie algebra associated to a Lie group; in particular, it will not depend on the choice of embedding into $\mathrm{GL}_n(\mathbb{R})$.

If G is a closed linear group, the exponential map restricts to a local diffeomorphism $\exp : \mathfrak{g} \to G$ where $\mathfrak{g}$ is the Lie algebra of G. Alternatively, we can consider the tangent space $T_g(G) \subseteq \mathcal{M}_n(\mathbb{R})$, which is the space of all derivatives $F'(0)$ where $F : [-\epsilon, \epsilon] \to G$ is a curve with $F(0) = g$. Then $\mathfrak{g} = T_e(G)$. Now we can associate to any smooth homomorphism $\Phi : G \to H$ its differential

$$\varphi = d\Phi|_e : \mathfrak{g} \to \mathfrak{h}. \tag{24}$$

Moreover, the diagram

$$\begin{array}{ccc} G & \xrightarrow{\Phi} & H \\ \uparrow{\scriptstyle \exp} & & \uparrow{\scriptstyle \exp} \\ \mathfrak{g} & \xrightarrow{\varphi} & \mathfrak{h} \end{array} \tag{25}$$

commutes. In particular, φ is a homomorphism of Lie algebras. Since exp is a local diffeomorphism, knowing φ determines Φ on a neighborhood of the identity in G. If G is connected, it follows that φ uniquely determines Φ.

Example What is the Lie algebra $\mathfrak{so}(n)$ of $\mathrm{SO}(n)$? We want A such that

$$\exp(At)\exp(At)^T = \exp(At)\exp(A^T t) = \mathrm{id}. \tag{26}$$

Differentiating, it follows that $A + A^T = 0$, and conversely this condition guarantees the aboe identity. So $\mathfrak{so}(n)$ consists of the skew-symmetric matrices. In particular, $\dim \mathfrak{so}(n) = \frac{n(n-1)}{2}$, so this is also the dimension of $\mathrm{SO}(n)$.

The functor $G \mapsto \mathfrak{g}$ preserves injective and surjective maps. It also reflects surjective maps if the codomain is connected: if an induced map $\varphi : \mathfrak{g} \to \mathfrak{h}$ on Lie algebras is surjective, then the corresponding map $\Phi : G \to H$ on Lie groups is surjective in a neighborhood of the origin, and since H is connected the image of Φ must therefore be all of H, so Φ is surjective. (But it does not reflect injective maps: the quotient map $\mathbb{R} \to S^1$ fails to be injective even though the corresponding map on Lie algebras is.)

In particular, suppose the map of Lie algebras $\mathfrak{su}(2) \to \mathfrak{so}(3)$ induced from the double cover $\mathrm{SU}(2) \to \mathrm{SO}(3)$ failed to be surjective. Since both Lie algebras have dimension 3, it must therefore have a nontrivial kernel, but this would be reflected in the kernel of the map $\mathrm{SU}(2) \to \mathrm{SO}(3)$ having positive dimension, which is not the case. Hence the map $\mathrm{SU}(2) \to \mathrm{SO}(3)$ is surjective.

4 Review of smooth manifolds

Let X be a Hausdorff topological space. A *chart* on X is an open subset $U \subseteq X$ together with a homeomorphism from U to an open subset $V \subseteq \mathbb{R}^n$. Two charts are *compatible* if, on their intersection, the corresponding map between open subsets of $\mathbb{R}^n$ is smooth. An *atlas* is a collection of compatible charts covering X.

Definition A *smooth manifold* is a Hausdorff topological space X as above equipped with a maximal atlas.

Every atlas is contained in a maximal atlas by Zorn's lemma, so to specify a smooth manifold it suffices to specify an atlas.

Example We can cover the sphere S^n with two charts as follows. Let $U = S^n \setminus \{(1, 0, ...0\}$ and let $U' = S^n \setminus \{(-1, 0, ...0\}$. Stereographic projection gives an identification of S^n minus a point with $\mathbb{R}^n$, and the two charts we get this way are compatible by explicit computation.

Definition Let X be a topological space. A *sheaf* on X consists of the following data satisfying the following conditions:

1. An assignment, to each open set U, of an object $F(U)$ (e.g. a set or an abelian group).
2. An assignment, to each inclusion $U \subseteq V$, of a restriction map $F(V) \to F(U)$.
3. The restriction maps should satisfy the property that if $U \subseteq V \subseteq W$ is a chain of inclusions, the composition $F(W) \to F(V) \to F(U)$ agrees with the restriction map $F(W) \to F(U)$.
4. If an open set U has an open cover $\bigsqcup_i U_i$ and $f, g \in F(U)$ such that $f|_{U_i}$ (the restriction of f to U_i) agrees with $g|_{U_i}$ for all i, then $f = g$.

Example The assignment, to every open subset U of $\mathbb{R}^n$, the algebra $C^\infty(U)$ of smooth functions $U \to \mathbb{R}$ is a sheaf of $\mathbb{R}$-algebras on $\mathbb{R}^n$.

Example More generally, if X is a smooth manifold then we can assign to X a sheaf C_X such that $C_X(U)$ consists of functions $U \to \mathbb{R}$ which are smooth on each chart. This is the sheaf of smooth functions on X.

Definition A *smooth map* is a continuous map $f : X \to Y$ between smooth manifolds such that, on each open set $U \subseteq Y$, the pullback map from functions on U to functions on $f^{-1}(U)$ sends smooth functions to smooth functions. A *diffeomorphism* is a smooth map with a smooth inverse.

Definition If X is a smooth manifold, its *tangent space* $T_x(X)$ at a point $x \in X$ consists of equivalence classes of curves $\varphi : \mathbb{R} \to X$ such that $\varphi(0) = x$. Two such curves φ_1, φ_2 are equivalent if, on some open neighborhood $U \ni x$, the identity

$$\frac{d}{dt}\left(f \circ \varphi_1(t)\right)|_{t=0} = \frac{d}{dt} f \circ \varphi_2(t)|_{t=0} \tag{27}$$

for all $f \in C_X(U)$.

Definition A *point derivative* at a point $x \in X$ is an $\mathbb{R}$-linear map

$$\delta : C_X(U) \to \mathbb{R} \tag{28}$$

such that $\delta(fg) = (\delta f)g(x) + f(x)(\delta g)$.

If $U \subseteq \mathbb{R}^n$ is open, every point derivative has the form

$$\delta f = \alpha_1 \frac{\partial}{\partial x_1} + ... + \alpha_n \frac{\partial}{\partial x_n}. \tag{29}$$

In particular, $\frac{d}{dt}\left(f \circ \varphi_1(t)\right)|_{t=0}$ is a point derivative, so in a chart we must have

$$\frac{d}{dt}\left(f \circ \varphi_1(t)\right)|_{t=0} = \sum_{i=1}^{n} \alpha_i \frac{\partial f}{\partial x_i} \tag{30}$$

where $\alpha_i = \frac{d}{dt} x_i(t)|_{t=0}$ and x_i are the coordinate functions on the chart.

A smooth map $f : X \to Y$ gives a pullback map $f^* : C_Y \to C_X$ on sheaves. This lets us pull back point derivatives on C_X to point derivatives on C_Y, giving a map $df_x : T_x(X) \to T_{f(x)}(Y)$ called the *derivative* of f at x.

Theorem 4.1. *(Inverse function theorem) Let $f : X \to Y$ be a smooth map and let $x \in X$. Then f is locally a diffeomorphism at $x \in X$ iff df_x is invertible.*

This is a local theorem, so it can be proved by working in a chart, hence in $\mathbb{R}^n$.

Definition The *sheaf of vector fields* on a smooth manifold X is the sheaf Vect such that Vect(U) consists of the *derivations* Der $C_X(U)$ on $C_X(U)$, namely those $\mathbb{R}$-linear maps

$$\delta : C_X(U) \to C_X(U) \tag{31}$$

satisfying the Leibniz rule $\delta(fg) = (\delta f)g + f(\delta g)$.

This is a sheaf of Lie algebras: if δ_1, δ_2 are two derivations on the functions on the same open subset U, their commutator or Lie bracket $[\delta_1, \delta_2]$ is also a derivation. If $x \in U$ is a point, a derivation gives a point derivation by evaluating its output on x. Hence a vector field on U is a smoothly varying collection of tangent vectors at each point of U.

Definition An *integral curve* for a vector field v on U is a map $\gamma(t) : (-a, b) \to U$ such that $\frac{d\gamma}{dt} = v(\gamma(t))$.

Integral curves exist and are unique by existence and uniqueness for ODEs. We can use integral curves to describe the geometric meaning of the Lie bracket $[v, w]$ of vector fields. Let γ be an integral curve for v with $\gamma(0) = x$. If f is a smooth function defined in a neighborhood of x, then there is a Taylor series

$$f(\gamma(t)) = ((\exp vt)f)(x) = f(x) + vf(x)t + ... \tag{32}$$

Starting from x, flow along an integral curve γ_1 for v for a time $-t$, then along an integral curve δ_1 for w for a time $-s$, then along an integral curve γ_2 for v for a time t, then along an integral curve δ_2 for w for a time s. This gets us to a point y, and we can compute that

$$\begin{aligned} f(y) - f(x) &= (\exp(vs)\exp(wt)\exp(-vs)\exp(-wt)f)(x) - f(x) && (33) \\ &= ([v, w]f)(x)st + O\left(\max(|s|, |t|)^3\right). && (34) \end{aligned}$$

Hence $[v, w]$ measures the failure of flowing along v to commute with flowing along w.

5 Lie algebras, again

Let G be a Lie group. Left multiplication $L_g : G \to G$ by a fixed element $g \in G$ induces a map $L_g^* : \text{Vect}(G) \to \text{Vect}(G)$. A vector field is called *left invariant* if it commutes with this map.

Definition The *Lie algebra* $\text{Lie}(G) \cong \mathfrak{g}$ of a Lie group G is the Lie algebra of left invariant vector fields on G.

Left invariance of a vector field implies that the value of the vector field at a point uniquely and freely determines its value everywhere; hence $\mathfrak{g} \cong T_e(G)$ as vector spaces. In particular, if G is a Lie algebra then there exists a nowhere-vanishing vector field on G (by translating a nonzero vector in some tangent space). This is not possible on S^2 by the hairy ball theorem, so in particular S^2 cannot be given the structure of a Lie group.

We can also talk about right invariant vector fields. This is another Lie algebra also isomorphic to $T_e(G)$ as a vector space. This gives two Lie brackets on $T_e(G)$, one of which is the negative of the other.

Example Let $G = \text{GL}_n(\mathbb{R})$. The assignment $X \mapsto -XA$ gives a left invariant vector field on G (where we identify the tangent space at any point of G with $\mathcal{M}_n(\mathbb{C})$ since it commutes with left multiplication.

Let G be a Lie group and $\mathfrak{g} \cong \text{Lie}(G) \cong T_e(G)$ be its Lie algebra. For $\xi \in T_e(G)$ we denote the associated left invariant vector field by v_ξ. Consider an integral curve of v_ξ, so a map $\varphi : \mathbb{R} \to G$ satisfying $\varphi(0) = e$ and

$$\frac{d\varphi}{dt} = v_\xi(\varphi). \tag{35}$$

A priori we can only guarantee that φ is locally defined, but we can show that φ is a homomorphism of Lie groups and then it extends to be defined everywhere. To see this, we use the fact that the left invariance of v_ξ implies that if φ is an integral curve then so is $g\varphi$. Now suppose φ is an integral curve with $g = \varphi(t), t > 0$. Then $g\varphi = \varphi$ by uniqueness. But this means that

$$\varphi(t)\varphi(s) = \varphi(t+s). \tag{36}$$

Let φ_ξ be the integral curve we have defined, and define the *exponential map*

$$\exp : \mathfrak{g} \ni \xi t \to \exp(\xi t) = \varphi_\xi(t) \in G. \tag{37}$$

This map is well-defined by existence and uniqueness. We can compute that $d\exp_e : \mathfrak{g} \to \mathfrak{g}$ is the identity, so exp is a local diffeomorphism near the identity.

Computing the differential at the identity of a homomorphism $\Phi : G \to H$ of Lie groups gives a map

$$\varphi : \mathfrak{g} \cong T_e(G) \to T_e(H) \cong \mathfrak{h}. \tag{38}$$

This turns out to be a homomorphism of Lie algebras, giving a functor

$$\text{LieGrp} \ni G \mapsto \text{Lie}(G) \cong \mathfrak{g} \in \text{LieAlg}. \tag{39}$$

There are various ways to see this. One is to work in local coordinates at the identity given by the exponential map. If $m : G \times G \to G$ is the multiplication map, then in local coordinates we can write

$$m(X, Y) = X + Y + B(X, Y) + \text{h.o.t.} \tag{40}$$

where B is some quadratic term. Then

$$m(Y, X) = X + Y + B(Y, X) + \text{h.o.t.} \tag{41}$$

and we can check that $m(X, Y) - m(Y, X) = [X, Y] + \text{h.o.t.}$.

It is somewhat surprising that the second term of this series completely determines it; this is the Baker-Campbell-Hausdorff formula. Hence the Lie bracket of $\mathfrak{g}$ locally determines the group multiplication of G.

5.1 The adjoint representation

G acts on itself by conjugation, giving for each $g \in G$ a conjugation map

$$G \ni x \mapsto gxg^{-1} \in G. \tag{42}$$

This map preserves the identity. Differentiating it at the identity in x, we obtain a map

$$\mathrm{Ad}(g) : T_e(G) \to T_e(G) \tag{43}$$

called the *adjoint representation* of G. As a function of g, the adjoint representation furnishes a map $G \to \mathrm{Aut}(\mathfrak{g})$. Differentiating this map, we now obtain a map

$$\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g}). \tag{44}$$

This turns out to be the Lie bracket; $X \in \mathfrak{g}$ is sent to the map $Y \mapsto [X, Y]$. We can verify this by computing that

$$\mathrm{Ad}(g)(Y) = Y \circ L_g^* \circ R_{g^{-1}}^* = L_g^* \circ v_Y \circ R_{g^{-1}}^*|_e \tag{45}$$

where v_Y is as above the left-invariant vector field associated to Y. Writing $g = \exp(Xt)$, which we can do for g sufficiently close to the identity, we can compute that (where the derivative with respect to t is evaluated at $t = 0$)

$$\begin{aligned} \mathrm{ad}(X)(Y) &= \frac{d}{dt} L_{\exp(Xt)}^* \circ v_Y \circ R_{\exp(-Xt)}^*|_e && (46)\\ &= \frac{d}{dt} L_{\exp(Xt)}^* \circ v_Y + v_Y \circ \frac{d}{dt} R_{\exp(-Xt)}^*|_e && (47)\\ &= (v_X \circ v_Y - v_Y \circ v_X)|_e && (48) \end{aligned}$$

as desired. Moreover, since this argument is natural in G, we can also use this argument to show that the map on Lie algebras induced by a morphism of Lie groups is a morphism of Lie algebras.

What can we say about the functor LieGrp $\to$ LieAlg? Since the exponential map is a local diffeomorphism, it is faithful. It would be interesting to see to what extent it is full: that is, we'd like to lift a morphism $\mathfrak{g} \to \mathfrak{h}$ of lie algebras up to a morphism $G \to H$ of Lie groups. This is always possible when G is connected and simply-connected. In fact, the functor from connected and simply-connected Lie groups to Lie algebras is an equivalence of categories.

5.2 Subgroups

Proposition 5.1. *A closed subgroup of a Lie group is a Lie group.*

Proof. Let H be a closed subgroup of a Lie group G with Lie algebra $\mathfrak{g}$, and let

$$\mathfrak{h} = \{\xi \in \mathfrak{g} | \exp \xi t \in H \forall t \in \mathbb{R}\}. \tag{49}$$

The same argument as in the case $G = \mathrm{GL}_n(\mathbb{R})$ shows that $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$. Now we claim that we can find neighborhoods U, V such that $U \cap H = \exp(V \cap \mathfrak{h})$. Once we have this, $V \cap \mathfrak{h}$ provides a chart in a neighborhood of the identity in H, which we can translate around to obtain charts on all of H.

To prove the claim, we'll write $\mathfrak{g} = \mathfrak{h} \oplus W$ for some complement W. Then in some neighborhood of the identity we have $\exp(\mathfrak{g}) = \exp(\mathfrak{h}) \exp(W)$. Assume the claim is false. Then we can find a sequence $w_n \in W$ such that $\exp w_n \in H$. The sequence $\frac{w_n}{|w_n|}$ has a convergent subsequence, so WLOG we can pass to it. For $t > 0$, let $a_n = \lceil \frac{t}{|w_n|} \rceil$. Then

$$\lim_{n \to \infty} \exp a_n w_n = \exp xt \in H \tag{50}$$

for some $x \in \mathfrak{h}$, but this contradicts $\mathfrak{h} \cap W = \{0\}$. □

The above implies, for example, that the center $Z(G)$ of any Lie group is also a Lie group. We can compute that its Lie algebra is the center

$$Z(\mathfrak{g}) = \{\xi \in \mathfrak{g} | \text{ad}_\xi = 0\} \tag{51}$$

of $\mathfrak{g}$. We can also compute that if H is a closed normal subgroup of G then $\mathfrak{h}$ is an ideal in $\mathfrak{g}$ (closed under bracketing with elements of $\mathfrak{g}$; all ideals are two-sided in the world of Lie algebras).

The functor from Lie groups to Lie algebras preserves injective and surjective maps. This is a consequence of the implicit function theorem.

5.3 Representations

A representation of a Lie group G on a vector space V is a morphism $G \to \text{GL}(V)$ of Lie groups. Differentiating this morphism gives a morphism $\mathfrak{g} \to \mathfrak{gl}(V) \cong \text{End}(V)$ of Lie algebras; this is (by definition) a representation of Lie algebras. In particular, differentiating the adjoint representation of G gives the adjoint representation of $\mathfrak{g}$.

If V is a representation then we write $V^G = \{v \in V | gv = v \forall g \in G\}$ for its invariant subspace. If G is connected, it suffices to check this condition in an arbitrarily small neighborhood of the identity, in particular a neighborhood such that the exponential map is a diffeomorphism, so equivalently we can use the invariant subspace

$$V^{\mathfrak{g}} = \{v \in V | \xi v = 0 \forall \xi \in \mathfrak{g}\} \tag{52}$$

as a $\mathfrak{g}$-representation.

If V, W are two representations of a Lie group G, then their tensor product $V \otimes W$ is also a representation of G with the action $g(v \otimes w) = gv \otimes gw$. Differentiating this condition gives

$$\xi(v \otimes w) = \xi v \otimes w + v \otimes \xi w \tag{53}$$

(the Leibniz rule), which defines the tensor product of representations of a Lie algebra.

6 Topology and Lie groups

Theorem 6.1. *Let G, H be Lie groups with G connected and simply connected. Let $\varphi : \mathfrak{g} \to \mathfrak{h}$ be a morphism between their Lie algebras. Then there exists a (unique) morphism $\Phi : G \to H$ such that $d\Phi_e = \varphi$.*

Uniqueness is a corollary of the fact that the exponential map is a local diffeomorphism and that a connected group is generated by a neighborhood of the identity, and we don't need simple connectedness for it.

Proof. Since G is connected and a smooth manifold, it is path-connected by smooth paths. Hence for any $g \in G$ we can find a smooth path $\gamma : [0, 1] \to G$ with $\gamma(0) = e, \gamma(1) = g$. We can write

$$\frac{d\gamma}{dt} = v_{\xi(t)} \tag{54}$$

where $\xi(t)$ is a path in $\mathfrak{g}$. Now we construct a path γ' in H satisfying

$$\frac{d\gamma'}{dt} = v_{\varphi(\xi(t))} \tag{55}$$

and set $\Phi(g) = \gamma'(1)$. *A priori* this depends on the choice of γ, but we will show that it does not. Let $\gamma_1, \gamma_2 : [0, 1] \to G$ be two smooth paths connecting e and g. Then there is a homotopy $h(t, s) : [0, 1] \times [0, 1] \to G$, which we can take to be smooth, such that $h(t, 0) = \gamma_1(t), h(t, 1) = \gamma_2(t)$. This homotopy h itself satisfies a differential equation, namely

$$\frac{\partial h}{\partial t} = v_{\xi(t,s)} \tag{56}$$

$$\frac{\partial h}{\partial s} = v_{\eta(t,s)} \tag{57}$$

and $\xi(t, s), \eta(t, s)$ are related. In fact,

$$\frac{\partial \eta}{\partial t} = \frac{\partial \xi}{\partial t} + [\xi, \eta]. \tag{58}$$

To prove this, let ψ be any smooth function on G. Then we can check that

$$\frac{\partial^2 \psi(h)}{\partial t \partial s} = v_{\frac{\partial \eta}{\partial t}} \psi + v_\xi \circ v_\eta \psi \tag{59}$$

and

$$\frac{\partial^2 \psi(h)}{\partial s \partial t} = v_{\frac{\partial \xi}{\partial t}} \psi + v_\eta \circ v_\xi \psi \tag{60}$$

and the conclusion follows by identifying these.

The image of the smooth homotopy h in H is another smooth homotopy f, and we can write down a similar differential equation

$$\frac{\partial f}{\partial t} = v_{\zeta(t,s)} \tag{61}$$

$$\frac{\partial f}{\partial s} = v_{\theta(t,s)} \tag{62}$$

satisfying the analogous relation

$$\frac{\partial \theta}{\partial t} = \frac{\partial \zeta}{\partial s} + [\zeta, \theta]. \tag{63}$$

But we know that $\zeta = \varphi(\xi)$, hence we know that $\theta = \varphi(\eta)$ is a solution to the above (hence it is the unique solution by existence and uniqueness). Now we compute that

$$\frac{\partial f}{\partial s}(1, s) = \theta(1, s) = \varphi(\eta(1, s)) = 0 \tag{64}$$

so the start and endpoints of the paths γ_1', γ_2' in H corresponding to γ_1, γ_2 are the same. Hence $\Phi : G \to H$ is well-defined by the simple connectedness of G.

It remains to show that Φ is a homomorphism, but this is clear by translating paths by left multiplication and using uniqueness. □

Corollary 6.2. *The category of finite-dimensional representations of a connected and simply connected Lie group G is isomorphic to the category of finite-dimensional representations of its Lie algebra $\mathfrak{g}$.*

6.1 Fiber bundles

Definition Let $\pi : X \to Y$ be a surjective map such that $\pi^{-1}(y) \cong Z$ for all $y \in Y$ and for some space Z. Then π is called a *fiber bundle* if it is locally trivial in the sense that for all $y \in Y$ there is an open neighborhood $U \ni y$ such that $\pi^{-1}(U) \cong U \times Z$ (compatibly with the projection down to U).

Assume that X is path connected (hence that Y is path connected). Let $\pi_0(Z)$ denote the set of path components of Z. If $y_0 \in Y$ and $x_0 \in X$ such that $\pi(x_0) = y_0$, then we get a sequence of maps

$$Z \xrightarrow{i} X \xrightarrow{\pi} Y \tag{65}$$

where $\pi(i(z)) = y_0 \forall z \in Z$. This induces a sequence of maps in π_1, and a part of the long exact sequence of homotopy groups shows that this sequence of maps fits into an exact sequence

$$\pi_1(Z) \xrightarrow{i^*} \pi_1(X) \xrightarrow{\pi^*} \pi_1(Y) \xrightarrow{\partial} \pi_0(Z) \to 0. \tag{66}$$

(Here 0 refers to the set $\{0\}$.) The interesting map here is ∂, and it is defined as follows. If $\gamma : [0, 1] \to Y$ is a path with $\gamma(0) = \gamma(1) = y_0$, then we can lift γ uniquely to a path $\tilde{\gamma} : [0, 1] \to X$ such that $\tilde{\gamma}(0) = x_0$. But $\tilde{\gamma}(1)$ may be somewhere else in the fiber $\pi^{-1}(y_0)$. However, $\tilde{\gamma}(1)$ only depends on γ up to homotopy, and this gives our map $\pi_1(Y) \to \pi_0(Z)$.

Example Let $\pi : \mathbb{R} \to S^1$ be the exponential map. This has fiber $\mathbb{Z}$. The corresponding exact sequence is

$$\pi_1(Z) \cong 0 \to \pi_1(X) \cong 0 \to \pi_1(S^1) \to \mathbb{Z} \to 0 \tag{67}$$

which (once we know that ∂ is a group homomorphism here) shows that $\pi_1(S^1) \cong \mathbb{Z}$ by exactness.

Example Consider again the map $\mathrm{SL}_2(\mathbb{R}) \to \mathbb{R}^2 \setminus \{0\}$ given by sending a matrix to its first column. This is a fiber bundle with fiber $\mathbb{R}$. The corresponding exact sequence is

$$\pi_1(Z) \cong 0 \to \pi_1(\mathrm{SL}_2(\mathbb{R})) \to \pi_1(\mathbb{R}^2 \setminus \{0\}) \cong \mathbb{Z} \to 0 \tag{68}$$

which shows that $\pi_1(\mathrm{SL}_2(\mathbb{R})) \cong \mathbb{Z}$ by exactness.

Definition A *covering map* $\pi : X \to Y$ is a fiber bundle map with discrete fiber. We say that X is a *covering space* of Y.

If G is a connected Lie group, we will construct a connected and simply connected Lie group $\tilde{G}$ (its *universal cover*) and a covering map $\pi : \tilde{G} \to G$ which has the following lifting property: if

$$\begin{array}{ccc} \tilde{G} & \dashrightarrow & H \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \pi'} \\ G & \xrightarrow[f]{} & K \end{array} \tag{69}$$

is a diagram of connected Lie groups with π' a covering space, then there is a map $\tilde{G} \to H$ making the diagram commute. $\tilde{G}$ is constructed by starting with the group of paths $\gamma : [0,1] \to G$ with $\gamma(0) = e$ and quotienting by homotopy relative to endpoints. The map $\tilde{G} \to G$ is given by evaluating at $\gamma(1)$, and its kernel is precisely the fundamental group $\pi_1(G)$ (based at e), giving an exact sequence

$$0 \to \pi_1(G) \to \tilde{G} \to G \to 0. \tag{70}$$

Hence we can think of $\pi_1(G)$ as a discrete normal subgroup of $\tilde{G}$. Since we know that such subgroups are central and, in particular, abelian, it follows that $\pi_1(G)$ is abelian.

By local triviality, the map $d\pi_e : \mathrm{Lie}(\tilde{G}) \to \mathrm{Lie}(G)$ is naturally an isomorphism, so their Lie algebras are the same. Conversely the following is true.

Proposition 6.3. *Let G_1, G_2 be two connected Lie groups. Then $\mathfrak{g}_1 \cong \mathfrak{g}_2$ iff $\tilde{G}_1 \cong \tilde{G}_2$.*

Proof. The isomorphism $\mathfrak{g}_1 \to \mathfrak{g}_2$ lifts to a map $\tilde{G}_1 \to \tilde{G}_2$ and the same is true for the map in the other direction. $\square$

In particular, there is at most one connected and simply connected Lie group G associated to a Lie algebra $\mathfrak{g}$. We will show later that there always exists such a Lie group.

Example Let G be a connected abelian Lie group. Then $\mathfrak{g} = \mathbb{R}^n$ with the trivial Lie bracket. Since $\mathbb{R}^n$ itself is a connected and simply connected Lie group with this Lie algebra, it follows that $\tilde{G} \cong \mathbb{R}^n$, so G must be the quotient of $\mathbb{R}^n$ by some discrete subgroup. These all have the form $\mathbb{Z}^m$, and the resulting quotients have the form $(S^1)^m \times \mathbb{R}^{n-m}$.

In the analytic category things get much more interesting. Even the study of quotients $\mathbb{C}/\Gamma$ (with Γ of full rank) is interesting in the analytic category; these are elliptic curves and there is a nontrivial moduli space $M_{1,1}$ of these of dimension 1. We get abelian varieties this way.

6.2 Homogeneous spaces

Let G be a Lie group and H a closed subgroup of G. Then there is a natural projection map $\pi : G \to G/H$ where G/H is the space of left cosets. It is possible to define a natural manifold structure on the latter space.

Definition A manifold G/H as above is a *homogeneous space* for G.

Since G acts transitively on G/H, it suffices to define this structure in a neighborhood of the identity coset, and locally we can identify a neighborhood of the identity coset with a complement of $\mathfrak{h}$ in $\mathfrak{g}$ via the exponential map.

Example The quotient $\text{SO}(n)/\text{SO}(n-1)$ is S^{n-1}.

Example As we saw earlier, the quotient of $\text{SL}_2(\mathbb{R})$ by the subgroup of matrices of the form $\begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix}$ is $\mathbb{R}^2 \setminus \{0\}$.

7 Structure theory

Below all Lie algebras are finite-dimensional.

Let G be a (path) connected Lie group. The commutator subgroup $[G, G]$ is path connected, but may not be closed in G.

Proposition 7.1. *Let G be connected and simply connected. Then $[G, G]$ is closed with Lie algebra the* commutator subalgebra

$$[\mathfrak{g}, \mathfrak{g}] = \mathit{span}([X, Y] : X, Y \in \mathfrak{g}). \tag{71}$$

Proof. The quotient map $\mathfrak{g} \to \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is the universal map from $\mathfrak{g}$ to an abelian Lie algebra $\mathbb{R}^n$, giving an exact sequence

$$0 \to [\mathfrak{g}, \mathfrak{g}] \to \mathfrak{g} \to \mathbb{R}^n \to 0. \tag{72}$$

We can lift these maps to an exact sequence of connected and simply connected Lie groups

$$0 \to K \to G \to \mathbb{R}^n \to 0 \tag{73}$$

where K is a closed subgroup of G. Since G/K is abelian, K contains $[G, G]$. We also know that $\text{Lie}([G, G])$ contains $[\mathfrak{g}, \mathfrak{g}]$ by computation. Hence the Lie algebra of K and $[G, G]$ coincide, and since K is connected and simply connected the conclusion follows. $\square$

7.1 Automorphisms

Let G be a connected and simply connected Lie group and let $\text{Aut}(G)$ be its group of automorphisms. Every $\varphi \in \text{Aut}(G)$ also acts on $\mathfrak{g}$, giving a morphism $\text{Aut}(G) \to \text{Aut}(\mathfrak{g})$ which is an isomorphism by the above results. This is not true if G fails to be simply connected. For example, if $G = V/\Gamma$ is a quotient of a finite-dimensional real vector space V by a discrete subgroup Γ, then $\text{Aut}(\tilde{G}) \cong \text{Aut}(V)$ is $\text{GL}(V)$, but $\text{Aut}(G)$ is the subgroup of $\text{GL}(V)$ consisting of those elements which send Γ to Γ.

The Lie algebra of $\text{Aut}(\mathfrak{g})$ is the Lie algebra $\text{Der}(\mathfrak{g})$ of derivations of $\mathfrak{g}$.

7.2 Semidirect products

Definition Let $\mathfrak{g}, \mathfrak{h}$ be Lie algebras and $\varphi : \mathfrak{g} \to \text{Der}(\mathfrak{h})$ be an action of $\mathfrak{g}$ on $\mathfrak{h}$. Their *semidirect product* $\mathfrak{g} \rtimes \mathfrak{h}$ is the direct sum $\mathfrak{g} \oplus \mathfrak{h}$ equipped with the bracket

$$[g_1 + h_1, g_2 + h_2] = [g_1, g_2] - \varphi(g_2)h + \varphi(g_1)h + [h_1, h_2]. \tag{74}$$

The three Lie algebras above fit into a short exact sequence

$$0 \to \mathfrak{h} \to \mathfrak{g} \rtimes \mathfrak{h} \to \mathfrak{g} \to 0 \tag{75}$$

of vector spaces, where all maps are Lie algebra homomorphisms. In particular, $\mathfrak{h}$ is an ideal in $\mathfrak{g} \rtimes \mathfrak{h}$. Conversely, given an ideal in a Lie algebra we can write down the above short exact sequence, and choosing a splitting of it gives a description of the Lie algebra as a semidirect product of the ideal and the quotient.

Suppose G, H are connected and simply connected and $\varphi : \mathfrak{g} \to \text{Der}(\mathfrak{h})$ is an action. Then we can lift it to an action $\Phi : G \to \text{Aut}(H)$, which lets us construct the semidirect product $G \rtimes H$ of Lie groups. This is the cartesian product as a set with group operation

$$(g_1, h_1)(g_2, h_2) = (g_1 g_2, \Phi(g_2^{-1})h_1 h_2). \tag{76}$$

7.3 Lie algebras of low dimension

We eventually want to prove that if $\mathfrak{g}$ is a finite-dimensional Lie algebra then there is a (unique up to isomorphism) connected and simply connected Lie group G with Lie algebra $\mathfrak{g}$ (Lie's third theorem). To do this we will need some structure theory.

As an exercise, let's classify Lie algebras of low dimension over $\mathbb{C}$.

1. If $\dim \mathfrak{g} = 1$, the abelian Lie algebra K is the only possibility.

2. If $\dim \mathfrak{g} = 2$, the Lie bracket defines a map $\Lambda^2(\mathfrak{g}) \to \mathfrak{g}$ with image the commutator subalgebra. The exterior square $\Lambda^2(\mathfrak{g})$ is 1-dimensional, so $\dim[\mathfrak{g}, \mathfrak{g}]$ is either 0 or 1.

   (a) If $\dim[\mathfrak{g}, \mathfrak{g}] = 0$ we get the abelian Lie algebra K^2.

   (b) If $\dim[\mathfrak{g}, \mathfrak{g}] = 1$, take $X \notin [\mathfrak{g}, \mathfrak{g}]$ and $Y \in [\mathfrak{g}, \mathfrak{g}]$. Then $[X, Y] = \lambda Y$ ($\lambda \neq 0$ or we reduce to the previous case) and by rescaling X we can take $[X, Y] = Y$. The resulting Lie algebra is isomorphic to the Lie algebra of 2×2 matrices of the form

$$\begin{bmatrix} * & * \\ 0 & 0 \end{bmatrix}. \tag{77}$$

3. If $\dim \mathfrak{g} = 3$, we will again work based on the dimension of the commutator subalgebra.

   (a) If $\dim[\mathfrak{g}, \mathfrak{g}] = 0$ we get the abelian Lie algebra $\mathbb{C}^3$.

   (b) If $\dim[\mathfrak{g}, \mathfrak{g}] = 1$, then the Lie bracket gives a skew-symmetric form $\mathfrak{g} \otimes \mathfrak{g} \to [\mathfrak{g}, \mathfrak{g}] \cong \mathbb{C}$ with kernel the center $Z(\mathfrak{g})$. Since $\dim \mathfrak{g} = 3$ and $\mathfrak{g}$ is not abelian, this kernel has dimension 1.

      i. If $Z(\mathfrak{g}) \cap [\mathfrak{g}, \mathfrak{g}] = 0$, then we can write $\mathfrak{g}$ as a direct sum $Z(\mathfrak{g}) \oplus \mathfrak{h}$ where $[\mathfrak{g}, \mathfrak{g}] \in \mathfrak{h}$. This exhibits $\mathfrak{g}$ as the direct sum of a 1-dimensional (abelian) Lie algebra and a 2-dimensional Lie algebra, which we've classified.

      ii. If $Z(g) = [\mathfrak{g}, \mathfrak{g}]$, then by taking $Z \in Z(\mathfrak{g})$ so that $Z = [X, Y]$ we have $[Z, X] = [Z, Y] = 0$, and this completely defines the Lie algebra since X, Y are independent. This is the Heisenberg Lie algebra of strictly upper triangular 3×3 matrices, e.g. those of the form

$$\begin{bmatrix} 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{bmatrix}. \tag{78}$$

   (c) If $\dim[\mathfrak{g}, \mathfrak{g}] = 2$, we claim that $\mathfrak{h} = [\mathfrak{g}, \mathfrak{g}]$ is abelian. Suppose otherwise. Then $\mathfrak{h}$ is spanned by X, Y such that $[X, Y] = Y$. If $Z \notin \mathfrak{h}$, then by the Jacobi identity

$$[Z, Y] = [Z, [X, Y]] = [[Z, X], Y] + [X, [Z, Y]] \in [\mathfrak{h}, \mathfrak{h}]. \tag{79}$$

   But $[\mathfrak{h}, \mathfrak{h}]$ is spanned by Y. It follows that $[Z, Y] = \lambda Y$ for some λ and $[Z, X] = aX + bY$. Since $\dim \mathfrak{h} = 2$ we know that $a \neq 0$. With respect to the basis X, Y, Z, the matrix of ad_X is

$$\begin{bmatrix} 0 & 0 & -a \\ 0 & 1 & -b \\ 0 & 0 & 0 \end{bmatrix} \tag{80}$$

and in particular it has trace 1, but ad of any element in the commutator (which X is in) is zero; contradiction.

Hence $\mathfrak{h}$ is abelian and $\mathfrak{g}$ is a semidirect product of $\mathbb{C}$ and $\mathbb{C}^2$, with a basis element $Z \in \mathbb{C}$ acting on $\mathbb{C}^2$ by some invertible 2×2 matrix (up to conjugation and scalar multiplication).

4. If $\dim[\mathfrak{g}, \mathfrak{g}] = 3$, we claim that $\mathfrak{g} \cong \mathfrak{sl}_2(\mathbb{C})$. For $X \in \mathfrak{g}$, consider $\text{ad}(X) : \mathfrak{g} \to \mathfrak{g}$. Since X is a commutator, $\text{tr}(\text{ad}(X)) = 0$. Furthermore, $X \in \ker(\text{ad}(X))$. Hence the eigenvalues of $\text{ad}(X)$ must be $0, a, -a$ for some a (since in this case $\mathfrak{g}$ has no center). If $a \neq 0$, then $\text{ad}(X)$ is semisimple; otherwise, $\text{ad}(X)$ is nilpotent.

   We claim there exists $X \in \mathfrak{g}$ such that $\text{ad}(X)$ is not nilpotent (hence semisimple). Suppose otherwise. If $[X, Y] = X$ for some Y then $\text{ad}(Y)(X) = -X$, so $\text{ad}(Y)$ would not be nilpotent. Hence we may assume that $X \notin \text{im}(\text{ad}(X))$ for all X. Hence $\text{ad}(X), X \neq 0$ has rank 1 and is nilpotent, so we can extend X to a basis X, Y, Z such that $\text{ad}(X)(Y) = 0$, so $[X, Y] = 0$, and $\text{ad}(X)(Z) = Y$, so $[X, Z] = Y$. But then $\dim[\mathfrak{g}, \mathfrak{g}] \leq 2$; contradiction.

   Hence we can find $H \in \mathfrak{g}$ with eigenvalues $0, 2, -2$ (by rescaling appropriately). If X, Y are the eigenvectors of eigenvalues $2, -2$, then $[H, X] = 2X, [H, Y] = -2Y$, and to completely determine $\mathfrak{g}$ it suffices to determine $[X, Y]$. Now,

   $$[H, [X, Y]] = [[H, X], Y] + [X, [H, Y]] = [2X, Y] - [X, 2Y] = 0 \quad (81)$$

   from which it follows that $[X, Y]$ is proportional to H. By rescaling we can arrange $[X, Y] = H$, and this is a presentation of $\mathfrak{sl}_2(\mathbb{C})$ as desired. More explicitly, $\mathfrak{sl}_2(\mathbb{C})$ has a basis

   $$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, Y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}. \quad (82)$$

The above argument in the perfect case $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ doesn't work over $\mathbb{R}$, since we used Jordan normal form. Over $\mathbb{R}$ both $\mathfrak{sl}_2(\mathbb{R})$ and $\mathfrak{so}(3) \cong \mathfrak{su}(2)$ are 3-dimensional and non-isomorphic perfect Lie algebras. However, by tensoring by $\mathbb{C}$ we get 3-dimensional perfect complex Lie algebras, which must be isomorphic to $\mathfrak{sl}_2(\mathbb{C})$. Our strategy for classifying real Lie algebras will be to classify complex Lie algebras and then to classify real Lie algebras which tensor up to our complex Lie algebras (their *real forms*).

7.4 Solvable and nilpotent Lie algebras

Let $D(\mathfrak{g}) = D^1(\mathfrak{g}) = D_1(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}]$ be the commutator. This is an ideal of $\mathfrak{g}$. We inductively define

$$D_n(\mathfrak{g}) = [\mathfrak{g}, D_{n-1}(\mathfrak{g})] \quad (83)$$

and

$$D^n(\mathfrak{g}) = [D^{n-1}(\mathfrak{g}), D^{n-1}(\mathfrak{g})]. \quad (84)$$

Definition A Lie algebra $\mathfrak{g}$ is *nilpotent* if $D_n(\mathfrak{g}) = 0$ for some n and *solvable* if $D^n(\mathfrak{g}) = 0$ for some n.

Example $\mathfrak{b}_n$, the Lie algebra of upper triangular matrices in $\mathfrak{gl}_n$, is solvable. $\mathfrak{n}_n$, the Lie algebra of strictly upper triangular matrices in $\mathfrak{gl}_n$, is nilpotent. We can compute that $[\mathfrak{b}_n, \mathfrak{b}_n] = \mathfrak{n}_n$, but $[\mathfrak{b}_n, \mathfrak{n}_n] = \mathfrak{n}_n$, so $\mathfrak{b}_n$ is not nilpotent.

Any subalgebra or quotient of a nilpotent resp. solvable Lie algebra is nilpotent resp. solvable. In a short exact sequence

$$0 \to \mathfrak{j} \to \mathfrak{g} \to \mathfrak{h} \to 0 \quad (85)$$

of Lie algebras, if the first and last two terms are solvable then so is the middle term. This is not true in the nilpotent case, since we have a short exact sequence

$$0 \to \mathfrak{n}_n \to \mathfrak{b}_n \to k^n \to 0 \quad (86)$$

where k is the base field. This generalizes: if $\mathfrak{g}$ is a solvable Lie algebra, then it has an ideal of codimension 1. (To see this it suffices to choose any subspace of codimension 1 containing $[\mathfrak{g}, \mathfrak{g}]$.) But now it follows that any solvable Lie algebra is an iterated semidirect product of abelian Lie algebras.

Corollary 7.2. *Let $\mathfrak{g}$ be a solvable Lie algebra over $\mathbb{R}$ resp. $\mathbb{C}$. Then there exists a connected and simply connected Lie group G with Lie algebra $\mathfrak{g}$. As a manifold, $G \cong \mathbb{R}^n$ resp. $\mathbb{C}^n$.*

Proof. We proceed by induction on the dimension. It suffices to observe that, when we lift a semidirect product of Lie algebras to a semidirect product of Lie groups, topologically the corresponding Lie group is the product of its factors. The only factors we will use here arise from abelian Lie algebras, hence are products of copies of $\mathbb{R}$ or $\mathbb{C}$. □

If $\mathfrak{g}$ is nilpotent, then $\text{ad}(X)$ is nilpotent for every $X \in \mathfrak{g}$. What about the converse?

Theorem 7.3. *(Engel) Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be a Lie subalgebra of a general linear Lie algebra. Suppose that every $X \in \mathfrak{g}$ is a nilpotent linear operator. Then the invariant subspace $V^{\mathfrak{g}}$ is nonzero; that is, there exists $v \in V$ such that $Xv = 0 \forall X \in \mathfrak{g}$.*

This is true over an arbitrary ground field.

Proof. First we make the following observation. Regard V as a representation of $\mathfrak{g}$. Then the dual V^* is again a representation, with

$$(X\varphi)(v) = \varphi(-Xv), v \in V, \varphi \in V^*. \quad (87)$$

Hence the tensor product $V \otimes V^*$ is again a representation, with

$$X(v \otimes \varphi) = Xv \otimes \varphi + v \otimes X\varphi. \tag{88}$$

But this is precisely the restriction of the adjoint representation of $\mathfrak{gl}(V)$ to $\mathfrak{g}$ (which in particular contains the adjoint representation of $\mathfrak{g}$). Now we compute that if $X^n = 0$, then

$$X^{2n}(v \otimes \varphi) = \sum_{k=0}^{2n} \binom{2n}{k} X^k v \otimes X^{2n-k}\varphi = 0 \tag{89}$$

hence if X is nilpotent, then $\text{ad}(X)$ is nilpotent.

Now the proof of the theorem proceeds by induction on $\dim \mathfrak{g}$. It is clear for $\dim \mathfrak{g} = 1$. In general, we claim that any maximal proper subalgebra $\mathfrak{h} \subsetneq \mathfrak{g}$ is an ideal of codimension 1. To see this, consider the adjoint representation $\text{ad} : \mathfrak{h} \to \mathfrak{gl}(\mathfrak{g})$. Since $\mathfrak{h}$ is invariant under this action, we can consider the quotient representation on $\mathfrak{g}/\mathfrak{h}$ and the representation on $\mathfrak{h}$ and apply Engel's theorem to these by the inductive hypothesis (since we know that if X is nilpotent then $\text{ad}(X)$ is nilpotent).

This gives us a nonzero element $X \not\in \mathfrak{h}$ such that $[\mathfrak{h}, X] \subseteq \mathfrak{h}$. Now $kX \oplus \mathfrak{h}$ is a bigger subalgebra of $\mathfrak{g}$ containing $\mathfrak{h}$, so $kX \oplus \mathfrak{h} = \mathfrak{g}$ and $\mathfrak{h}$ is an ideal of codimension 1 as desired.

Now we finish the proof. With $\mathfrak{g}, V$ as in the problem, $\mathfrak{g}$ has a maximal proper subalgebra $\mathfrak{h}$ which must be an ideal of codimension 1. By the inductive hypothesis, $V^{\mathfrak{h}}$ is nonzero. For $h \in \mathfrak{h}, v \in V^{\mathfrak{h}}, Y \in \mathfrak{g} \setminus \mathfrak{h}$, write

$$hYv = Yhv + [h, Y]v = 0 \tag{90}$$

since $hv = 0$ and $[h, Y] \in \mathfrak{h}$. Hence Y maps $V^{\mathfrak{h}}$ to $V^{\mathfrak{h}}$. Since Y is nilpotent, it has an invariant subspace $0 \neq V^{\mathfrak{g}} \subseteq V^{\mathfrak{h}}$, and the conclusion follows. □

Corollary 7.4. *Under the above hypotheses, there exists a complete flag*

$$0 = V_0 \subset V_1 \subset ... \subset V_n = V \tag{91}$$

of subspaces of V *(where* $\dim V_i = i$*) such that* $\mathfrak{g}V_i \subseteq V_{i-1}$.

This follows by induction. By choosing a basis compatible with such a flag, we may equivalently say that $\mathfrak{g}$ can be conjugated to be a subalgebra of the strictly upper triangular matrices $\mathfrak{n}(V)$.

Corollary 7.5. *Let* $\mathfrak{g}$ *be a Lie algebra such that ad*(X) *is nilpotent for all* $X \in \mathfrak{g}$*. Then* $\mathfrak{g}$ *is nilpotent.*

Proof. $\mathfrak{g}$ is the middle term of a short exact sequence

$$0 \to Z(\mathfrak{g}) \to \mathfrak{g} \to \text{im}(\text{ad}) \to 0 \tag{92}$$

where $Z(\mathfrak{g})$ is central, in particular abelian, and $\text{im}(\text{ad})$ is the image of the adjoint representation $\mathfrak{g} \to \text{Der}(\mathfrak{g})$. We can apply the corollary of Engel's theorem to $\text{im}(\text{ad})$, and it follows that im(ad) is nilpotent. If im(ad) is nilpotent of order n, then $D_n(\mathfrak{g}) \subseteq Z(\mathfrak{g})$, hence $D_{n+1}(\mathfrak{g}) = 0$ as desired. □

Lie's theorem is an analogue of Engel's theorem for solvable Lie algebras, but we need to assume that our ground field k has characteristic zero and is algebraically closed.

Theorem 7.6. *(Lie) Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be a solvable subalgebra. Then every element of $\mathfrak{g}$ acting on V has a common eigenvector; that is, there is a nonzero $v \in V$ such that $Xv = \lambda(X)v$ for all $X \in \mathfrak{g}$ and some scalars $\lambda(X)$.*

Here $\lambda(X)$ is also referred to as a *weight*. Note that λ must be linear and must factor through the abelianization $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$.

Proof. We proceed by induction on $\dim \mathfrak{g}$. Let $\mathfrak{h} \subseteq \mathfrak{g}$ be an ideal of codimension 1. If $\lambda \in \mathfrak{h}^*$, consider the *weight space*

$$V_\lambda = \{v \in V : hv = \lambda(h)v \forall h \in \mathfrak{h}\} \tag{93}$$

as well as the *generalized weight space*

$$\tilde{V_\lambda} = \{v \in V : (h - \lambda(h))^n v = 0 \forall h \in \mathfrak{h}\} \tag{94}$$

(we can take $n = \dim V$). These are always invariant under the action of $\mathfrak{h}$. By the inductive hypothesis, we can find λ such that V_λ is nonzero.

We claim that $\tilde{V_\lambda}$ is $\mathfrak{g}$-invariant. To see this, for $h \in \mathfrak{h}, X \in \mathfrak{g}, v \in \tilde{V_\lambda}$ consider

$$(h - \lambda(h))Xv = X(h - \lambda(h))v + [h, X]v \in \tilde{V_\lambda} \equiv X(h - \lambda(h)) \bmod \tilde{V_\lambda} \tag{95}$$

which gives

$$(h - \lambda(h))^n Xv \equiv X(h - \lambda(h))^n v \bmod \tilde{V_\lambda} \tag{96}$$

and by multiplying both sides by $(h - \lambda(h))^m$ for m sufficiently large m, as well as taking n sufficiently large, it follows that $Xv \in \tilde{V_\lambda}$ as desired.

Now we claim that V_λ is $\mathfrak{g}$-invariant. We know that if $h \in \mathfrak{h}$ is a commutator then its trace is zero. On the other hand, the trace of h acting on $\tilde{V_\lambda}$ is $(\dim \tilde{V_\lambda})\lambda(h)$, hence $\lambda(h) = 0$ in this case. Now, with $h \in \mathfrak{h}, X \in \mathfrak{g}, v \in V_\lambda$ we have

$$hXv = Xhv + [h, X]v = X\lambda(h)v + \lambda([h, X])v = \lambda(h)Xv \tag{97}$$

hence $Xv \in V_\lambda$ as desired.

Now we finish the proof. We know that $\mathfrak{g} = kX \oplus \mathfrak{h}$ and that $\mathfrak{g}$ has an invariant subspace V_λ where $\mathfrak{h}$ acts on V_λ by the weight $\lambda(h)$. Now we can find an eigenvector for the action of X in V_λ, and the conclusion follows. □

Corollary 7.7. *With hypotheses as above, there exists a complete flag*

$$0 = V_0 \subsetneq V_1 \subsetneq ... \subsetneq V_n = V \tag{98}$$

such that $\mathfrak{g}(V_i) \subseteq V_i$. *Equivalently,* $\mathfrak{g}$ *is conjugate to a subalgebra of the Borel subalgebra* $\mathfrak{b}(V)$ *of upper triangular matrices.*

Corollary 7.8. *With hypotheses as above,* $[\mathfrak{g}, \mathfrak{g}]$ *is nilpotent.*

(We only need to use the adjoint representation here; the fact that the adjoint representation is faithful is not an issue because its kernel is the center.) By extension of scalars this result is also true for a ground field of characteristic 0, not necessarily algebraically closed.

7.5 Semisimple Lie algebras

Lemma 7.9. *If* $\mathfrak{m}, \mathfrak{n}$ *are solvable ideals in* $\mathfrak{g}$*, then* $\mathfrak{m} + \mathfrak{n}$ *is also solvable.*

Proof. Observe that

$$[\mathfrak{m} + \mathfrak{n}, \mathfrak{m} + \mathfrak{n}] \subseteq [\mathfrak{m}, \mathfrak{m}] + [\mathfrak{n}, \mathfrak{n}] + \mathfrak{m} \cap \mathfrak{n} \tag{99}$$

so we can proceed by induction on the solvability degrees of $\mathfrak{m}$ and $\mathfrak{n}$. □

Consequently every Lie algebra has a maximal solvable ideal.

Definition The maximal solvable ideal of a Lie algebra $\mathfrak{g}$ is its *radical* $\text{rad}(\mathfrak{g})$. A Lie algebra is *semisimple* if $\text{rad}(\mathfrak{g}) = 0$.

Proposition 7.10. $\mathfrak{g}/rad(\mathfrak{g})$ *is semisimple.*

Proof. We have a short exact sequence

$$0 \to \text{rad}(\mathfrak{g}) \to \mathfrak{g} \xrightarrow{p} \mathfrak{g}/\text{rad}(\mathfrak{g}) \to 0. \tag{100}$$

If $\mathfrak{h} = \text{rad}(\mathfrak{g}/\mathfrak{g})$ then this restricts to a short exact sequence

$$0 \to \text{rad}(\mathfrak{g}) \to p^{-1}(\mathfrak{h}) \to \mathfrak{h} \to 0. \tag{101}$$

But $p^{-1}(\mathfrak{h})$ is solvable, hence the map from $\text{rad}(\mathfrak{g})$ is an isomorphism and $\mathfrak{h} = 0$ as desired. □

To continue our study of structure theory we will decompose an arbitrary Lie algebra into its radical and its semisimple quotient, and then we will classify semisimple Lie algebras.

Definition A symmetric bilinear form B on a Lie algebra $\mathfrak{g}$ is *invariant* if

$$B(\text{ad}_X(Y), Z) + B(Y, \text{ad}_X(Z)) = 0 \tag{102}$$

or equivalently if

$$B([X, Y], Z) = B(X, [Y, Z]). \tag{103}$$

This is the infinitesimal form of the invariance condition $B(\text{Ad}_g(Y), \text{Ad}_g(Z)) = B(Y, Z)$ one would expect on representations of groups.

An invariant form has the following properties.

1. $\text{Ker}(B)$ is an ideal.

2. $B(Z(\mathfrak{g}), [\mathfrak{g}, \mathfrak{g}]) = 0$.

3. If $\mathfrak{j}$ is an ideal, its *orthogonal complement*

$$\mathfrak{j}^\perp = \{X \in \mathfrak{g} : B(J, X) = 0 \forall J \in \mathfrak{j}\} \tag{104}$$

   is also an ideal.

Example Let $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of $\mathfrak{g}$. Then the form

$$B_V(X, Y) = \text{tr}_V(\rho(X)\rho(Y)) \tag{105}$$

is invariant. To verify this it suffices to verify this for $\mathfrak{gl}(V)$, but

$$\text{tr}([X, Y]Z) = \text{tr}(XYZ) - \text{tr}(YXZ) = \text{tr}(X[Y, Z]) = \text{tr}(XYZ) - \text{tr}(XZY) \tag{106}$$

by the cyclic invariance of the trace.

Definition The *Killing form* on a Lie algebra $\mathfrak{g}$ is the invariant form obtained as above from the adjoint representation:

$$B_{\mathfrak{g}}(X, Y) = \text{tr}_{\mathfrak{g}}(\text{ad}_X \text{ad}_Y). \tag{107}$$

Proposition 7.11. *Let $\mathfrak{j} \subseteq \mathfrak{g}$ be an ideal. Then the restriction of the Killing form $B_{\mathfrak{g}}$ to $\mathfrak{j}$ agrees with $B_{\mathfrak{j}}$.*

Proof. The adjoint representation of $\mathfrak{j}$ on $\mathfrak{g}$ has $\mathfrak{j}$ as a subrepresentation. $\mathfrak{j}$ acts trivially on the quotient representation, so the contribution to the trace from the quotient is zero. Hence the only contribution is from the adjoint representation of $\mathfrak{j}$. □

Proposition 7.12. *Under the same hypotheses as in Lie's theorem, $B_V(\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}]) = 0$.*

Proof. By Lie's theorem $\mathfrak{g}$ is conjugate to a subalgebra of $\mathfrak{b}(V)$, hence $[\mathfrak{g}, \mathfrak{g}]$ is conjugate to a subalgebra of $\mathfrak{n}(V)$. The product of an upper-triangular and strictly upper-triangular matrix is strictly upper-triangular, hence has trace zero. □

Theorem 7.13. *(Cartan's criterion) The converse holds. That is, a subalgebra $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ is solvable iff $B_V(\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}]) = 0$.*

Cartan's criterion follows in turn from the following.

Proposition 7.14. *If $B_V(\mathfrak{g}, \mathfrak{g}) = 0$, then $\mathfrak{g}$ is solvable.*

To see this it suffices to observe that under the hypotheses of the nontrivial direction of Cartan's criterion we have

$$B_V([\mathfrak{g}, \mathfrak{g}], [\mathfrak{g}, \mathfrak{g}]) = 0 \tag{108}$$

so $[\mathfrak{g}, \mathfrak{g}]$ is solvable. But then $\mathfrak{g}$ is solvable. It suffices to prove the proposition. First we will need the following lemma.

Lemma 7.15. *(Jordan decomposition) If $X \in End(V)$, where V is a finite-dimensional vector space over an algebraically closed ground field k, then there is a unique decomposition*

$$X = X_s + X_n \tag{109}$$

*where X_s is diagonalizable (*semisimple*), X_n is nilpotent, and $[X_s, X_n] = 0$. Moreover, there exist $p(t), q(t) \in k[t]$ with $p(0) = q(0) = 0$ such that $X_s = p(X), X_n = q(X)$.*

Proof. Let $\lambda_1, ...\lambda_k$ be the eigenvalues of X. Then V is a direct sum of generalized eigenspaces

$$V = \bigoplus_{i=1}^{k} \tilde{V}_{\lambda_k}. \tag{110}$$

Then the restriction of X_s to each generalized eigenspace must be λ_k, and so the restriction of X_n to each generalized eigenspace must be $X - X_s$. This gives uniqueness, and now it suffices to prove existence. To get existence, let $p(t)$ be a polynomial such that

$$p(t) \equiv \lambda_k \bmod (t - \lambda_k)^{\dim \tilde{V}_{\lambda_k}} \tag{111}$$

and, if none of the λ_i are zero, also such that $p(t) \equiv 0 \bmod t$. Such a p exists by the Chinese remainder theorem. We can compute $p(X)$ blockwise, and $p(X)$ restricted to $\tilde{V}_{\lambda_k}$ is λ_k, hence we can take $X_s = p(X), X_n = X - p(X)$ as desired. □

Now we will prove the proposition.

Proof. Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be a Lie subalgebra, with hypotheses as above. Then $X \in \mathfrak{g}$ admits a Jordan decomposition $X_s + X_n$. We readily check that $\text{ad}(X_s)$ is semisimple as an element of $\text{End}(\mathfrak{gl}(V))$ and that $\text{ad}(X_n)$ is nilpotent. Hence $\text{ad}(X)_s = \text{ad}(X_s)$ and $\text{ad}(X)_n = \text{ad}(X_n)$, so we can write $\text{ad}(X_s) = p(\text{ad}(X))$ for some polynomial p. In particular, since $\text{ad}_X(\mathfrak{g}) \subseteq \mathfrak{g}$, we conclude that

$$\text{ad}_{X_s}(\mathfrak{g}) \subseteq \mathfrak{g}. \tag{112}$$

So the adjoint representation extends to semisimple and nilpotent parts.

We want to show that $X \in [\mathfrak{g}, \mathfrak{g}]$ is nilpotent. Then we would be done by Engel's theorem. To see this, write $X = X_s + X_n$ and decompose V as a direct sum of eigenspaces V_{λ_i} of X_s (we'll count the eigenvalues λ_i according to their multiplicities $n_i = \dim V_{\lambda_i}$).

Writing $\mathrm{End}(V) = V^* \otimes V$, the action of X_s restricted to $V^*_{\lambda_i}$ is as multiplication by $-\lambda_i$, hence the decomposition of $\mathrm{End}(V)$ under the action of X_s takes the form

$$V^* \otimes V = \bigoplus_{\mu} \bigoplus_{\lambda_i - \lambda_j = \mu} V^*_{\lambda_j} \otimes V_{\lambda_i}. \tag{113}$$

Define $\overline{X}_s$ to be the matrix with eigenvalues the conjugate of those of X_s (and with the same eigenvectors). The decomposition of $\mathrm{End}(V)$ induces a decomposition

$$\mathfrak{g} = \bigoplus_{\mu} \mathfrak{g}_\mu \tag{114}$$

of $\mathfrak{g}$ into subspaces preserved by X_s. These are therefore also preserved by $\overline{X}_s$, so $\mathrm{ad}(\overline{X}_s)(\mathfrak{g}) \subseteq \mathfrak{g}$.

We compute that

$$B_V(X, \overline{X}_s) = \sum n_i^2 |\lambda_i|^2. \tag{115}$$

By assumption, $X \in [\mathfrak{g}, \mathfrak{g}]$, so we can write $X = \sum_i [Y_i, Z_i]$. Then by invariance

$$B_V(X, \overline{X}_s) = \sum_i B_V([Y_i, Z_i], \overline{X}_s) = \sum_i B_V(Y_i, [Z_i, \overline{X}_s]). \tag{116}$$

But by the hypothesis this is equal to zero. Hence the λ_i must be equal to zero and $X = X_n$ as desired. □

Corollary 7.16. *If $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ is semisimple, then B_V is nondegenerate on $\mathfrak{g}$.*

Proof. The kernel of B_V is an ideal, and by the Cartan criterion it is a solvable ideal. □

Corollary 7.17. *$\mathfrak{g}$ is semisimple iff the Killing form $B_{\mathfrak{g}}$ is nondegenerate.*

Proof. Assume that $B_{\mathfrak{g}}$ is nondegenerate but that $\mathrm{rad}(\mathfrak{g}) \neq 0$. Then $\mathfrak{g}$ has a nonzero abelian ideal, but any such ideal must lie in the kernel of $B_{\mathfrak{g}}$, since as before there is no contribution to the trace from either the ideal or the quotient. □

Definition A Lie algebra $\mathfrak{g}$ is *simple* if $\mathfrak{g}$ is not abelian and has no proper nonzero ideals.

Corollary 7.18. *A semisimple Lie algebra is the direct sum of its simple ideals.*

Proof. Let $\mathfrak{g}$ be semisimple and let $\mathfrak{h}$ be a minimal nonzero ideal. The radical of $\mathfrak{h}$ is $\mathfrak{g}$-invariant, so must vanish, hence $\mathfrak{h}$ is semisimple. The restriction of the Killing form of $\mathfrak{g}$ to $\mathfrak{h}$ is the Killing form of $\mathfrak{h}$, which is nondegenerate on $\mathfrak{h}$, so we can decompose $\mathfrak{g}$ into a direct sum

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^\perp \tag{117}$$

where $\mathfrak{h}^\perp$ is also an ideal. But then this must be a direct sum of Lie algebras, and we conclude that $\mathfrak{g}$ is the direct sum of some simple ideals

$$\mathfrak{g} = \bigoplus_i \mathfrak{g}_i. \tag{118}$$

But then any ideal of $\mathfrak{g}$ must be a direct sum of some of the $\mathfrak{g}_i$ by simplicity. In particular, all of the simple ideals must appear and this decomposition is unique. □

8 Basic representation theory

Let V be a representation of $\mathfrak{g}$, or equivalently a $\mathfrak{g}$-module; we will use the two terms interchangeably. V is not necessarily finite-dimensional. Recall that the direct sum $V \oplus W$ and tensor product $V \otimes W$ of representations gives another representation, with

$$X(v \otimes w) = Xv \otimes w + v \otimes Xw \tag{119}$$

and that the dual V^* of a representation is another representation, with

$$Xf(v) = -f(Xv) \tag{120}$$

(where the negative sign corresponds to taking inverses when defining dual representations of group representations). Moreover, the space $\text{Hom}_k(V, W)$ of k-linear maps between V and W is a representation with

$$X\varphi = [X, \varphi]. \tag{121}$$

If V is finite-dimensional this is just $V^* \otimes W$.

There is a functor

$$\mathfrak{g}\text{-mod} \ni V \mapsto V^{\mathfrak{g}} \in \text{Vect} \tag{122}$$

which sends a representation to its invariant subspace. This functor is left exact; in fact it is just the functor $\text{Hom}_{\mathfrak{g}}(k, V)$ where k is the 1-dimensional trivial $\mathfrak{g}$-module. Hence it has right derived functors, and we will use these to define Lie algebra cohomology. Note that

$$\text{Hom}_{\mathfrak{g}}(V, W) = \text{Hom}_k(V, W)^{\mathfrak{g}}. \tag{123}$$

Definition A $\mathfrak{g}$-module V is *simple* or *irreducible* if V is nonzero and has no proper nonzero submodules.

Note that $\mathfrak{g}$ is a simple Lie algebra if and only if its adjoint representation is irreducible.

Lemma 8.1. *(Schur) Let V, W be simple $\mathfrak{g}$-modules. Then:*

1. *Any $\varphi \in Hom_{\mathfrak{g}}(V, W)$ is either zero or an isomorphism.*

2. *In particular,* $End_{\mathfrak{g}}(V)$ *is a division algebra over* k.

3. *In particular, if* $\dim V < \infty$ *and* k *is algebraically closed then* $End_{\mathfrak{g}}(V) \cong k$.

Proof. The kernel and image of φ are submodules, so they must be all of V resp. W or zero. The second claim is clear. For the third claim it suffices to observe that any finite-dimensional division algebra over an algebraically closed field k is k. □

Definition A $\mathfrak{g}$-module V is *semisimple* or *completely reducible* if the following equivalent conditions are satisfied:

1. Any submodule $W \subseteq V$ has a complement $W' \subseteq V$ such that $V \cong W \oplus W'$.
2. $V = \bigoplus_i V_i$ where each V_i is simple.
3. $V = \sum_j V_j$ where each V_j is simple.

Theorem 8.2. *Any finite-dimensional representation of a semisimple Lie algebra* $\mathfrak{g}$ *is semisimple.*

To prove this theorem we will use a little Lie algebra cohomology.

9 Lie algebra cohomology

Let $\mathfrak{g}$ be a Lie algebra and V a module.

Definition $C^n(\mathfrak{g}, V)$ (*n-cochains* of $\mathfrak{g}$ with coefficients in V) is the module $\text{Hom}_k(\Lambda^n(\mathfrak{g}), V)$ of alternating multilinear functions $\mathfrak{g}^n \to V$. There is a differential $d : C^n(\mathfrak{g}, V) \to C^{n+1}(\mathfrak{g}, V)$ making this a cochain complex given by

$$(d\varphi)(X_1, ...X_{n+1}) = \sum_{1 \le i < j \le n+1} (-1)^{i+j-1} \varphi([X_i, X_j], X_1, ..., \widehat{X_i}, ..., \widehat{X_j}, ..., X_{n+1}) \quad (124)$$

$$+ \sum_i (-1)^i X_i \varphi(X_1, ..., \widehat{X_i}, ..., X_{n+1}) \quad (125)$$

where $\widehat{X}$ indicates that X is omitted.

It is an exercise to check that $d^2 = 0$. It helps to write $d = d_0 + d_1$ where d_0 is the first sum above and d_1 is the second. Then $d_0^2 = 0$ by the Jacobi identity and $d_0 d_1 + d_1 d_0 + d_1^2 = 0$ because V is a module.

Definition The (n^{th}) *cohomology* $H^n(\mathfrak{g}, V)$ of $\mathfrak{g}$ with coefficients in V is the n^{th} cohomology of the above cochain complex. Explicitly, it is the quotient of $\ker(d) \cap C^n(\mathfrak{g}, V)$ (the *n-cocycles* $Z^n(\mathfrak{g}, V)$) by $\text{im}(d) \cap C^n(\mathfrak{g}, V)$ (the *n-coboundaries* $B^n(\mathfrak{g}, V)$).

What is $H^0(\mathfrak{g}, V)$? The beginning of the cochain complex is

$$0 \to V \xrightarrow{d} \mathrm{Hom}_k(\mathfrak{g}, V) \to ... \tag{126}$$

where d acts on V by $(dv)(X) = -Xv$. Hence

$$H^0(\mathfrak{g}, V) \cong V^{\mathfrak{g}} \tag{127}$$

is the invariant submodule. In general Lie algebra cohomology is the derived functor of taking invariants.

What is $H^1(\mathfrak{g}, V)$? The differential $C^1 \to C^2$ has the form

$$(d\varphi)(X, Y) = \varphi([X, Y]) - X\varphi(Y) + Y\varphi(X). \tag{128}$$

If V is a trivial module, the second two terms disappear, so 2-cocycles are maps φ which vanish on the commutator $[\mathfrak{g}, \mathfrak{g}]$ and 2-coboundaries are trivial. Hence in particular

$$H^1(\mathfrak{g}, k) \cong (\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}])^*. \tag{129}$$

What if V is the adjoint representation? Then 2-cocycles are maps $\varphi : \mathfrak{g} \to \mathfrak{g}$ such that

$$\varphi([X, Y]) = [X, \varphi(Y)] + [\varphi(X), Y] \tag{130}$$

hence are precisely derivations. 2-coboundaries are derivations of the form $X \mapsto [Y, X]$, hence are the *inner* derivations. Hence

$$H^1(\mathfrak{g}, \mathfrak{g}) \cong \mathrm{Der}(\mathfrak{g})/\mathrm{Inn}(\mathfrak{g}) \cong \mathrm{Out}(\mathfrak{g}) \tag{131}$$

consists of outer derivations of $\mathfrak{g}$.

H^1 controls extensions of modules in the following sense. Suppose

$$0 \to W \to U \to V \to 0 \tag{132}$$

is a short exact sequence of $\mathfrak{g}$-modules. Any such sequence splits as a short exact sequence of vector spaces, so we can write $U = V \oplus W$ (as vector spaces). The action of $\mathfrak{g}$ on U then takes the form

$$X(v, w) = (Xv, Xw + \varphi(X)v) \tag{133}$$

where $\varphi(X)$ is some linear map $V \to W$, since the action must be compatible with both the inclusion of W and the quotient into V. The condition that this is an action places additional constaints on φ: we compute that

$$\begin{aligned} XY(v, w) &= X(Yv, Yw + \varphi(Y)v) & (134)\\ &= (XYv, XYw + X\varphi(Y)v + \varphi(X)Yv) & (135) \end{aligned}$$

and similarly that

$$YX(v,w) = (YXv, YXw + Y\varphi(X)v, \varphi(Y)Xv) \tag{136}$$

and

$$[X,Y](v,w) = ([X,Y]v, [X,Y]w + \varphi([X,Y])v). \tag{137}$$

The condition that we have an action gives $(XY - YX)(v,w) = [X,Y](v,w)$, which gives

$$\varphi([X,Y])v = [X, \varphi(Y)]v + [Y, \varphi(X)]v. \tag{138}$$

This is precisely the condition that $\varphi \in Z^1(\mathfrak{g}, \text{Hom}_k(V,W))$ is a 1-cocycle with coefficients in $\text{Hom}_k(V,W)$. But there is also a natural equivalence relation on the possible choices of φ: namely, by choosing a different splitting of U, we we can apply automorphisms of U of the form $R(v,w) = (v, w + \theta(v))$, where $\theta \in \text{Hom}_k(V,W)$. We want to say that two actions X_φ, X_ψ associated to two cocycles φ, ψ are equivalent if

$$R \circ X_\varphi = X_\psi \circ R. \tag{139}$$

Expanding gives

$$(Xv, \theta(Xv) + \varphi(X)v + Xw) = (Xv, X\theta(v) + Xw + \psi(X)v) \tag{140}$$

or

$$\varphi(X) - \psi(X) = [X, \theta] = d\theta(X). \tag{141}$$

Hence, up to equivalence, short exact sequences with first and last term V and W are classified by $H^1(\mathfrak{g}, \text{Hom}_k(V,W))$.

9.1 Semisimplicity

Theorem 9.1. *The following conditions on a Lie algebra $\mathfrak{g}$ are equivalent:*

1. *Every finite-dimensional $\mathfrak{g}$-module is semisimple (completely reducible).*
2. *$H^1(\mathfrak{g}, M) = 0$ for any finite-dimensional $\mathfrak{g}$-module M.*
3. *$H^1(\mathfrak{g}, M) = 0$ for any finite-dimensional simple $\mathfrak{g}$-module M.*

Proof. $1 \Leftrightarrow 2$ is clear by the computation we did above: both are equivalent to the claim that any short exact sequence of finite-dimensional $\mathfrak{g}$-modules splits. $2 \Rightarrow 3$ is straightforward. It remains to show $3 \Rightarrow 2$, which we can do by using long exact sequences as follows. If

$$0 \to K \xrightarrow{i} M \xrightarrow{p} N \to 0 \tag{142}$$

is a short exact sequence of modules, then we get a long exact sequence

$$... \to H^n(\mathfrak{g}, K) \xrightarrow{i^*} H^n(\mathfrak{g}, M) \xrightarrow{p^*} H^n(\mathfrak{g}, N) \xrightarrow{\partial} H^{n+1}(\mathfrak{g}, K) \to ... \tag{143}$$

which we can use to show $3 \Rightarrow 2$ by induction on the dimension. □

Below we will assume that we are working over a field k of characteristic zero.

Theorem 9.2. *Let $\mathfrak{g}$ be a semisimple Lie algebra and V a finite-dimensional irreducible faithful representation of $\mathfrak{g}$. Then $H^i(\mathfrak{g}, V) = 0$ for all $i \geq 0$.*

Proof. Assume first that k is algebraically closed.

The inclusion $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ gives us an invariant bilinear form $B_V : \mathfrak{g} \times \mathfrak{g} \to k$. Since B_V is non-degenerate, if $e_1, ..., e_n$ is any basis of $\mathfrak{g}$ we can associate to it a dual basis $f_1, ..., f_n$ such that $B_V(e_i, f_j) = \delta_{ij}$. Now we can define the *Casimir operator*

$$C = \sum_{i=1}^{n} e_i f_i \in \text{End}_k(V). \tag{144}$$

It is an exercise to show that the Casimir operator does not depend on the choice of basis e_i.

Lemma 9.3. *$C \in End_{\mathfrak{g}}(V)$. That is, C commutes with the action of $\mathfrak{g}$.*

Proof. For $X \in \mathfrak{g}$ we compute that

$$\begin{aligned}
[X, \sum_i e_i f_i] &= \sum_i \Big([X, e_i] f_i + \sum e_i [X, f_i] \Big) && (145)\\
&= \sum_{i,j} B_V([X, e_i], f_j) e_j f_i + \sum_{i,j} e_i B_V([X, f_i], e_j) f_j && (146)\\
&= \sum_{i,j} \left(B_V([X, e_i], f_j) + B_V([X, f_j], e_i) \right) e_j f_i && (147)
\end{aligned}$$

which vanishes since B_V is invariant. Here we use the fact that

$$\begin{aligned}
X &= \sum B_V(X, f_i) e_i && (148)\\
&= \sum B_V(X, e_i) f_i. && (149)
\end{aligned}$$

□

Since k is algebraically closed, C acts as a scalar by Schur's lemma. By taking the trace we can show that this scalar is $\frac{\dim \mathfrak{g}}{\dim V}$; in particular, it is nonzero.

We will now prove the theorem by writing down an explicit contracting homotopy $h : C^p(\mathfrak{g}, V) \to C^{p-1}(\mathfrak{g}, V)$ as follows:

$$h\varphi(X_1, ...X_{p-1}) = \sum_i f_i \varphi(e_i, X_1, ...X_{p-1}). \tag{150}$$

Lemma 9.4. $(hd + dh)\varphi = -C\varphi$.

Proof. Exercise. The calculation is similar to the calculation we did above. □

It follows that $C^\bullet(\mathfrak{g}, V)$ is contractible. □

Theorem 9.5. *The above is true without the assumption that V is faithful.*

Proof. If V is such a module, we can write $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ where $\mathfrak{m}$ is the kernel of the representation and $\mathfrak{h}$ acts faithfully on V (because we can take the complement with respect to the Killing form). We have

$$C^k(\mathfrak{g}, V) = \bigoplus_{p+q=k} C^p(\mathfrak{h}, V) \otimes C^q(\mathfrak{m}, k). \tag{151}$$

The RHS can be upgraded to a bicomplex $C^{p,q}$, with differentials $d_{1,0} : C^{p,q} \to C^{p+1,q}$ and $d_{0,1} : C^{p,q} \to C^{p,q+1}$, and then the LHS is its total complex equipped with the differential $d = d_{1,0} + d_{0,1}$.

Let $\varphi \in C^k(\mathfrak{g}, V)$ be a cocycle. Write $\varphi = \sum \varphi_p$ where $\varphi_p \in C^p(\mathfrak{h}, V) \otimes C^q(\mathfrak{m}, k)$. Choose p maximal such that $\varphi \neq 0$. Since we know that $C^\bullet(\mathfrak{h}, V)$ is contractible, $d\varphi = 0$ implies that $d_{1,0}\varphi_p = 0$, hence $\varphi_p = d_{1,0}(\psi_{p-1})$. Then $\varphi - d(\psi_{p-1})$ is also a cocycle, but with a smaller value of p. Eventually there is no such p. □

Corollary 9.6. *(Weyl) Every finite-dimensional representation of a semisimple Lie algebra is completely reducible.*

Proof. We now know that $H^1(\mathfrak{g}, V) = 0$ for $i \geq 0$ and V a nontrivial simple module. In the trivial case $H^1(\mathfrak{g}, k) \cong (\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}])^* = 0$, so we conclude that $H^1(\mathfrak{g}, V) = 0$ for any finite-dimensional simple V, and this gives semisimplicity of the finite-dimensional representation theory. □

Corollary 9.7. *The above is true without the assumption that k is algebraically closed.*

Proof. It suffices to observe that

$$H^i(\mathfrak{g}, V) \otimes \bar{k} \cong H^i(\mathfrak{g} \otimes_k \bar{k}, V \otimes_k \bar{k}). \tag{152}$$

□

Corollary 9.8. *If $\mathfrak{g}$ is a semisimple Lie algebra (over $\mathbb{R}$), then there exists a Lie group G with Lie algebra $\mathfrak{g}$.*

Proof. If $\mathfrak{g}$ is semisimple then

$$H^1(\mathfrak{g}, \mathfrak{g}) = \mathrm{Der}(\mathfrak{g})/\mathrm{Inn}(\mathfrak{g}) = 0 \tag{153}$$

from which it follows that $\mathrm{Der}(\mathfrak{g}) \cong \mathfrak{g}$. But now $\mathrm{Der}(\mathfrak{g})$ is the Lie algebra of the Lie group $\mathrm{Aut}(\mathfrak{g})$. □

Corollary 9.9. *If $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ is semisimple and $X \in \mathfrak{g}$ has Jordan decomposition $X_s + X_n$, then $X_s, X_n \in \mathfrak{g}$.*

Proof. $\mathrm{ad}(X_s) \in \mathrm{Der}(\mathfrak{g})$, hence must be $\mathrm{ad}(Y)$ for some $Y \in \mathfrak{g}$. □

Corollary 9.10. *If $\mathfrak{g}$ is semisimple, then $X \in \mathfrak{g}$ is semisimple (resp. nilpotent) iff ad(X) is semisimple (resp. nilpotent).*

Corollary 9.11. *If $\mathfrak{g}$ is semisimple, $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is a representation, and $X \in \mathfrak{g}$ is semisimple, then $\rho(X)$ is semisimple.*

9.2 Extensions

Let

$$0 \to \mathfrak{a} \to \widehat{\mathfrak{g}} \to \mathfrak{g} \to 0 \tag{154}$$

be a short exact sequence of Lie algebras with $\mathfrak{a}$ abelian. The the adjoint representation restricts to a representation $\widehat{\mathfrak{g}} \to \mathfrak{gl}(\mathfrak{a})$ which factors through a representation $\mathfrak{g} \to \mathfrak{gl}(\mathfrak{a})$. Thus studying such short exact sequences is equivalent to studying $\mathfrak{g}$-modules $\mathfrak{a}$ together with brackets on $\widehat{\mathfrak{g}} = \mathfrak{g} \oplus \mathfrak{a}$ (here the vector space direct sum) compatible with the action. Explicitly, the action should have the form

$$[(X_1, Y_1), (X_2, Y_2)] = ([X_1, X_2], X_1 Y_2 - X_2 Y_1 + \varphi(X_1, X_2)) \tag{155}$$

where φ is some bilinear map. The Lie algebra axioms then impose conditions on φ, and it turns out that these conditions say precisely that φ is a 2-cocycle in $Z^2(\mathfrak{g}, \mathfrak{a})$. This follows from the Jacobi identity. Two such cocycles give isomorphic extensions if and only if they are related by a map of the form $R(X, Y) = (X, Y - \theta(X))$, and we can compute that this turns out to be the case if and only if $\varphi_1 - \varphi_2 = d\theta$ is a 2-coboundary in $B^2(\mathfrak{g}, \mathfrak{a})$.

Hence to any short exact sequence above we can associate an element of $H^2(\mathfrak{g}, \mathfrak{a})$, and this element classifies extensions of $\mathfrak{g}$ by $\mathfrak{a}$ (provided that the action is fixed). In particular, the short exact sequence splits as a semidirect product $\mathfrak{g} \rtimes \mathfrak{a}$ if and only if this element is equal to zero.

Proposition 9.12. *(Whitehead) If $\mathfrak{g}$ is semisimple and V is a finite-dimensional $\mathfrak{g}$-module, then $H^2(\mathfrak{g}, V) = 0$.*

Proof. By previous results, it suffices to prove this in the case $V = k$. Here we are dealing with a *central* extension

$$0 \to k \to \widehat{\mathfrak{g}} \to \mathfrak{g} \to 0. \tag{156}$$

The adjoint representation factors through a representation $\mathfrak{g} \to \mathfrak{gl}(\widehat{\mathfrak{g}})$ by centrality, hence it is completely reducible, so we can write $\widehat{\mathfrak{g}} \cong k \oplus \mathfrak{g}$ as $\mathfrak{g}$-modules and hence as ideals.

An alternative proof proceeds by constructing an isomorphism $H^1(\mathfrak{g}, \mathfrak{g}) \cong H^1(\mathfrak{g}, \mathfrak{g}^*) \cong H^2(\mathfrak{g}, k)$ using the Killing form. □

In particular, any extension of a semisimple Lie algebra is a semidirect product.
Once again the characteristic of the ground field is zero.

Theorem 9.13. *(Levi) Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. Then $\mathfrak{g}$ is a semidirect product $\mathfrak{g}_s \rtimes rad(\mathfrak{g})$ where $\mathfrak{g}_s \cong \mathfrak{g}/rad(\mathfrak{g})$ is semisimple (a* Levi subalgebra *of $\mathfrak{g}$). Any two Levi subalgebras are conjugate by the Lie subgroup of $Aut(\mathfrak{g})$ corresponding to the adjoint representation.*

Proof. We induct on the length of the derived series of $\mathfrak{g}$. We already know this in the case that $\text{rad}(\mathfrak{g})$ is abelian. In general, let $\mathfrak{r} = \text{rad}(\mathfrak{g})$. The short exact sequence

$$0 \to \mathfrak{r}/[\mathfrak{r},\mathfrak{r}] \to \mathfrak{g}/[\mathfrak{r},\mathfrak{r}] \to \mathfrak{g}_s \to 0 \tag{157}$$

splits because $\mathfrak{r}/[\mathfrak{r},\mathfrak{r}]$ is abelian. Let $p : \mathfrak{g} \to \mathfrak{g}/[\mathfrak{r},\mathfrak{r}]$ be the quotient map. Then the splitting above lets us think of $\mathfrak{g}_s$ as living in $\mathfrak{g}/[\mathfrak{r},\mathfrak{r}]$, and from here we can write down a short exact sequence

$$0 \to [\mathfrak{r},\mathfrak{r}] \to p^{-1}(\mathfrak{g}_s) \to \mathfrak{g}_s \to 0. \tag{158}$$

By the inductive hypothesis this splits.

Let $\mathfrak{g}_s, \mathfrak{g}_s'$ be two Levi subalgebras. The particular form of the isomorphism $\mathfrak{g}_s \rtimes \mathfrak{r} \cong \mathfrak{g}_s' \rtimes \mathfrak{r}$ implies that

$$\mathfrak{g}_s' = \{X + \varphi(X) | X \in \mathfrak{g}_s\} \tag{159}$$

for some $\varphi \in \text{Hom}_k(\mathfrak{g}_s, \mathfrak{r})$. We again induct. If $\mathfrak{r}$ is abelian, then

$$[X,Y] + \varphi([X,Y]) = [X + \varphi(X), Y + \varphi(Y)] \tag{160}$$

implies that φ is a cocycle in $Z^1(\mathfrak{g}_s, \mathfrak{r})$. Since H^1 vanishes, φ is a coboundary, hence has the form $\varphi(X) = [\xi, X]$ where $\xi \in \mathfrak{r}$. But then $\mathfrak{g}_s' = \exp(\text{ad}_\xi)\mathfrak{g}_s$.

It follows that in the general case we have

$$\mathfrak{g}_s \cong g(\mathfrak{g}_s') \bmod [\mathfrak{r},\mathfrak{r}]. \tag{161}$$

But then the induction works the same as before. □

Corollary 9.14. *(Lie's third theorem) Every finite-dimensional Lie algebra $\mathfrak{g}$ over $\mathbb{R}$ (resp. $\mathbb{C}$) is the Lie algebra of a real (resp. complex) Lie group G.*

Proof. We already know this for solvable Lie algebras and for semisimple Lie algebras. The general case follows by Levi decomposition. □

Example $\text{SL}_2(\mathbb{R})$ is a Lie group. It is not simply connected: its fundamental group is $\mathbb{Z}$. Its universal cover is the simply connected Lie group associated to $\mathfrak{sl}_2(\mathbb{R})$, and it is a larger group $\widetilde{\text{SL}_2}(\mathbb{R})$ fitting into a short exact sequence

$$0 \to \mathbb{Z} \to \widetilde{\mathrm{SL}_2}(\mathbb{R}) \to \mathrm{SL}_2(\mathbb{R}) \to 0. \tag{162}$$

The group $\widetilde{\mathrm{SL}_2}(\mathbb{R})$ does not have a faithful (finite-dimensional) representation! To see this, note that by the semisimplicity of $\mathfrak{sl}_2(\mathbb{R})$, any complex representation V is a direct sum of irreducible representations $V_1 \oplus ... \oplus V_n$. By Schur's lemma, a generator γ of $\mathbb{Z}$ acts by some scalar λ_i on each V_i. But since $\mathfrak{sl}_2(\mathbb{R})$ is perfect (equal to its own commutator), any representation lands in traceless matrices, so any representation of the group lands in SL. It follows that $\lambda_i^{\dim V_i} = 1$, so the action of $\mathbb{Z}$ cannot be faithful.

As it turns out, this does not happen over $\mathbb{C}$.

10 Universal enveloping algebras

Below by algebra we mean associative algebra.

Any algebra A is a Lie algebra under the commutator bracket $[a, b] = ab - ba$. In particular, if V is a vector space then $\mathrm{End}(V)$ is a Lie algebra which we have been caling $\mathfrak{gl}(V)$. This defines a functor from algebras to Lie algebras, and the universal enveloping algebra construction is the left adjoint of this functor.

Definition An algebra $U(\mathfrak{g})$ is the *universal enveloping algebra* of $\mathfrak{g}$ if it is equipped with a morphism $\varepsilon : \mathfrak{g} \to U(\mathfrak{g})$ of Lie algebras which is universal in the sense that any homomorphism $\mathfrak{g} \to A$ of Lie algebras, where A is an algebra, factors uniquely through ε.

$U(\mathfrak{g})$ is unique up to unique isomorphism. It can be explicitly constructed as follows: we first take the tensor algebra

$$T(\mathfrak{g}) = \bigoplus_{n \ge 0} \mathfrak{g}^{\otimes n} \cong \bigoplus_{n \ge 0} T^n(\mathfrak{g}), \tag{163}$$

which is the free algebra on $\mathfrak{g}$ regarded as a vector space, and then quotient by the ideal generated by the commutation relations $XY - YX = [X, Y]$, where $X, Y \in \mathfrak{g}$.

There is a canonical equivalence $\mathfrak{g}$-mod $\cong U(\mathfrak{g}$-mod which in principle reduces the study of the representation theory of Lie algebras to the study of the module theory of noncommutative rings. This is because morphisms $\mathfrak{g} \to \mathfrak{gl}(V)$ of Lie algebras can be naturally identified with morphisms $U(\mathfrak{g}) \to \mathrm{End}(V)$ of algebras.

By construction we have $U(\mathfrak{g} \oplus \mathfrak{h}) \cong U(\mathfrak{g}) \otimes U(\mathfrak{h})$. The diagonal embedding $\Delta : \mathfrak{g} \mapsto \mathfrak{g} \oplus \mathfrak{g}$ induces a map $\Delta : U(\mathfrak{g}) \to U(\mathfrak{g}) \otimes U(\mathfrak{g})$ given by $\Delta(X) = X \otimes 1 + 1 \otimes X$ on elements $X \in \mathfrak{g}$. This map makes $U(\mathfrak{g})$ a bialgebra.

$U(\mathfrak{g})$ has three natural $\mathfrak{g}$-module structures. The first one is given by left multiplication $\varepsilon(X)Y$, the second one is given by right multiplication $-Y\varepsilon(X)$ (we take negatives to convert left modules to right modules), and the third one is given by the adjoint action $\varepsilon(X)Y - Y\varepsilon(X)$.

Example Let $\mathfrak{g}$ be abelian. Then $U(\mathfrak{g}) \cong S(\mathfrak{g})$, or equivalently the algebra of polynomial functions on $\mathfrak{g}^*$.

In general $U(\mathfrak{g})$ can be thought of as a noncommutative deformation of $S(\mathfrak{g})$ as follows. We can define

$$U_t(\mathfrak{g}) \cong T(\mathfrak{g})/(XY - YX - t[X,Y]) \tag{164}$$

for all t. When $t \neq 0$ these are all isomorphic to the universal enveloping algebra, but when $t = 0$ we get the symmetric algebra.

Definition Let A be an algebra. A *filtration* on A is an increasing sequence $F^0(A) \subseteq F^1(A) \subseteq ...$ of subspaces of A with $A = \bigcup F^k(A)$ such that $F^i(A)F^j(A) \subseteq F^{i+j}(A)$. A *filtered algebra* is an algebra together with a filtration. The *associated graded* of a filtered algebra is

$$\text{Gr}(A) = \bigoplus_{k=0}^{\infty} F^k(A)/F^{k-1}(A). \tag{165}$$

The associated graded is naturally a graded algebra with product inherited from that of A. There is a natural map $A \to \text{Gr}(A)$ which is an isomorphism of vector spaces. Using this isomorphism we can think of the multiplication on A as a multiplication on $\text{Gr}(A)$ which begins with its given multiplication but continues with lower order terms.

Example Let V be any vector space. The subspaces

$$F^k(T(V)) = \bigoplus_{i=0}^{k} V^{\otimes i} \tag{166}$$

form a filtration of $T(V)$. Then we have a canonical isomorphism $T(V) \cong \text{Gr}(T(V))$.

Example The universal enveloping algebra $U(\mathfrak{g})$ inherits a filtration from $T(\mathfrak{g})$ as above. Its associated graded is never isomorphic to $U(\mathfrak{g})$ if $\mathfrak{g}$ is nonabelian. In this case $U(\mathfrak{g})$ is noncommutative, but $\text{Gr}(U(\mathfrak{g}))$ is always commutative. It suffices to check this on generators $X, Y \in \mathfrak{g}$. Here

$$XY = YX + [X,Y] \tag{167}$$

where $XY, YX \in F^2$ but $[X,Y] \in F^1$, hence $XY \equiv YX \bmod F^1$.

Below we make no assumptions about the base field or about the dimension of $\mathfrak{g}$.

Theorem 10.1. *(Poincaré-Birkhoff-Witt) Gr$(U(\mathfrak{g})) \cong S(\mathfrak{g})$.*

Proof. Choose a totally ordered basis $\{X_i : i \in I\}$ of $\mathfrak{g}$. This induces a basis of $S(\mathfrak{g})$ consisting of monomials $X_{i_1}...X_{i_k}$ where $i_1 \leq i_2 \leq ... \leq i_k$. The PBW theorem is equivalent to the assertion that these products also serve as a basis of $\text{Gr}(U(\mathfrak{g}))$, hence, by transporting along the map $U(\mathfrak{g}) \to \text{Gr}(U(\mathfrak{g}))$, a basis of $U(\mathfrak{g})$. This is what we will prove.

First we want to show that the monomials span $U(\mathfrak{g})$. If $X_j X_{i_1} ... X_{i_k} X_{i_{k+1}} ... X_{i_n}$ is such that $i_k \leq j \leq i_{k+1}$, then we we can commute it past to get

$$X_{i_1} X_j X_{i_2} ... X_{i_n} + [X_j, X_i] X_{i_2} ... X_{i_n}. \tag{168}$$

Repeating this k times we get that this is equivalent to $X_{i_1} ... X_{i_k} X_j X_{i_{k+1}} ... X_{i_n}$ modulo terms of lower order. So by induction on the order we get spanning.

The hard part is to show that the monomials are linearly independent. To do this we will construct a $\mathfrak{g}$-module structure on $S(\mathfrak{g})$. We will do this inductively and using linearity. We will also denote the action by $X \cdot v$ and we will distinguish elements $X \in \mathfrak{g}$ from elements $\overline{X} \in S(\mathfrak{g})$. First, define

$$X_i \cdot v = \overline{X_i}. \tag{169}$$

Next, define

$$X_i \cdot X_j \cdot v = \begin{cases} \overline{X_i X_j} v \text{ if } i \leq j \\ X_j \cdot X_i \cdot v + [X_i, X_j] \cdot v \text{ if } i > j \end{cases}. \tag{170}$$

Secretly we are reconstructing the left action of $U(\mathfrak{g})$ on itself. But the hard part now is to check that this is actually an action, or equivalently that

$$X_i \cdot X_j \cdot v - X_j \cdot X_i \cdot v = [X_i, X_j] \cdot v. \tag{171}$$

We will check this by inducting on the degree of v and using linearity. This reduces to checking that

$$X_i \cdot X_j \cdot \overline{X_k} v - X_j \cdot X_i \cdot \overline{X_k} v = [X_i, X_j] \cdot \overline{X_k} v. \tag{172}$$

There are three cases $i \leq j \leq k, i \leq k \leq j$, and $k \leq i \leq j$. We will only explicitly write down what happens in the third case. Here we get

$$\begin{aligned} X_j \cdot X_i \cdot \overline{X_k} v &= X_j \cdot \overline{X_i X_k} v & (173) \\ &= X_i \cdot X_j \cdot \overline{X_k} v + [X_j, X_i] \cdot \overline{X_k} v & (174) \end{aligned}$$

where $\overline{X_k} v$ is already a monomial. But in general $X_k v = \overline{X_k} v +$ lower order terms, and by the inductive hypothesis everything is fine for lower order terms, so we conclude that

$$X_j \cdot X_i \cdot X_k \cdot v - X_j \cdot X_i \cdot X_k \cdot v = [X_i, X_j] \cdot X_k \cdot v \tag{175}$$

regardless of the relationship between X_k and v. Of the three cyclic permutations of this identity, we know two of them by the cases above. It suffices now to verify that the sum of the three identities holds. The LHS of this sum is

$$X_k \cdot [X_i, X_j] \cdot v + X_i \cdot [X_j, X_k] \cdot v + X_j \cdot [X_k, X_i] v. \tag{176}$$

By the inductive hypothesis we can write this as $[X_k, [X_i, X_j]] \cdot v + [X_i, X_j] \cdot X_k \cdot v +$ cyclic permutations, and when we add these cyclic permutations the triple commutator disappears by the Jacobi identity. What remains is the RHS of the sum, so we have verified the identity.

Now suppose there is a linear dependence $\sum a_{i_1 ... i_k} X_{i_1} ... X_{i_k} = 0$ among the monomials. Then

$$\sum a_{i_1 ... i_k} X_{i_1} \cdot ... \cdot X_{i_k} \cdot 1 = \sum a_{i_1 ... i_k} \overline{X_{i_1} ... X_{i_k}} = 0 \tag{177}$$

implies that the coefficients $a_{i_1 ... i_k}$ are all equal to 1 as desired. □

In particular, the natural map $\varepsilon : \mathfrak{g} \to U(\mathfrak{g})$ is an injection. From now on we will regard $\mathfrak{g}$ as a subspace of $U(\mathfrak{g})$.

Various properties of a filtered ring A are inherited from its associated graded. For example, if $\text{Gr}(A)$ has no zero divisors or is left or right Noetherian, then the same is true of A. In particular $U(\mathfrak{g})$ has no zero divisors and is left and right Noetherian.

If $\mathfrak{h} \subseteq \mathfrak{g}$ is a Lie subalgebra there is an inclusion $U(\mathfrak{h}) \subseteq U(\mathfrak{g})$. In fact $U(\mathfrak{g})$ is a free $U(\mathfrak{h})$-module; this follows from a suitable choice of PBW basis, which allows us to write $U(\mathfrak{g}) \cong U(\mathfrak{h}) \otimes S(\mathfrak{g}/\mathfrak{h})$ as a $U(\mathfrak{h})$-module.

Let $f : A \to B$ be a morphism of rings. Then there is a restriction functor $\text{Res} : B\text{-mod} \to A\text{-mod}$ and an induction functor

$$\text{Ind} : A\text{-mod} \ni M \mapsto B \otimes_A M \in B\text{-mod} \tag{178}$$

which is its left adjoint; that is,

$$\text{Hom}_B(\text{Ind}(M), N) \cong \text{Hom}_A(M, \text{Res}(N)). \tag{179}$$

In the particular case that f is an inclusion $U(\mathfrak{h}) \to U(\mathfrak{g})$, the isomorphism above gives

$$\text{Ind}(M) \cong U(\mathfrak{g}) \otimes_{U(\mathfrak{h})} M \cong S(\mathfrak{g}/\mathfrak{h}) \otimes_k M \tag{180}$$

(k the base field). Even if M is finite-dimensional this will be infinite-dimensional in general, but it is the appropriate notion of induction for representations of Lie algebras.

Example Consider $\mathfrak{sl}_2$ with basis H, X, Y as usual. Let $\mathfrak{b}$ be the Borel subalgebra, spanned by H and X. This is solvable, so its simple modules are all 1-dimensional. Write $C_\lambda = kv$ for the module spanned by a vector v satisfying $Hv = \lambda v, Xv = 0$, and let

$$M_\lambda = U(\mathfrak{g}) \otimes_{U(\mathfrak{b})} C_\lambda \cong S(kY) \otimes_k C_\lambda. \tag{181}$$

What is the module structure on this? We have $Y(Y^n \otimes v) = Y^{n+1} \otimes v$ and

$$\begin{aligned} H(Y^n \otimes v) &= \text{ad}_H(Y^n) \otimes v + Y^n \otimes Hv && (182)\\ &= (-2nY^n) \otimes V + Y^n \otimes \lambda v && (183)\\ &= (\lambda - 2n) Y^n \otimes v && (184) \end{aligned}$$

where we used the Leibniz rule

$$\mathrm{ad}_H(XY) = \mathrm{ad}_H(X)Y + X\mathrm{ad}_H(Y). \tag{185}$$

Finally we have

$$X(Y^n \otimes v) = \mathrm{ad}_X(Y^n) \otimes v + Y^n \otimes Xv. \tag{186}$$

The second term vanishes. To compute the first term we apply the Leibniz rule again, using the fact that $\mathrm{ad}_X(Y) = H$, to get

$$\begin{aligned}\mathrm{ad}_X(Y^n) \otimes v &= \sum_{k=0}^{n-1} Y^k H Y^{n-k-1} \otimes v & (187)\\ &= n(\lambda - (n-1))Y^{n-1} \otimes v. & (188)\end{aligned}$$

If λ is not a non-negative integer, the action of X eventually reproduces v starting with any vector, and so M_λ is irreducible. If λ is a non-negative integer m, the action of X eventually annihilates a vector. The vector we obtain before annihilation gives an inclusion $M_{\lambda-2m-2} \to M_\lambda$. The quotient V_λ is irreducible, and any finite-dimensional irreducible $\mathfrak{sl}_2$-module is of this form.

To see this, let V be such a module. Then V contains an eigenvector v for the action of $\mathfrak{b}$. This gives a nontrivial map $C_\lambda \to V$ of $U(\mathfrak{b})$-modules, hence by adjointness a nontrivial map $M_\lambda \to V$ of $U(\mathfrak{sl}_2)$-modules, and V_λ is the unique nontrivial finite-dimensional quotient of M_λ.

The above discussion can be summarized as follows.

Proposition 10.2. *The irreducible finite-dimensional representations of $\mathfrak{sl}_2$ are in bijection with the non-negative integers. On the representation V_n corresponding to n, H acts with eigenvalues $n, n-2, ..., -n+2, -n$. In particular,* $\dim V_n = n + 1$.

It turns out that $V_n = S^n(V_1)$.

10.1 Free resolutions

We want to resolve the trivial $\mathfrak{g}$-module k as a $U(\mathfrak{g})$-module, by free $U(\mathfrak{g})$-modules. It turns out that there exists a particularly nice such resolution, the *Koszul complex*

$$\cdots \to U(\mathfrak{g}) \otimes \Lambda^2(\mathfrak{g}) \to U(\mathfrak{g}) \otimes \mathfrak{g} \to U(\mathfrak{g}) \to k \to 0. \tag{189}$$

The differentials in this resolution are given by

$$\partial(Y \otimes X_1 \wedge ... \wedge X_k) = \sum_{i=1}^{k} (-1)^i Y X_i \wedge X_1 \wedge ... \wedge \widehat{X_i} \otimes X_n \tag{190}$$

$$+ \sum_{1 \le i < j \le k} (-1)^{i+j-1} Y \otimes [X_i, X_j] \wedge ... \wedge \widehat{X_i} \wedge ... \wedge \widehat{X_j} \wedge ... \wedge X_n \tag{191}$$

There are three things to verify: ∂ is a morphism of $\mathfrak{g}$-modules, $\partial^2 = 0$, and the complex is exact. We will verify the third thing. First, note that the filtration $F^d(U(\mathfrak{g}))$ on $U(\mathfrak{g})$ induces a filtration $F^d(K) = F^d(U(\mathfrak{g})) \otimes \Lambda^\bullet(\mathfrak{g})$ on the Koszul complex K. Then K has an associated graded complex $\text{Gr}(K)$ on which an associated graded version $\text{Gr}(\partial) = \delta$ of ∂ acts. This is precisely the Koszul complex $S(\mathfrak{g}) \otimes \Lambda^\bullet(\mathfrak{g})$, which is well-known to be exact.

To see this, we break up δ using a second grading. Its components have the form $S^p \otimes \Lambda^q \to S^{p+1} \otimes \Lambda^{q-1}$, and we can write down a homotopy going the other way of the form

$$h(x_1 ... x_{p+1} \otimes y_1 \wedge ... \wedge y_{q-1}) = \sum_i (x_1 ... \widehat{x_i} ... x_{p+1} \otimes x_1 \wedge y_1 \wedge ... \wedge y_{q-1}). \tag{192}$$

Then $h\partial + \partial h = (p+q)\text{id}$. But by induction on the degree, if the associated graded of a filtered complex is exact then so is the original complex.

Another way to state this argument is to use $U_t(\mathfrak{g})$, the version of the universal enveloping algebra depending on a parameter, to write down a one-parameter family of complexes. When $t = 0$ this is the commutative Koszul complex, and when $t \neq 0$ this is isomorphic to the noncommutative Koszul complex. When we do this we are studying $\text{Ker}(\partial_t)/\text{Im}(\partial_t)$ where ∂_t is a linear operator depending polynomially on a parameter t. But the dimension of this can only increase at a particular value of t (relative to its value in a neighborhood), not decrease; in other words it is upper-semicontinuous. Hence to verify that there is no cohomology anywhere it suffices to verify it at $t = 0$.

Now we can compute the derived functor $\text{Ext}^i_{U(\mathfrak{g})}(k, V)$ by resolving k using the Koszul resolution, applying $\text{Hom}_{U(\mathfrak{g})}(-, V)$, and taking cohomology. But $\text{Hom}_{U(\mathfrak{g})}(U(\mathfrak{g}) \otimes \Lambda^k(\mathfrak{g}), V) \cong \text{Hom}(\Lambda^k(\mathfrak{g}), V)$; moreover, the induced differential agrees with the differential we wrote down earlier to define Lie algebra cohomology. Hence

$$H^i(\mathfrak{g}, V) \cong \text{Ext}^i_{U(\mathfrak{g})}(k, V). \tag{193}$$

It follows that $H^i(\mathfrak{g}, V)$ are the derived functors of invariants $H^0(\mathfrak{g}, V) \cong \text{Hom}_{\mathfrak{g}}(\mathfrak{g}, V) \cong V^{\mathfrak{g}}$. What we did earlier involving extensions amounted to showing this identity when $i = 1$ by hand, after noting that

$$\text{Ext}^i_{U(\mathfrak{g})}(M, N) \cong \text{Ext}^i_{U(\mathfrak{g})}(k, M^* \otimes N) \tag{194}$$

if M is finite-dimensional.

10.2 Bialgebras

The diagonal embedding $\mathfrak{g} \to \mathfrak{g} \oplus \mathfrak{g}$ induces a comultiplication $\Delta : U(\mathfrak{g}) \to U(\mathfrak{g}) \otimes U(\mathfrak{g})$ given by $\Delta(X) = X \otimes 1 + 1 \otimes X$ on elements $X \in \mathfrak{g}$. Together with this comultiplication $U(\mathfrak{g})$ becomes a bialgebra.

Definition A is a *bialgebra* if it is an algebra equipped with a comultiplication $\Delta : A \to A \otimes A$ such that $\Delta(ab) = \Delta(a)\Delta(b)$ and such that Δ is coassociative in the sense that the diagram

$$\begin{array}{ccc} A & \xrightarrow{\Delta} & A \otimes A \\ {\scriptstyle\Delta}\downarrow & & \downarrow{\scriptstyle\Delta\otimes\mathrm{id}} \\ A \otimes A & \xrightarrow{\mathrm{id}\otimes\Delta} & A \otimes A \otimes A \end{array} \tag{195}$$

commutes.

(We also need a counit $\varepsilon : A \to k$ but this is less important for the time being. For $U(\mathfrak{g})$ this is the augmentation map $U(\mathfrak{g}) \to k$ which kills all positive-degree terms.)

Definition Let A be a bialgebra. An element $x \in A$ is *primitive* or *Lie-like* if $\Delta x = x \otimes 1 + 1 \otimes x$. The primitive elements of A are denoted $\mathrm{Prim}(A)$.

The subspace of primitive elements of a bialgebra is closed under the commutator bracket, making it a Lie algebra. To see this, we compute that

$$\Delta[X,Y] = \Delta X \Delta Y - \Delta Y \Delta X \tag{196}$$
$$= (X \otimes 1 + 1 \otimes X)(Y \otimes 1 + 1 \otimes Y) - (Y \otimes 1 + 1 \otimes Y)(X \otimes 1 + 1 \otimes X) \tag{197}$$
$$= XY \otimes 1 + 1 \otimes XY - YX \otimes 1 - 1 \otimes YX \tag{198}$$
$$= [X,Y] \otimes 1 + 1 \otimes [X,Y]. \tag{199}$$

Proposition 10.3. *If the ground field has characteristic zero, then* $\mathrm{Prim}(U(\mathfrak{g})) = \mathfrak{g}$.

Proof. First let $\mathfrak{g}$ be abelian. In this case $U(\mathfrak{g}) \cong S(\mathfrak{g})$ is the algebra of polynomial functions on $\mathfrak{g}^*$ and the comultiplication Δ exhibits the addition map $\mathfrak{g}^* \times \mathfrak{g}^* \to \mathfrak{g}^*$ in the sense that $(\Delta f)(X,Y) = f(X+Y)$. Hence f is primitive iff $f(X+Y) = f(X) + f(Y)$. By degree considerations this is true iff f is linear, and the linear f are precisely the elements of $\mathfrak{g}$. (In positive characteristic we might have $f(X) = X^p$.)

Consider the filtration on $U(\mathfrak{g}) \otimes U(\mathfrak{g})$ given by

$$F^i(U(\mathfrak{g}) \otimes U(\mathfrak{g})) = \sum_{p+q=i} F^p(U(\mathfrak{g})) \otimes F^q(U(\mathfrak{g})). \tag{200}$$

Then the comultiplication induces a map $\mathrm{Gr}(U(\mathfrak{g})) \to \mathrm{Gr}(U(\mathfrak{g})) \otimes \mathrm{Gr}(U(\mathfrak{g}))$ on associated gradeds. Furthermore, if $X \in U(\mathfrak{g})$ is primitive, then its leading term $\overline{X} \in \mathrm{Gr}(U(\mathfrak{g}))$ (we need to make some choices here) is also primitive. ∎

10.3 The Baker-Campbell-Hausdorff formula

Let k be a ground field of characteristic zero.

Definition Let V be a vector space. The *free Lie algebra* on V is the universal Lie algebra $L(V)$ equipped with a linear map $V \to L(V)$ from V. The free Lie algebra on a set S is the free Lie algebra on the free vector space on S.

What is $U(L(V))$? There is a canonical map $L(V) \to U(L(V))$ and a canonical map $V \to L(V)$ giving a canonical map $V \to U(L(V))$. This map exhibits $U(L(V))$ as the free algebra $T(V)$ on V by looking at the universal properties ($\text{Hom}(U(L(V)), -)$ is linear maps from V). Hence by PBW we can think of $L(V)$ as the Lie subalgebra of $T(V)$ generated by V under commutators. Moreover, we can recover $L(V)$ inside $T(V)$ as the Lie algebra of primitive elements, since $T(V) \cong U(L(V))$ is a Hopf algebra. Here the comultiplication is generated by

$$T(V) \supseteq V \ni x \mapsto x \otimes 1 + 1 \otimes x \in T(V) \otimes T(V). \tag{201}$$

The *Baker-Campbell-Hausdorff formula* is a formula of the form

$$\log e^X e^Y = X + Y + \frac{[X, Y]}{2} + \frac{[X, [X, Y]]}{12} + \frac{[Y, [Y, X]]}{12} + ... \tag{202}$$

which expresses the group multiplication on a Lie group in a suitably small neighborhood of the identity entirely in terms of the Lie bracket. We can show the existence of the BCH formula using the above, but we need to produce a completion. The ideal $m = \bigoplus_{i=1}^{\infty} T^i(V)$ is a two-sided ideal in $T(V)$, and we want to take the m-adic completion of $T(V)$. This is done by giving $T(V)$ a topology where the neighborhoods of the identity are the powers m^n and translating to get neighborhoods of any point. This topology is metrizable by a natural m-adic metric and we can consider the completion with respect to this metric. This gives a ring

$$\widehat{T(V)} = \prod_{i=0}^{\infty} T^i(V) \tag{203}$$

of noncommutative formal power series with a two-sided ideal $\hat{m}$ of formal power series with zero constant term. Now, the comultiplication $\Delta : T(V) \to T(V) \otimes T(V)$ respects gradings (if $T(V) \otimes T(V)$ is given its natural grading). In $T(V) \otimes T(V)$ we can consider the ideal $n = m \otimes 1 + 1 \otimes m$ and complete with respect to it, giving a completion $T(\widehat{V) \otimes T}(V)$. The comultiplication then extends to a continuous map $\widehat{T(V)} \to T(\widehat{V) \otimes T}(V)$.

Example If $\dim V = 1$, we get formal power series $k[[x]]$ in one variable. The comultiplication $\Delta : k[x] \to k[x, y]$ then extends to a continuous map $k[[x]] \to k[[x, y]]$.

We can now consider a completion $\widehat{L(V)}$ given by the closure of $L(V)$ in $\widehat{T(V)}$. These are formal power series in Lie monomials, and can be identified with $\text{Prim}(\widehat{T(V)})$.

We have a well-defined exponential map $\exp : \hat{m} \to 1 + \hat{m}$ defined by the usual formula and similarly a well-defined inverse map, the logarithm $\log : 1 + \hat{m} \to \hat{m}$.

Definition An element a of a bialgebra A is *grouplike* if $\Delta a = a \otimes a$.

We'll denote the set of all grouplike elements of a bialgebra A by $\text{Grp}(A)$. In $\widehat{T(V)}$, we have

$$\Delta X = X \otimes 1 + 1 \otimes X \Rightarrow \Delta e^X = e^{\Delta X} = e^X \otimes e^X \tag{204}$$

so the exponential of a primitive element is grouplike, and conversely the logarithm of a grouplike element is primitive. Moreover, the grouplike elements are closed under multiplication.

Now, let V be a 2-dimensional vector space with basis X, Y. In $\widehat{T(V)}$, e^X and e^Y are grouplike, hence so is $e^X e^Y$, hence $\log(e^X e^Y)$ is a primitive element and so must be a sum of Lie monomials.

This establishes the existence, abstractly, of the Baker-Campbell-Hausdorff formula. But how do we actually compute terms in it? The idea is to construct a projection

$$\pi : \hat{m} \to \widehat{L(V)} = \text{Prim}(\widehat{T(V)}) \tag{205}$$

as follows. First, let

$$\varphi(X_1 X_2 ... X_n) = [X_1, [X_2, [..., X_n]]] \tag{206}$$

and let $L^n(V) = L(V) \cap T^n(V)$, where $L(V)$ is regarded as a subspace of $T(V)$.

Lemma 10.4. *If $u \in L^n(V)$, then $\varphi(u) = nu$.*

Proof. We compute that $\varphi(uv) = \text{ad}_u(\varphi(v))$ (it suffices to check this on generators). Now inductively, we may assume WLOG that $u = [v, w]$ where $v \in L^p(V), w \in L^q(V)$, and $p + q = n$. Then

$$\begin{aligned} \varphi([v,w]) &= \varphi(vw - wv) && (207)\\ &= \text{ad}_v(\varphi(w)) - \text{ad}_w(\varphi(v)) && (208)\\ &= q[v,w] - p[w,v] && (209)\\ &= (p+q)[v,w]. && (210) \end{aligned}$$

□

Now we can define π by defining π on elements of $L^n(V)$ as $\frac{\varphi(u)}{n}$. This is continuous and hence extends to a map on completions $\pi : \hat{m} \to \widehat{L(V)}$.

Now we can compute terms in the Baker-Campbell-Hausdorff formula by expanding out

$$\log\left(1 + X + \frac{X^2}{2!} + ...\right)\left(1 + Y + \frac{Y^2}{2!} + ...\right) \tag{211}$$

and applying π to the result. This is an algorithmic, if tedious, solution, and gives another method for proving the existence of Lie groups with Lie algebra a given finite-dimensional Lie algebra.

11 Structure theory of semisimple Lie algebras

11.1 Weights, roots, and Cartan subalgebras

Let V be a finite-dimensional vector space over an algebraically closed field k of characteristic zero and let $X \in \text{End}(V)$. Then we have a Jordan decomposition into *weight spaces* $V = \bigoplus V_\lambda^X$ where

$$V_\lambda^X = \ker(X - \lambda)^{\dim V} = \bigcup_{m=1}^{\infty} \ker(x - \lambda)^m. \tag{212}$$

If $\mathfrak{n}$ is a nilpotent Lie subalgebra of $\text{End}(V)$ then we have a similar decomposition $V = \bigoplus V_\lambda$, but instead of ranging over eigenvalues λ now ranges over linear functionals $\lambda \in \mathfrak{n}^*$ (*weights*), and the weight spaces have the form

$$V_\lambda = \bigcap_{X \in \mathfrak{n}} \ker(X - \lambda(X))^{\dim V}. \tag{213}$$

We prove this by observing that $V_\lambda^{X_1}$ is invariant under X_2 if $[X_1, X_2] = 0$, where $X_i \in \mathfrak{n}$, and inducting.

If $V = \bigoplus V_\lambda$ and $W = \bigoplus W_\mu$, then the weight space decomposition of their tensor product has the form

$$(V \otimes W)_\nu = \bigoplus_{\lambda+\mu=\nu} V_\lambda \otimes W_\mu. \tag{214}$$

We would like to apply this idea in general. We will do this by looking at nilpotent subalgebras of a Lie algebra.

Definition A *Cartan subalgebra* of a Lie algebra $\mathfrak{g}$ is a nilpotent subalgebra $\mathfrak{h}$ equal to its *normalizer*

$$N(\mathfrak{h}) = \{X \in \mathfrak{g} | [X, \mathfrak{h}] \subseteq \mathfrak{h}\}. \tag{215}$$

If $\mathfrak{g}$ has a Cartan subalgebra $\mathfrak{h}$, then the adjoint action of $\mathfrak{h}$ on $\mathfrak{g}$ induces a decomposition

$$\mathfrak{g} = \bigoplus_{\lambda \in \mathfrak{h}^*} \mathfrak{g}_\lambda. \tag{216}$$

Because $\mathfrak{h}$ is its own centralizer we have $\mathfrak{g}_0 = \mathfrak{h}$, so the weight spaces associated to nonzero weights are the interesting ones.

Definition A *root* is a nonzero weight λ such that $\dim \mathfrak{g}_\lambda \neq 0$.

The above decomposition of $\mathfrak{g}$ is the *root decomposition.*

Example Let $\mathfrak{g} = \mathfrak{sl}_3$. The subalgebra $\mathfrak{h}$ of diagonal matrices turns out to be a Cartan subalgebra. If e_{ij} is the matrix with all entries 0 except the ij-entry, then we compute that

$$\left[\begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix}, e_{ij} \right] = (d_i - d_j) e_{ij}. \tag{217}$$

The weight space $\mathfrak{h}^*$ contains three linear functionals d_1, d_2, d_3 with relation $d_1 + d_2 + d_3 = 0$. In this space, the roots are the linear functionals $d_i - d_j, i \neq j$; in particular there are 6 of them. They can be drawn as the vertices of a regular hexagon (there is a Euclidean structure here but it is not obvious yet).

Note that $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subseteq \mathfrak{g}_{\lambda+\mu}$. Hence the Cartan subalgebra induces a nice grading of $\mathfrak{g}$.

Theorem 11.1. *Every finite-dimensional Lie algebra has a Cartan subalgebra. If G is a Lie group with Lie algebra $\mathfrak{g}$ then any two Cartan subalgebras of $\mathfrak{g}$ are conjugate by the adjoint action of G.*

We will prove this over $\mathbb{C}$ (it is false over $\mathbb{R}$). It is actually true over an algebraically closed field of characteristic zero, but we need to replace Lie groups with algebraic groups.

We first need some preliminary facts. Let $X \in \mathfrak{g}$. Then $\text{ad}_X \in \text{End}(\mathfrak{g})$ is never invertible because X is in its nullspace. Write

$$\det(t - \text{ad}_X) = t^n + c_{n-1}(X) t^{n-1} + ... + c_m(X) t^m. \tag{218}$$

Definition The *rank* of $\mathfrak{g}$ is the minimal r such that $c_r(X)$ is not identically zero. $X \in \mathfrak{g}$ is *regular* if $c_r(X) \neq 0$, where r is the rank of $\mathfrak{g}$.

Example Let $\mathfrak{g} = \mathfrak{sl}_2$. We compute that

$$\det(t - \text{ad}_H) = (t-2)(t+2)t = t^3 - 4t. \tag{219}$$

The rank cannot be zero, so it is 1. The set of regular elements of $\mathfrak{sl}_2$ is the set of X such that $\det(X) \neq 0$.

The set $\mathfrak{g}^{\text{reg}}$ of regular elements of $\mathfrak{g}$ is the complement of the variety $c_r(X) = 0$. In particular, it is open and dense in the Euclidean topology. Since the variety has codimension 2, it is also connected.

Now we prove the theorem.

Proof. Let X be a regular element of $\mathfrak{g}$ and consider the Jordan decomposition

$$\mathfrak{g} = \bigoplus \mathfrak{g}_\lambda^X \tag{220}$$

under the adjoint action of X. We claim that $\mathfrak{h} = \mathfrak{g}_0^X$ is a Cartan subalgebra. Write $\mathfrak{g} = \mathfrak{h} \oplus W$ where W is the sum of the nonzero root spaces. Since $[\mathfrak{h}, \mathfrak{g}_\lambda] \subseteq \mathfrak{g}_\lambda$, for any $Y \in \mathfrak{h}$ we can write

$$\text{ad}(Y) = \text{ad}_1(Y) + \text{ad}_2(Y) \tag{221}$$

where ad_1 is the part acting on $\mathfrak{h}$ and ad_2 is the part acting on W. Let

$$V = \{Y \in \mathfrak{h} | \text{ad}_1(Y) \text{ nilpotent}\} \tag{222}$$

$$U = \{Y \in \mathfrak{h} | \text{ad}_2(Y) \text{ invertible}\}. \tag{223}$$

Since X is regular we have $U \subseteq V$. But U is open and V is Zariski closed, so in fact $V = \mathfrak{h}$.

So $\mathfrak{h}$ is nilpotent. The normalizing condition is straightforward to check, and it follows that $\mathfrak{h}$ is a Cartan subalgebra.

We now need a lemma.

Lemma 11.2. *Let $\mathfrak{h}$ be a Cartan subalgebra, $\mathfrak{g} = \mathfrak{h} \oplus W$ be as above, ad_1, ad_2 be as above, and U as above. Then U is nonempty.*

But this is straightforward: W is a sum of weight spaces associated to finitely many weights $\lambda_1, ..., \lambda_m$, and we can take the complement of the union of their kernels, which are hyperplanes.

Now consider the map

$$\Phi : U \times W \ni (v, w) \mapsto \exp(\text{ad}(w))v \in \mathfrak{g}. \tag{224}$$

The differential $d\Phi(v, 0)$ may be regarded as an operator $\mathfrak{h} \oplus W \to \mathfrak{h} \oplus W$, and as such an operator it is the identity on $\mathfrak{h}$ and $\text{ad}_2(v)$ on W. In particular, it is surjective. Hence Φ is locally surjective.

It follows that any Cartan subalgebra in $\mathfrak{g}$ equals $\mathfrak{g}_0^X$ for some $X \in \mathfrak{g}^{\text{reg}}$. To see this, we use the fact that $(\text{Ad}(G))(U)$ contains an open set, hence contains a regular element Y, and then $\text{Ad}(g)(\mathfrak{h})$ ($\mathfrak{h}$ a Cartan subalgebra) is $\mathfrak{g}_0^Y$ for some g.

Consider the equivalence relation on $\mathfrak{g}^{\text{reg}}$ given by $X \sim Y$ if $\mathfrak{g}_0^X$ is conjugate to $\mathfrak{g}_0^Y$. We want to show that there is only one equivalence class. But by the local surjectivity of Φ it follows that each equivalence class is open, hence by the connectedness of $\mathfrak{g}^{\text{reg}}$ the conclusion follows. (This is the only place where the theorem fails over $\mathbb{R}$; in this case $\mathfrak{g}^{\text{reg}}$ is not necessarily connected.) □

In this situation it is typical to concentrate attention on the semisimple case.

11.2 The semisimple case

Let $\mathfrak{g}$ be semisimple, B be the Killing form, and

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha \tag{225}$$

be the root decomposition, where $\mathfrak{h}$ is a Cartan subalgebra and $\Delta \subset \mathfrak{h}^* \setminus \{0\}$ is the set of roots.

Proposition 11.3. *1. $B(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ unless $\alpha + \beta = 0$.*

2. The restriction of B to $\mathfrak{h}$ is nondegenerate.

3. $\mathfrak{h}$ is abelian.

4. $\mathfrak{h}$ consists only of semisimple elements.

Proof. To establish the first claim we compute that, for $H \in \mathfrak{h}, X \in \mathfrak{g}_\alpha, Y \in \mathfrak{g}_\beta$, if X, Y are eigenvectors for H we have

$$B([H, X], Y) + B(X, [H, Y]) = (\alpha(H) + \beta(H))B(X, Y) = 0. \tag{226}$$

If $\alpha + \beta \neq 0$ there exists H such that $\alpha(H) + \beta(H) \neq 0$, hence $B(X, Y) = 0$. To establish the claim for generalized eigenvectors we work by induction.

The second claim follows from the first claim, which shows that $\mathfrak{h}$ is orthogonal to $\bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha$ with respect to the Killing form.

The third claim follows from the second claim, since $B(\mathfrak{h}, [\mathfrak{h}, \mathfrak{h}]) = 0$ implies $[\mathfrak{h}, \mathfrak{h}] = 0$.

To establish the fourth claim, recall that for $X \in \mathfrak{g}$ the semisimple and nilpotent parts X_s, X_n have the property that $\text{ad}(X_s), \text{ad}(X_n)$ are polynomials in $\text{ad}(X)$. In particular, if $X \in \mathfrak{h}$ then $\text{ad}(X)$ preserves the root decomposition, hence so do polynomials in it, hence (by the self-normalizing condition) $X_s, X_n \in \mathfrak{h}$ as well. But if $X_n \in \mathfrak{h}$ then $X_n = 0$ since by Lie's theorem every $\text{ad}(X), X \in \mathfrak{h}$ can be simultaneously upper-triangularized and $\text{ad}(X_n)$ is strictly upper-triangular. $\square$

Corollary 11.4. *In a semisimple Lie algebra, regular elements are semisimple.*

The definition of regular is sometimes different in the literature: it sometimes means that the dimension of the centralizer is the minimum possible. These elements are also called principal.

Corollary 11.5. *$\mathfrak{g}_\alpha$ is the usual (not generalized) weight space*

$$\{X \in \mathfrak{g} | [H, X] = \alpha(H) X \forall H \in \mathfrak{h}\}. \tag{227}$$

By the nondegeneracy of B, if $\alpha \in \Delta$ is a root then so is $-\alpha$. Let $\mathfrak{h}_\alpha = [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subseteq \mathfrak{h}$. We compute that if $H \in \mathfrak{h}, X \in \mathfrak{g}_\alpha, y \in \mathfrak{g}_{-\alpha}$, then

$$B(H, [X, Y]) = B([H, X], Y) = \alpha(h)B(X, Y). \tag{228}$$

In particular, $B(\ker(\alpha), \mathfrak{h}_\alpha) = 0$, and by nondegeneracy of B it follows that $\dim \mathfrak{h}_\alpha = 1$.

Suppose by contradiction that α is zero when restricted to $\mathfrak{h}_\alpha$. Then we can choose $H \neq 0 \in \mathfrak{h}_\alpha, X \in \mathfrak{g}_\alpha, Y \in \mathfrak{g}_{-\alpha}$ such that $[X, Y] = H, [H, X] = 0, [H, Y] = 0$. These generate the Heisenberg algebra, which is nilpotent, so $\text{ad}(H)$ acts nilpotently. But this contradicts semisimplicity.

Hence it follows that α is nonzero when restricted to $\mathfrak{h}_\alpha$. Let $H_\alpha \in h_\alpha$ be such that $\alpha(H_\alpha) = 2$. Then we can find $X_\alpha \in \mathfrak{g}_\alpha, Y_\alpha \in \mathfrak{g}_{-\alpha}$ such that

$$H_\alpha = [X_\alpha, Y_\alpha] \tag{229}$$

$$[H_\alpha, X_\alpha] = 2X_\alpha \tag{230}$$

$$[H_\alpha, Y_\alpha] = -2Y_\alpha. \tag{231}$$

This is a copy $(\mathfrak{sl}_2)_\alpha$ of $\mathfrak{sl}_2$ in $\mathfrak{g}$ which we will call an $\mathfrak{sl}_2$-triple. Because $\mathfrak{sl}_2$ is semisimple, every finite-dimensional representation of it is completely reducible. We also know that the irreducible representations $V(m)$ of $\mathfrak{sl}_2$ are parameterized by the non-negative integers, and on $V(m)$, H acts semisimply with integer eigenvalues. In fact we know that $V(m)$ admits a basis on which H acts with eigenvalues $-m, -(m-2), ..., m-2, m$, and X and Y shift up and down weight spaces by 2 respectively. Moreover, if $Hv = \lambda v$ and $Yv = 0$ then $\lambda \leq 0$, and if $Hv = \lambda v$ and $Xv = 0$ then $\lambda \geq 0$. Applied to $(\mathfrak{sl}_2)_\alpha$, it follows in particular that $\beta(H_\alpha) \in \mathbb{Z}$ for all roots β.

Proposition 11.6. $\dim \mathfrak{g}_\alpha = 1$.

Proof. Suppose otherwise. Consider $H_\alpha, X_\alpha, Y_\alpha$ as above. Then there exists $X \in \mathfrak{g}_\alpha$ such that $[Y_\alpha, X] = 0$, so $\text{ad}(H_\alpha)(X) = 2X$, but this contradicts our observation above that if $Yv = 0$ then $\lambda \leq 0$. □

A homework problem implies that the H_α span $\mathfrak{h}$. Let $\mathfrak{h}_\mathbb{R}$ be the real span of the H_α. Then the Killing form B is positive definite on $\mathfrak{h}_\mathbb{R}$, since for all $H \in \mathfrak{h}_\mathbb{R}$ we have $B(H, H) = \text{tr}(\text{ad}(H)^2) > 0$ since we know H acts with real (in fact integer) eigenvalues on $\mathfrak{g}$.

Let α, β be roots. Suppose $\beta(H_\alpha) = m > 0$. Then we know from properties of $\mathfrak{sl}_2$ actions that $(\text{ad}(Y_\alpha))^m X_\beta \in \mathfrak{g}_{\beta - m\alpha}$ is nonzero, so $\beta - m\alpha$ is a root. Similarly, if $\beta(H_\alpha) = m < 0$, then $(\text{ad}(X_\alpha))^m X_\beta \in \mathfrak{g}_{\beta + m\alpha}$ is nonzero, so $\beta + m\alpha$ is a root. We find that

$$\alpha, \beta \in \Delta \Rightarrow \beta - \beta(H_\alpha)\alpha \in \Delta. \tag{232}$$

We now define certain linear transformations r_α on the space spanned by the roots as follows. Define $r_\alpha(\beta) = \beta - \beta(H_\alpha)\alpha$ if $\beta \neq \alpha$, and define $r_\alpha(\alpha) = -\alpha$. This is reflection across the hyperplane perpendicular to α with respect to the following inner product.

First, notice that since B is nondegenerate we can identify $\mathfrak{h}$ with $\mathfrak{h}^*$ using it, and moreover B induces a nondegenerate form $(-,-)$ on $\mathfrak{h}^*$. The image of H_α under this identification must be proportional to α, and in fact it is precisely $\frac{2\alpha}{(\alpha,\alpha)} = \alpha^\vee$ (the *coroot* corresponding to the root α). We can now rewrite r_α as

$$r_\alpha(\xi) = \xi - \frac{2(\xi,\alpha)}{(\alpha,\alpha)}\alpha \tag{233}$$

which is precisely a reflection. In particular, $r_\alpha^2 = \text{id}$. (Here we are thinking of the real span of the H_α and the $\alpha^\vee$.) These reflections generate a finite group W acting on Δ called the *Weyl group* of $\mathfrak{g}$.

The situation now is that we have identified various interesting data - roots and a Weyl group - associated to a semisimple Lie algebra. We can now attempt to classify all such data in order to classify semisimple Lie algebras.

Definition Let V be a finite-dimensional real inner product space. A *root system* is a finite subset $\Delta \subset V \setminus \{0\}$ such that

1. Δ spans V,
2. If $\alpha, \beta \in \Delta$, then $\frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{Z}$,
3. If $\alpha, \beta \in \Delta$, then $r_\alpha(\beta) \in \Delta$.

The *rank* of Δ is $\dim V$.

Example Let $\mathfrak{g} = \mathfrak{sl}_3$. All roots α satisfy $(\alpha,\alpha) = 2$. We can find roots α, β with $(\alpha,\beta) = 1$ (they are 60° apart). The Weyl group turns out to be S_3.

Example Let $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$. A Cartan subalgebra $\mathfrak{h}$ is given by the diagonal matrices. These have n diagonal entries. Let $\tilde{\mathfrak{h}}$ be the diagonal matrices in $\mathfrak{gl}_n(\mathbb{C})$; then we can think of the n diagonal entries as n elements $\varepsilon_1, ..., \varepsilon_n \in \tilde{\mathfrak{h}}^*$, which define elements $\varepsilon_i - \varepsilon_j \in \mathfrak{h}^*$ for $i \neq j$ Each of these are roots with root space spanned by the elementary matrix e_{ij}. The ε_i are an orthogonal basis with respect to the inner product on $\mathfrak{h}^*$.

We can now compute the Weyl group. The simple reflection $r_{\varepsilon_i - \varepsilon_j}$ acting on $\tilde{\mathfrak{h}}^*$, which we identify with $\mathbb{C}^n$ via the basis ε_i, permutes the entries i and j. The same is true of the action on $\mathfrak{h}^*$, which is a quotient of the above. Hence the Weyl group in this case is S_n.

The root system associated to $\mathfrak{sl}_n(\mathbb{C})$ is the one of type A_{n-1}; in general the index is the rank.

Example The root decomposition of $\mathfrak{g} = \mathfrak{so}(n)$ depends on the parity of n. First we need to choose a Cartan subalgebra. It is convenient to use the diagonal subalgebra in some basis, but the obvious basis is bad. When $n = 2k$ we will instead do the following. Let B be the block matrix

$$B = \begin{bmatrix} 0_n & 1_n \\ 1_n & 0_n \end{bmatrix}. \tag{234}$$

Then $\mathfrak{so}(n)$ is isomorphic to the Lie algebra of real $n \times n$ matrices X satisfying $X^T B + BX = 0$. These are precisely the block matrices of the form

$$X = \begin{bmatrix} A & C \\ D & -A^T \end{bmatrix} \tag{235}$$

where C, D are skew-symmetric. In particular, we get a diagonal subalgebra of dimension k which is a Cartan subalgebra.

The root system here is the one of type D_k. First, we have k weights $\varepsilon_1, ..., \varepsilon_k \in \mathfrak{h}^*$ coming from looking at diagonal entries giving us roots $\varepsilon_i - \varepsilon_j, i \neq j$ with root spaces the (block-diagonal) elementary matrices as before. However, there are also root spaces coming from the C and D parts, giving us roots $\varepsilon_i + \varepsilon_j, i \neq j$ and $-\varepsilon_i - \varepsilon_j, i \neq j$. Some examples:

1. When $k = 1$ we have a degenerate case since $\mathfrak{so}(2)$ is abelian; there are no roots.

2. When $k = 2$ we have 4 roots $\pm\varepsilon_1 \pm \varepsilon_2$. Drawing these gives the four vertices of a square; in particular, two of the roots are orthogonal to the other two, so the corresponding root spaces commute; in other words, there is a decomposition $D_2 \cong A_1 \times A_1$ of root systems. This reflects the isomorphism $\mathfrak{so}(4) \cong \mathfrak{so}(3) \times \mathfrak{so}(3)$.

3. When $k = 3$ we have 12 roots $\pm\varepsilon_i \pm \varepsilon_j$. This is the same number of roots as in the root system A_3 associated to $\mathfrak{sl}_4$, and in fact the root systems are isomorphic. This reflects the isomorphism $\mathfrak{so}(6) \otimes \mathbb{C} \cong \mathfrak{sl}_4(\mathbb{C})$ which we can see as follows. Let V be the defining 4-dimensional representation of $\text{SL}_4(\mathbb{C})$. Then $\Lambda^2(V)$ is a 6-dimensional representation equipped with a nondegenerate bilinear form given by the wedge product $\Lambda^2(V) \otimes \Lambda^2(V) \to \Lambda^4(V)$. The latter is the trivial representation of $\text{SL}_4(\mathbb{C})$ because it is the determinant representation, so SL_4 preserves a nondegenerate bilinear form on $\Lambda^2(V)$, giving a homomorphism $\text{SL}_4 \to \text{SO}(6, \mathbb{C})$. The kernel of this map is ± 1 and the two Lie algebras have the same dimension, so this is a covering map and induces an isomorphism of Lie algebras.

When $n = 2k + 1$ is odd the picture is less nice. Now we will take B to be the block matrix

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0_k & 1_k \\ 0 & 1_k & 0_k \end{bmatrix}. \tag{236}$$

Then the elements of $\mathfrak{so}(2k+1)$ have the form

$$X = \begin{bmatrix} 0 & x_1, ..., x_k & y_1, ..., y_k \\ -y_1, ..., y_k & A & C \\ -x_1, ..., x_k & D & -A^T \end{bmatrix}. \tag{237}$$

There is an inclusion $\mathfrak{so}(2k) \to \mathfrak{so}(2k+1)$ which induces an inclusion of root systems; we have the same root systems as above but we also have the roots $\pm\varepsilon_1, ..., \pm\varepsilon_k$.

The root system here is the one of type B_k. We can compute the Weyl group as follows. The simple reflections $r_{\varepsilon_i - \varepsilon_j}$ again act by transposition. The simple reflections $r_{\varepsilon_i+\varepsilon_j}$ act by transposition and then a -1 on the transposed entries. The simple reflections r_{ε_i} act by -1 on the i^{th} entry. The corresponding group $W(B_k)$ is the hyperoctahedral group, or equivalently the group of symmetries of a k-cube. It may be presented as permutation matrices with entries ± 1, or equivalently as a wreath product $\mathbb{Z}_2 \wr S_k$, which is a semidirect product $(\mathbb{Z}_2)^k \rtimes S_k$.

This lets us compute the group $W(D_k)$: it is the subgroup of index 2 of the above group where there are an even number of sign changes. Abstractly it is a semidirect product $(\mathbb{Z}_2)^{k-1} \rtimes S_k$.

In particular, the isomorphism $D_3 \cong A_3$ of root systems implies that $(\mathbb{Z}_2)^2 \rtimes S_3$ is isomorphic to S_4.

Example Let $\mathfrak{g} = \mathfrak{sp}(2n)$ be the Lie algebra of $2n \times 2n$ matrices preserving a skew-symmetric bilinear form. With respect to the block matrix

$$J = \begin{bmatrix} 0_n & 1_n \\ -1_n & 0_n \end{bmatrix} \tag{238}$$

we can write $\mathfrak{g}$ as the collection of matrices X satisfying $X^T J + JX = 0$. Here we again can write all such matrices in the form

$$X = \begin{bmatrix} A & C \\ D & -A^T \end{bmatrix} \tag{239}$$

but now C, D are symmetric. Again we have a diagonal Cartan subalgebra with associated roots $\pm\varepsilon_i \pm \varepsilon_j$, but the diagonal entries of C and D give us new roots $\pm 2\varepsilon_i$.

The root system here is the one of type C_n. It can be obtained from the one of type B_n by multiplying the short roots $\pm\varepsilon_i$ by 2, so in particular $W(C_n) \cong W(D_n)$. There is an isomorphism of root systems $B_2 \cong C_2$ reflecting an isomorphism $\mathfrak{so}(5) \otimes \mathbb{C} \cong \mathfrak{sp}(4) \otimes \mathbb{C}$.

11.3 Root systems

Definition A root system Δ is called *irreducible* if it cannot be written as a disjoint union $\Delta_1 \sqcup \Delta_2$ where $(\Delta_1, \Delta_2) = 0$.

Every root system is uniquely a disjoint union of irreducible root systems.

Definition A root system is *reduced* if $\alpha, p\alpha \in \Delta$ implies $p = \pm 1$.

We proved on the homework that all root systems arising from Lie algebras have this property.

Example The root system $BC_n = B_n \cup C_n$ is non-reduced.

If Δ is a root system and $\alpha \in \Delta$ is a root, define $\alpha^\vee = \frac{2\alpha}{(\alpha,\alpha)}$, so that $(\alpha, \beta^\vee) = \frac{2(\alpha,\beta)}{(\beta,\beta)} \in \mathbb{Z}$. Then

$$\Delta^\vee = \{\alpha^\vee | \alpha \in \Delta\} \tag{240}$$

is also a root system. A_n and D_n are self-dual, while $B_n^\vee = C_n$. This is related to Langlands duality.

Let $\alpha, \beta \in \Delta$ be two roots in a root system. Define

$$\begin{aligned} m(\alpha,\beta) &= (\alpha,\beta^\vee)(\beta,\alpha^\vee) && (241)\\ &= \frac{4(\alpha,\beta)^2}{(\alpha,\alpha)(\beta,\beta)} && (242)\\ &= 4\cos^2\theta && (243) \end{aligned}$$

where θ is the angle between α and β. We know that this quantity must be an integer. It follows that we can only have $m(\alpha, \beta) = 0, 1, 2, 3, 4$. This gives $\theta = \frac{\pi}{2}, \frac{\pi}{3}, \frac{2\pi}{3}, \frac{\pi}{4}, \frac{3\pi}{4}, \frac{\pi}{6}, \frac{5\pi}{6}, 0, \pi$ (but in the last two cases α, β are proportional and this is uninteresting).

If $m(\alpha, \beta) = 1, 2, 3$ then of the two numbers $(\alpha, \beta^\vee)$ and $(\beta, \alpha^\vee)$ one of them must be ± 1 and the other must be $\pm m(\alpha, \beta)$.

Corollary 11.7. *If* $(\alpha, \beta) > 0$*, then* $\alpha - \beta \in \Delta$*.*

Proof. In this case $(\alpha, \beta^\vee) = 1$, so $r_\beta(\alpha) = \alpha - \beta \in \Delta$. □

We already have enough information to classify rank 2 root systems. Here there are two roots α, β which are not a scalar multiple of each other. If $m(\alpha, \beta) = 0$ then we get $A_1 \sqcup A_1$. If $m(\alpha, \beta) = 1$ then we get A_2. If $m(\alpha, \beta) = 2$ then we get $B_2 = C_2$. The most interesting case is when $m(\alpha, \beta) = 3$, in whih case we get an exceptional root system G_2 associated to an exceptional Lie algebra.

If S is a subset of the roots of Δ and we intersect Δ with the span of S, the result is another root system.

Definition A subset $S = \{\alpha_1, ..., \alpha_n\}$ of Δ is a *base* if

1. S is a basis of the ambient vector space V,
2. If $\beta \in \Delta$ then $\beta = \sum n_i \alpha_i$ where either all $n_i \in \mathbb{Z}_{\geq 0}$ or all $n_i \in \mathbb{Z}_{\leq 0}$.

Bases can be constructed as follows.

Definition Choose $t \in V$ such that $(t, \alpha) \neq 0$ for all $\alpha \in \Delta$. Let

$$\Delta^+ = \{\alpha \in \Delta | (t, \alpha) > 0\} \tag{244}$$

be the *positive* roots and let

$$\Delta^- = \{\alpha \in \Delta | (t, \alpha) < 0\} \tag{245}$$

be the *negative* roots. A positive root $\alpha \in \Delta^+$ is *simple* if α cannot be written in the form $\beta + \gamma$ where $\beta, \gamma \in \Delta^+$.

Proposition 11.8. *The set S of simple roots is a base.*

Proof. The second property of a base follows more or less by construction. The first property is more difficult. We want to show that if $\alpha, \beta \in S$ then $(\alpha, \beta) \leq 0$. Suppose otherwise; then $\alpha - \beta \in \Delta$ by the above. WLOG $\alpha - \beta \in \Delta^+$, but then $\alpha = \beta + (\alpha - \beta)$ is not simple.

Any linear dependence among the elements of S can be written in the form $\sum a_i \alpha_i = \sum b_j \alpha_j$ where $a_i, b_j \geq 0$. Taking inner products we get

$$(\sum a_i \alpha_i, \sum a_i \alpha_i) = (\sum b_j \alpha_j, \sum a_i \alpha_i) \leq 0 \tag{246}$$

hence all the a_i are equal to zero, and similarly all of the b_j are equal to zero. □

Example In A_{n-1}, a choice of $t = t_1 \varepsilon_1 + ... + t_n \varepsilon_n$ satisfies the above condition iff $t_i \neq t_j$ for all $i \neq j$. Let's take $t_1 > t_2 > ... > t_n$, so

$$\Delta^+ = \{\varepsilon_i - \varepsilon_j, i < j\} \tag{247}$$

and the simple roots are $\varepsilon_i - \varepsilon_{i+1}, i = 1, 2, ...n - 1$. The reflections associated to simple roots are called *simple reflections*, and here these are the simple transpositions swapping i with $i + 1$. In this case we see that W is generated by simple reflections; this will be true in general. It is also true in general that W acts freely and transitively on the set of bases.

From now on we assume that Δ is reduced.

Lemma 11.9. *Fix a choice of Δ^+. Let S be the set of simple roots and let $\alpha \in S$. Then*

$$r_\alpha(\Delta^+) = (\Delta^+ \setminus \{\alpha\}) \cup \{-\alpha\}. \tag{248}$$

Proof. We induct on the number of simple roots needed to add up to a positive root. Let $\beta \in S$. Then $r_\alpha(\beta) = \beta - (\beta, \alpha^\vee)\alpha = \beta + m\alpha$ where $m \geq 0$. Hence this is a positive root.

Let β be a non-simple positive root. Then $\beta = \beta_1 + \beta_2$ where β_1, β_2 are positive. If neither is equal to α then we are done by induction. If $\beta = \beta_1 + \alpha$ then $r_\alpha(\beta) = r_\alpha(\beta_1) - \alpha$.

The following general fact is true: if $\gamma \in \Delta^+$ and $\alpha \in S$ with $\gamma \neq \alpha$, then $\gamma - \alpha \in \Delta$ implies $\gamma - \alpha \in \Delta^+$. This is true since otherwise $\alpha - \gamma \in \Delta^+$ contradicts α simple.

It follows that $r_\alpha(\beta_1) - \alpha \in \Delta^+$ as desired. □

Lemma 11.10. *Let t_1, t_2 be associated to two choices Δ_1^+, Δ_2^+ of positive roots. Then there exists some $w \in W$ such that $\Delta_2^+ = w(\Delta_1^+)$.*

Proof. Let $d = |\Delta_1^+ \cap \Delta_2^-|$. We induct on d. If $d = 0$ then $\Delta_1^+ = \Delta_2^+$. If $d > 0$ then we can find $\alpha \in S_1$ such that $\alpha \notin \Delta_2^+$. Then $|r_\alpha(\Delta_1^+) \cap \Delta_2^-| = d - 1$, and by the inductive hypothesis we are done. □

Moreover, the above construction of bases exhausts all possible bases. This is because every base must be contained in some half-space.

Corollary 11.11. *W acts transitively on the set of bases.*

Note that any root is contained in some base.

Corollary 11.12. *W is generated by simple reflections.*

(Exercise.)

We can now attempt to classify root systems using simple roots. Recall that we know there are only a few possibilities for the angle between any two such simple roots α_i, α_j, namely $90°, 120°, 135°, 150°$. We can encode this information as follows.

Definition The *Coxeter graph* of a root system has vertices a set S of simple roots such that the number of edges between two vertices α_i, α_j is $0, 1, 2, 3$ respectively in the above four cases. Equivalently, the number of edges between α_i, α_j is

$$m(\alpha_i, \alpha_j) = (\alpha_i, \alpha_j^\vee)(\alpha_j, \alpha_i^\vee). \tag{249}$$

The *Dynkin diagram* is the Coxeter graph together with a direction on edges if α_i and α_j have different lengths, from the longer root to the shortest.

Example Consider A_2. After choosing a hyperplane we get two simple roots an angle of $120°$ from each other, so the Coxeter graph is $\bullet \text{---} \bullet$.

Example Consider B_2. After choosing a hyperplane we get two simple roots an angle of $135°$ from each other, so the Coxeter graph is $\bullet = \bullet$. The Dynkin diagram is $\bullet \Longrightarrow \bullet$ because the two simple roots have different lengths.

Definition A root system is *simply laced* if its Dynkin diagram has no multiple edges or directed edges and is connected.

Definition The *Cartan matrix* $C(\Delta)$ of a root system is the matrix with entries

$$a_{ij} = (\alpha_j, \alpha_i^\vee) = \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)}. \tag{250}$$

Note that $a_{ii} = 2$ and $a_{ij} \le 0$ if $i \neq j$. If $\Delta^\vee$ is the dual root system then $C(\Delta^\vee) = C(\Delta)^T$. Furthermore, the data of the Cartan matrix is equivalent to the data of the Dynkin diagram. In one direction this is clear since the number of edges between α_i and α_j is $a_{ij}a_{ji}$.

On your homework you'll show that the Coxeter graph of Δ is connected iff Δ is irreducible. We'll try to classify root systems by classifying irreducible root systems, or equivalently connected Dynkin diagrams. A sequence of observations (all Coxeter graphs and Dynkin diagrams below are connected):

1. A full subgraph (a subgraph on some set of vertices with every edge between those vertices) of a Dynkin diagram is again a Dynkin diagram.

2. A Coxeter graph cannot contain cycles. To see this, let u_i be the unit vector in the direction α_i. If $u_1, ..., u_m$ is a cycle in a Coxeter graph Γ, then

$$(u_1 + ... + u_m, u_1 + ... + u_m) = m + 2\sum_{i<j}(u_i, u_j). \tag{251}$$

   By cyclicity, at least m of the terms (u_i, u_j) are nonzero. They are all nonpositive, and the nonzero terms are less than or equal to $-\frac{1}{2}$, so this scalar product is nonpositive; contradiction.

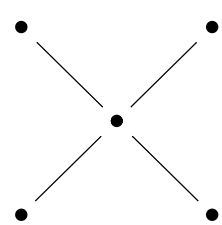
3. If Γ is a Coxeter graph obtained by joining a vertex α_1 of a graph Γ_1 to a vertex α_p of a graph Γ_2 by a bridge $\alpha_1, ..., \alpha_p$, then the graph obtained by wedging Γ_1 and Γ_2 along the vertices α_1, α_p is also a Coxeter graph. To see this, we replace $\alpha_1, ..., \alpha_p$ with $\beta = \alpha_1 + ... + \alpha_p$, which has the same inner product with every vertex in Γ_1 as α_1 and has the same inner product with every vertex in Γ_2 as α_p.

4. $\bullet \equiv \bullet - \bullet$ is not a Coxeter graph. If u_1, u_2, u_3 are the corresponding unit vectors associated to the roots then the angles between them are $90°, 120°, 150°$, which add up to $360°$, so the vectors must lie in a plane and this contradicts linear independence. Similarly any graph containing the above (with more multiple edges) or the graph $\bullet = \bullet = \bullet$ is not a Coxeter graph.

5. If Γ has a triple edge, then Γ is G_2 with Coxeter graph $\bullet \equiv \bullet$. This follows from the above.

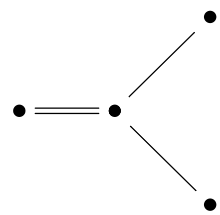
6. If Γ has a double edge, it has at most one. This again follows from the above (we consider an appropriate full subgraph containing two double edges and a bridge between them, then contract the bridge).

7. The product of the Cartan matrix C by the matrix with diagonal entries $\frac{(\alpha_i, \alpha_i)}{2}$ is the matrix of the scalar product in S. In particular, $\det C > 0$. In particular, the graphs



(252)

   and



(253)

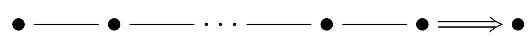
are not Coxeter graphs (the corresponding determinants vanish).

The following lemma is useful for computing determinants of Cartan matrices. Let Γ be a Dynkin diagram. Let α be a vertex of Γ which is connected to exactly one other vertex β. Let $\Gamma \setminus \alpha$ denote the Dynkin diagram obtained by removing α (and any corresponding edges), and let $\Gamma \setminus \{\alpha, \beta\}$ denote the Dynkin diagram obtained by removing α and β (and any corresponding edges). For Γ a Dynkin diagram, let $C(\Gamma)$ denote its Cartan matrix. Then

$$\det C(\Gamma) = 2 \det C(\Gamma \setminus \alpha) - m(\alpha, \beta) \det C(\Gamma \setminus \{\alpha, \beta\}). \tag{254}$$

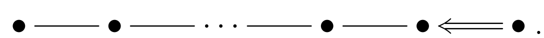
8. By contracting bridges, we see that a Coxeter graph which is not G_2 must consist of a path of length p, a double edge, and a path of length q, or else must consist of a path of length p, a path of length q, and a path of length r connected at a vertex. Now we just need to compute the determinants of the Cartan matrices.

9. In the double edge case, we get B_n or C_n if one of the paths has length zero, depending on which direction the double edge is oriented. Here $\det C(B_n) = \det C(C_n) = 2$. The Dynkin diagrams have the form



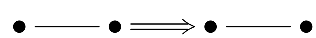
(255)

and



(256)

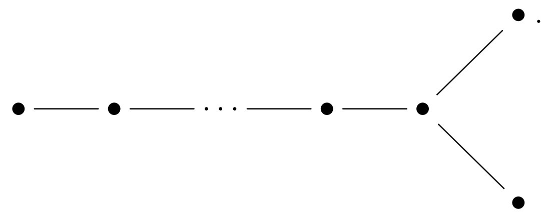
There is one additional case, the Dynkin diagram



(257)

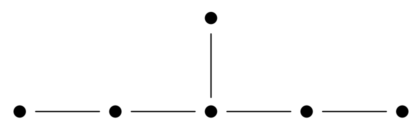
which is associated to a root system called F_4. This is the subject of a homework problem.

10. In the triple edge case, we compute that $\det C = 2pqr - qr(p-1) - pq(r-1) - pr(q-1)$. This is positive iff $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} > 1$. There is an infinite family $(2, 2, n)$ of solutions to this inequality which give the root systems D_n with Coxeter graphs of the form

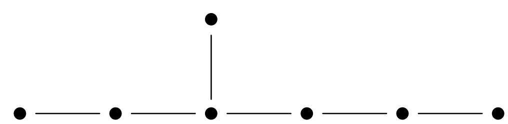


(258)

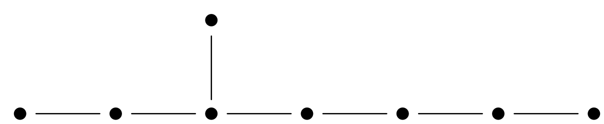
There are three exceptional solutions $(2,3,3),(2,3,4),(2,3,5)$ giving root systems E_6, E_7, E_8 whose Dynkin diagrams have the following forms:



(259)



(260)



(261)

To get the corresponding root systems it suffices to construct E_8 since E_6, E_7 are full subgraphs of E_8. The root lattice is the lattice in $\mathbb{R}^8$ obtained by considering the lattice L spanned by $\varepsilon_1, ..., \varepsilon_8, \frac{\varepsilon_1+...+\varepsilon_8}{2}$ and then considering the sublattice L' of vectors $\sum a_i \varepsilon_i \in L$ such that $\sum a_i \in 2\mathbb{Z}$. The root system is $\Delta = \{\alpha \in L' | (\alpha, \alpha) = 2\}$.

The determinants of Cartan matrices contain important information. For example, $\det C(A_n) = n+1$ is the order of the center of the simply connected complex Lie group with Lie algebra $\mathfrak{sl}_{n+1}(\mathbb{C})$. The corresponding Dynkin diagrams are paths

(262)

There are two additional questions: which root systems actually arise from semisimple Lie algebras, and does the root system actually determine the Lie algebra?

Theorem 11.13. *Let Δ be a reduced irreducible root system. Then there exists a simple Lie algebra $\mathfrak{g}$, unique up to isomorphism, whose root system is Δ.*

Proof. The idea is to assume that $\mathfrak{g}$ is a simple Lie algebra with root system Δ and to see what it must look like. Let $S = \{\alpha_1, ..., \alpha_n\}$ be a base. For a simple root α_i let (H_i, X_i, Y_i) be the corresponding $\mathfrak{sl}_2$-triple. We have the following relations among these, where $a_{ij} = (\alpha_i^\vee, \alpha_j)$:

- $[H_i, X_j] = a_{ij} X_j$ (by definition)
- $[H_i, Y_j] = -a_{ij} Y_j$ (by definition)
- $[H_i, H_j] = 0$ (since the H_i are in $\mathfrak{h}$)
- $[X_i, Y_j] = \delta_{ij} H_i$ (since $\alpha_i - \alpha_j$ is not a root).

These are the Chevalley relations, and they are not enough. To get more, we will use more facts about $\mathfrak{sl}_2$-representations. Recall that if V is a finite-dimensional such representation and $v \in V$ with $Hv = \lambda v$, then $Xv = 0$ implies $\lambda \in \mathbb{Z}_{\geq 0}$ and $Y^{(1+\lambda)} v = 0$, and $Yv = 0$ implies $\lambda \in \mathbb{Z}_{\leq 0}$ and $X^{(1-\lambda)} v = 0$. Applied to the $\mathfrak{sl}_2$-triple (H_i, X_i, Y_i) with $v = Y_j, j \neq i$, we get additional relations

- $(\text{ad}_{Y_i})^{1-a_{ij}}(Y_j) = 0$
- $(\text{ad}_{X_i})^{1-a_{ij}}(X_j) = 0$.

These are the Serre relations. In particular, if $a_{ij} = 0$, then $[Y_i, Y_j] = [X_i, X_j] = 0$.

The Chevalley and Serre relations describe an abstract Lie algebra. Our goal is to show that it is finite-dimensional and simple with root system Δ.

First, we will only impose the Chevalley relations, getting a Lie algebra $\tilde{\mathfrak{g}}$ which is the quotient of the free Lie algebra on the (H_i, X_i, Y_i) by the ideal generated by the Chevalley relations.

Lemma 11.14. *1. $\tilde{\mathfrak{h}} = span(H_1, ..., H_n)$ is an abelian subalgebra of $\tilde{\mathfrak{g}}$ which acts semisimply on $\tilde{\mathfrak{g}}$.*

2. Let $\tilde{\mathfrak{n}}^+$ resp. $\tilde{\mathfrak{n}}^-$ be the subalgebras in $\tilde{\mathfrak{g}}$ generated by $X_1, ..., X_n$ resp. $Y_1, ..., Y_n$. Then $\tilde{\mathfrak{g}}$ is the vector space direct sum

$$\tilde{\mathfrak{n}}^+ \oplus \tilde{\mathfrak{h}} \oplus \tilde{\mathfrak{n}}^-. \tag{263}$$

Moreover, $\tilde{\mathfrak{n}}^+$ and $\tilde{\mathfrak{n}}^-$ are free Lie algebras on the X_i resp. the Y_i.

Proof. 1. The first statement is clear. For the second statement, let $H \in \tilde{\mathfrak{h}}$ and suppose X, Y are eigenvectors of ad_H with eigenvalues α, β. Then $[X, Y]$ is an eigenvector with eigenvalue $\alpha + \beta$. Since $\tilde{\mathfrak{g}}$ is generated under commutator by eigenvectors of ad_H, the conclusion follows.

2. First observe that $\tilde{\mathfrak{g}}$ admits an involution σ given by $\sigma(H_i) = -H_i, \sigma(X_i) = Y_i, \sigma(Y_i) = X_i$. Hence to prove something about $\tilde{\mathfrak{n}}^{\pm}$ it suffices to prove it for $\tilde{\mathfrak{n}}^+$. Observe that by induction we have

$$[X_i, \tilde{\mathfrak{n}}^+] \subset \tilde{\mathfrak{n}}^+, [\tilde{\mathfrak{h}}, \tilde{\mathfrak{n}}^+] \subset \tilde{\mathfrak{n}}^+, [Y_i, \tilde{\mathfrak{n}}^+] \subset \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}^+ \tag{264}$$

by applying the Chevalley relations. The same conclusion for $\tilde{\mathfrak{n}}^-$ shows that $\tilde{\mathfrak{g}}$ is the not-necessarily-direct sum of the three terms we want.

Note that $\tilde{\mathfrak{g}}$ has a weight decomposition

$$\tilde{\mathfrak{g}} = \tilde{\mathfrak{h}} \oplus \bigoplus \tilde{\mathfrak{g}}_\lambda \tag{265}$$

with respect to $\tilde{\mathfrak{h}}$, giving us a set $\tilde{\Delta}$ of roots of $\tilde{\mathfrak{g}}$. Using the weight decomposition, we now construct a module M over $\tilde{\mathfrak{g}}$ which, as a vector space, is the tensor product

$$M = T(\text{span}(Y_1, ..., Y_n)) \otimes S(\text{span}(H_1, ..., H_n)) \tag{266}$$

of a tensor algebra and a symmetric algebra. Each Y_i acts by left multiplication. Each H_i acts on terms $Y \otimes v$, where Y is a monomial, via

$$H_i(Y \otimes v) = \lambda(H_i)Y \otimes v + Y \otimes H_i v \tag{267}$$

where λ is the weight of Y (1 has weight 0). Each X_i acts as follows: inductively define $X_i(1 \otimes v) = 0$ and

$$X_i(Y_j Y \otimes v) = \delta_{ij} H_i(Y \otimes v) + Y_j(X_i(Y \otimes v)). \tag{268}$$

We will not check it in class, but it is true, that this is a $\tilde{\mathfrak{g}}$-module. Using this action we can now check that $\tilde{\mathfrak{n}}^- \cap (\tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}^+) = 0$ and that $\tilde{\mathfrak{h}} \cap \tilde{\mathfrak{n}}^+$. Moreover, $U(\tilde{\mathfrak{n}}^-)(1 \otimes 1)$ is the tensor algebra on $Y_1, ..., Y_n$, and hence $\tilde{\mathfrak{n}}^-$ is free (hence $\tilde{\mathfrak{n}}^+$ is free).

□

With respect to the root decomposition, we have

$$\tilde{\mathfrak{n}}^\pm = \bigoplus_{\lambda \in \tilde{\Delta}^\pm} \mathfrak{g}_\lambda \tag{269}$$

where $\tilde{\Delta}^\pm$ is a choice of positive resp. negative roots for $\tilde{\Delta}$.

Lemma 11.15. *Let $\mathfrak{j} \subseteq \tilde{\mathfrak{g}}$ be a proper ideal. Then $\mathfrak{j} \cap \tilde{\mathfrak{h}} = 0$ and*

$$\mathfrak{j} = (\mathfrak{j} \cap \tilde{\mathfrak{n}}^+) \oplus (\mathfrak{j} \cap \tilde{\mathfrak{n}}^-). \tag{270}$$

Proof. Suppose by contradiction that $H \in \mathfrak{j} \cap \tilde{\mathfrak{h}}$ is nonzero. Then there exists α_i such that $\alpha_i(H) \neq 0$. Hence $X_i, Y_i \in \mathfrak{j}$, so $H_i \in \mathfrak{j}$, and there exists α_j such that $\alpha_j(H_i) \neq 0$. By repeating this argument and using the connectedness of the Coxeter graph we see that $\mathfrak{j} = \tilde{\mathfrak{g}}$.

Now by semisimplicity $\mathfrak{j}$ has a root decomposition

$$\mathfrak{j} = \bigoplus_{\lambda \in \tilde{\Delta}} (\mathfrak{j} \cap \mathfrak{g}_\lambda) \tag{271}$$

and we can split it into positive and negative parts. □

Corollary 11.16. $\tilde{\mathfrak{g}}$ *has a unique proper maximal ideal* $\mathfrak{j}$. *Hence the quotient* $\mathfrak{g} = \tilde{\mathfrak{g}}/\mathfrak{j}$ *is the unique simple Lie algebra satisfying the Chevalley relations associated to the root system* Δ.

It remains to show that the root system of $\mathfrak{g}$ is the root system Δ we started with and that $\mathfrak{g}$ is finite-dimensional.

First we claim that $\mathfrak{g}$ satisfies the Serre relations. Equivalently, if $T_{ij} = (\text{ad}_{Y_i})^{1-a_{ij}} Y_j \in \tilde{\mathfrak{g}}$, then $T_{ij} \in \mathfrak{j}$. It suffices to check that $[X_k, T_{ij}] = 0$, since the H_k act by scalars and the Y_k preserve $\tilde{\mathfrak{n}}^-$; this implies that T_{ij} generates a proper ideal.

If $k \neq i, j$ then X_k commutes with Y_i, Y_j so there is no issue. If $k = j$ then

$$\begin{aligned} \text{ad}(X_j)(T_{ij}) &= \text{ad}(X_j)\text{ad}(X_i^{1-a_{ij}})(Y_j) && (272)\\ &= \text{ad}(X_i^{1-a_{ij}})[X_j, Y_j] && (273)\\ &= \text{ad}(X_i^{1-a_{ij}})H_j. && (274)\end{aligned}$$

If $a_{ij} \leq -1$ then this vanishes and if $a_{ij} = 0$ then $[X_i, H_j] = -a_{ij}X_i = 0$. If $k = i$ then

$$\text{ad}(X_i)\text{ad}(Y_i)^{1-a_{ij}}(Y_j) = 0 \quad (275)$$

by inspecting the action of the $\mathfrak{sl}_2$-triple (H_i, X_i, Y_i) on Y_j.

The proof above also applies to the Chevalley relations involving the X_i

Recall that a linear operator T is *locally nilpotent* if for every v there is some n such that $T^n v = 0$. We claim that $\text{ad}(X_i)$ and $\text{ad}(Y_i)$ are locally nilpotent. They are locally nilpotent acting on the H_i, X_i, Y_i by a straightforward calculation, and another calculation shows that $\text{ad}(X_i)^n A = 0, \text{ad}(X_i)^m B = 0$, then $\text{ad}(X_i)^{n+m}[A, B] = 0$, so the conclusion follows (the same argument applies to the Y_i).

By local nilpotence $\exp \text{ad}(X_i), \exp \text{ad}(Y_i)$ are well-defined automorphisms of $\mathfrak{g}$. Let

$$s_i = \exp \text{ad}(X_i) \exp(-\text{ad}(Y_i)) \exp \text{ad}(X_i). \quad (276)$$

We leave it as an exercise that $s_i(H_j) = H_j - a_{ji}H_i$. In particular, s_i preserves the Cartan subalgebra $\mathfrak{h}$. Hence it must permute weight spaces, and in fact the above implies that

$$s_i(\mathfrak{g}_\alpha) = \mathfrak{g}_{r_i(\alpha)} \quad (277)$$

where $r_i \in W$ is the corresponding simple reflection. It follows that the root system $\Delta(\mathfrak{g})$ of $\mathfrak{g}$ is W-invariant and that $\dim \mathfrak{g}_{w(\alpha)} = \dim \mathfrak{g}_\alpha$ where $w \in W$. (In fact we can define the Weyl group in this way.) Hence Δ embeds into $\Delta(\mathfrak{g})$. Moreover, the root spaces of the simple roots (spanned by the X_i and Y_i) are 1-dimensional, so if we can prove that $\Delta(\mathfrak{g}) = \Delta(\mathfrak{g})$ then all of the root spaces are 1-dimensional and $\mathfrak{g}$ is finite-dimensional as desired.

Write $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$ where each piece is the quotient of the corresponding piece for $\tilde{\mathfrak{g}}$. Let $\Delta^\pm(\mathfrak{g})$ be the roots associated to $\mathfrak{n}^\pm$. We claim that

$$w(\Delta^+(\mathfrak{g})) = \Delta^+(\mathfrak{g}) \setminus A \cup -A \quad (278)$$

where $A \subset \Delta$ (not $\Delta(\mathfrak{g})$). To see this we prove using the same proof as before that $r_i(\Delta^+(\mathfrak{g})) = \Delta^+(\mathfrak{g}) \setminus \{\alpha_i\} \cup \{-\alpha_i\}$ and induct on the length. Hence W preserves the positive roots of $\Delta(\mathfrak{g})$ not in Δ, and similarly for the negative roots.

Now suppose by contradiction that Δ is not all of $\Delta(\mathfrak{g})$. Then there is some $\delta \in \Delta^+(\mathfrak{g}) \setminus \Delta^+$. There is also some $w_0 \in W$ such that $w_0(\Delta^+) = \Delta^-$. But $w_0(\delta) \in \Delta^+(\mathfrak{g})$. So $w_0(\delta)$ is both positive and negative, and this is a contradiction after writing δ as a sum of positive roots. □

We can repeat the above argument with a more general Cartan matrix associated to a more general Dynkin diagram. If it is not the Cartan matrix or Dynkin diagram of a finite-dimensional simple or semisimple Lie algebra then we get interesting Lie algebras, e.g. affine and Kac-Moody Lie algebras.

11.4 Kac-Moody algebras

Let a_{ij} be a square matrix with $a_{ii} = 2$ and $a_{ij} \in \mathbb{Z}_{\le 0}$ if $i \neq j$ such that if $a_{ij} = 0$ then $a_{ji} = 0$. We can still consider the quotient by the Chevalley relations to get a Lie algebra $\tilde{\mathfrak{g}}$ with decomposition $\tilde{\mathfrak{n}}^- \oplus \tilde{\mathfrak{h}} \oplus \tilde{\mathfrak{n}}^+$ as above; our proof did not depend on any other properties of the matrix a_{ij}. Rather than quotient by the unique maximal ideal, in general we want to quotient by the unique maximal ideal $\mathfrak{j}$ such that $\mathfrak{j} \cap \tilde{\mathfrak{h}} = 0$, which exists. This gives a Lie algebra $\mathfrak{g}$, and the same proof as above shows that the Serre relations hold in $\mathfrak{g}$.

We can even consider the action of Weyl group on roots as before, but what happens in the general case is that $\Delta(\mathfrak{g})$ usually has more roots than can be obtained by considering W acting on simple roots.

Example Consider the Cartan matrix $\left[\begin{array}{cc} 2 & -2 \\ -2 & 2 \end{array} \right]$. Here $C = H_1 + H_2$ is central. In addition we have $\alpha_1 = -\alpha_2$ and this seems bad. To fix this we introduce a new element D with trivial commutator with every generator except $[D, X_2] = X_2, [D, Y_2] = -Y_2$. This gives us a 3-dimensional Cartan subalgebra $\text{span}(H_1, H_2, D)$.

When applying reflections we get $r_1(\alpha_2) = \alpha_2 + 2\alpha_1$ and similarly $r_2(\alpha_1) = \alpha_1 + 2\alpha_2$. The root $\delta = \alpha_1 + \alpha_2$ is fixed, so $r_1(\alpha_1 + 2\alpha_2) = 3\alpha_1 + 2\alpha_2$, and so forth. We get roots of the form $n\alpha_1 + (n \pm 1)\alpha_2$ (*real* roots) and roots of the form $n\alpha_1 + n\alpha_2$ (*imaginary* roots). In particular $r_1 r_2$ has infinite order and we cannot reach an imaginary root from α_1 or α_2.

A geometric construction of this Lie algebra is as follows. Consider the *loop algebra*

$$\mathfrak{sl}_2 \otimes \mathbb{C}[t, t^{-1}] \oplus \mathbb{C}c \tag{279}$$

with c central and commutator

$$[a \otimes t^k, b \otimes t^\ell] = [a, b] \otimes t^{k+\ell} + k\delta_{k,-\ell}\text{tr}(ab)c. \tag{280}$$

With respect to the usual presentation of $\mathfrak{sl}_2$ we have

$$H_1 = H, X_1 = X, Y = Y \tag{281}$$

$$H_2 = H_1 + C, X_2 = Yt, Y_2 = Xt^{-1}. \tag{282}$$

The extra element D corresponds to the derivation $t\frac{\partial}{\partial t}$.

This algebra is called the loop algebra because we can think of $\mathfrak{sl}_2 \otimes \mathbb{C}[t, t^{-1}]$ as a model of the Lie algebra of the group of loops $S^1 \to \mathrm{SL}_2$. We get a more general construction by replacing $\mathfrak{sl}_2$ with a simple Lie algebra $\mathfrak{g}$ and replacing $\mathrm{tr}(ab)$ with the Killing form.

Example Now consider the Cartan matrix $\begin{bmatrix} 2 & -3 \\ -3 & 2 \end{bmatrix}$. The two roots α_1, α_2 satisfy $(\alpha_1, \alpha_1) = (\alpha_2, \alpha_2) = 2$ and $(\alpha_1, \alpha_2) = -3$, so the quadratic form has eigenvalues $-1, 5$, hence the roots we obtain lie on a hyperbola. In fact the real roots give us all solutions to the Pell's equation $5x^2 - y^2 = 4$ (by diagonalizing the quadratic form suitably). They can be identified with the $\alpha \in \mathbb{Z}\left[\frac{1+\sqrt{5}}{2}\right]$ of norm -1, and the Weyl group can be identified with the unit group. The other roots have positive norm.

The multiplicities $\dim \mathfrak{g}_\alpha$ are not known in general if α is not real.

11.5 More about root systems

Let Δ be a reduced root system. Let $\mathfrak{h}_\mathbb{R}$ be the real span of the roots and $(-,-)$ the inner product on $\mathfrak{h}_\mathbb{R}$. The Weyl group W acts transitively on bases.

Theorem 11.17. *1. The action of W on bases is free (so the stabilizer is trivial).*

2. The Weyl chamber

$$C^+ = \{\xi \in \mathfrak{h}_\mathbb{R} | (\alpha, \xi) \geq 0 \forall \alpha \in \Delta^+\} \tag{283}$$

is a fundamental domain for the action of W on $\mathfrak{h}_\mathbb{R}$ (any W-orbit intersects in exactly one point).

Proof. Suppose w stabilizes Δ^+. Write $w = r_1 ... r_\ell$ where the r_i are simple reflections and ℓ is minimal possible. Let $r_i = r_{\alpha_i}$, let $w_p = r_1 ... r_p$, and let $\beta_p = w_{p-1}(\alpha_p)$. Then β_p is a simple root for $w_{p-1}(\Delta^+)$, and using the relation $wr_\alpha = r_{w(\alpha)} w$ we can write

$$w = r_{\beta_\ell} ... r_{\beta_1}. \tag{284}$$

We have

$$r_{\beta_p}(w_{p-1}(\Delta^+)) = w_{p-1}(\Delta^+) \setminus \beta_p \cup \{-\beta_p\} \tag{285}$$

so each r_{β_i} acts by negating a root. Since w preserves Δ^+ then there must be some p such that $\beta_1 = \beta_p$; let's take the minimal such p. Then

$$w = r_1 u r_1 v = uv \tag{286}$$

but this contradicts the minimality of ℓ. This argument in fact shows that the length function

$$W \ni w \mapsto \ell(w) \in \mathbb{Z}_{\geq 0}, \tag{287}$$

which assigns to w the length of the minimal product representation of w as a product $r_1 ... r_{\ell(w)}$ of simple reflections, can be computed as

$$\ell(w) = |w\Delta^+ \cap \Delta^-|. \tag{288}$$

Example Let $\Delta = A_n$, so $W = S_{n+1}$. With respect to the choice of positive root $\{\varepsilon_i - \varepsilon_j | i < j\}$ we have that $|w(\Delta^+) \cap \Delta^-|$ is the number of inversions of w.

We also know now that if w is not the identity then $w\Delta^+$ has at least one root in Δ^-, hence

$$\text{int}(wC^+) \cap \text{int}(C^+) = \emptyset. \tag{289}$$

So the interiors of the translates of the Weyl chambers are disjoint. Because W acts transitively on bases (and every element of $\mathfrak{h}_{\mathbb{R}}$ is positive with respect to some basis) w esee that any W-orbit intersects C^+.

It remains to show that every W-orbit intersects C^+ exactly once. This is clear for interior points. For points on the boundary we can show that if a point on the boundary is sent by the action of W to another point on the boundary then we can perturb it to a point in the interior, and possibly after applying another reflection the perturbed point is sent to another point in the interior. □

It follows that we can cover $\mathfrak{h}_{\mathbb{R}}$ by translates of the Weyl chamber, labeling the interior of each region by a unique element of W.

We can associate to any root system Δ two lattices.

Definition The *root lattice* Q is the lattice given by the integer span of the roots. The *weight lattice* P is

$$P = \{\xi \in \mathfrak{h}_{\mathbb{R}} | (\xi, \alpha_i^\vee) \in \mathbb{Z}\}. \tag{290}$$

The basis of P dual to the basis $\alpha_1^\vee, ..., \alpha_n^\vee$ is the basis of *fundamental weights* $\omega_1, ..., \omega_n$.

The coroot lattice is the dual of the root lattice of $\Delta^\vee$. Note that by definition $Q \subset P$, and in fact it has finite index. In terms of the fundamental weights, the simple roots can be expressed as

$$\alpha_i = \sum a_{ji} \omega_j \tag{291}$$

where a_{ij} is the Cartan matrix. Hence $|P/Q| = \det(a_{ij})$.

Example Let $\Delta = A_n$. We realized the roots as $\varepsilon_i - \varepsilon_j, i \neq j$ with simple roots $\varepsilon_i - \varepsilon_{i+1}, i = 1, ..., n$. We can write

$$Q = \{\sum n_i \varepsilon_i | \sum n_i = 0\}. \tag{292}$$

Write $\varepsilon \mapsto \overline{\varepsilon}$ for the orthogonal projection of the ε_i from $\mathbb{R}^{n+1}$ to $\mathbb{R}^n$, so

$$\overline{\varepsilon_1} = \varepsilon_1 - \frac{\varepsilon_1 + ... + \varepsilon_{n+1}}{n+1} \tag{293}$$

and so forth. Then we can write

$$\omega_1 = \overline{\varepsilon_1}, \omega_2 = \overline{\varepsilon_1} + \overline{\varepsilon_2}, ..., \omega_n = \overline{\varepsilon_1} + ... + \overline{\varepsilon_n}. \tag{294}$$

The root lattice Q has index $n+1$ in the weight lattice P, so $|P/Q| = \det(a_{ij}) = n+1$. This reflects the fact that the center of $\mathrm{SL}_{n+1}(\mathbb{C})$, the simply connected complex Lie group with Lie algebra $\mathfrak{sl}_{n+1}(\mathbb{C})$, is $\mathbb{Z}/(n+1)\mathbb{Z}$.

Definition The *Weyl vector* ρ is

$$\rho = \omega_1 + ... + \omega_n = \frac{1}{2} \sum_{\alpha \in \Delta^+} \alpha. \tag{295}$$

We have $(\rho, \alpha_i^\vee) = 1$ for $i = 1, ..., n$ and hence $r_i(\rho) = \rho - \alpha_i$.

Definition The set of *dominant weights* is $P^+ = C^+ \cap P$.

Note that every W-orbit of P intersects P^+ exactly once. We will show next time that when Δ is the root system of a semisimple Lie algebra $\mathfrak{g}$, the dominant weights P^+ parameterize the set of finite-dimensional irreducible representations of $\mathfrak{g}$.

12 Highest weight theory

Let $\mathfrak{g}$ be a semisimple finite-dmensional Lie algebra over $\mathbb{C}$.

Definition A *weight module* is a $\mathfrak{g}$-module M which is semisimple as an $\mathfrak{h}$-module, hence which has a weight decomposition

$$M = \bigoplus_{\lambda \in \mathfrak{h}^*} M_\lambda \tag{296}$$

where

$$M_\lambda = \{m \in M | Hm = \lambda(h) m \forall h \in \mathfrak{h}\}. \tag{297}$$

Some comments:

- In particular, since $\mathfrak{h}$ consists of semisimple elements and these act semisimply in finite-dimensional representations, every finite-dimensional $\mathfrak{g}$-module is a weight module.

- By the standard calculation, $\mathfrak{g}_\alpha M_\lambda \subseteq M_{\lambda+\alpha}$.

- The direct sum $M \oplus N$ of two weight modules M, N is a weight module with weights $P(M \oplus N)$ the union $P(M) \cup P(N)$ (with multiplicity) of the weights of M, N. The tensor product $M \otimes N$ of two weight modules is also a weight module with weights the sum of the sets of weights of M, N; more precisely,

$$(M \otimes N)_\lambda \cong \bigoplus_{\mu+\nu=\lambda} M_\mu \otimes N_\nu. \tag{298}$$

- Submodules and quotient modules of weight modules are weight modules.

- If M is a finite-dimensional $\mathfrak{g}$-module then $P(M)$ lies in the weight lattice P. To see this, note that $\lambda \in P$ iff $(\lambda, \alpha_i^\vee) \in \mathbb{Z}$ iff $\lambda(H_i) \in \mathbb{Z}$, and this follows from looking at the action of the $\mathfrak{sl}_2$-triple (H_i, X_i, Y_i).

Example Earlier we saw an infinite-dimensional example of a weight module, namely the Verma module

$$M(\lambda) = U(\mathfrak{sl}_2) \otimes_{U(\mathfrak{b})} C_\lambda \tag{299}$$

for $\mathfrak{sl}_2$ associated to λ. The weights of this module were $\lambda - 2n$ for $n \in \mathbb{Z}_{\geq 0}$.

Fix a choice of positive roots Δ^+ and consider the corresponding triangular decomposition

$$\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+ \tag{300}$$

of $\mathfrak{g}$. Then $\mathfrak{n}^+$ is nilpotent (repeatedly adding positive roots eventually gets us out of the root system) and $\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}^+$, the *Borel subalgebra*, is solvable, and $\mathfrak{n}^+ = [\mathfrak{b}, \mathfrak{b}]$, hence the abelianization of $\mathfrak{b}$ is $\mathfrak{h}$.

We can apply Lie's theorem to obtain the following.

Lemma 12.1. *If M is a finite-dimensional $\mathfrak{g}$-module, there is some λ and some $v \in M_\lambda$ such that $hv = \lambda(h)v$ for all $h \in \mathfrak{h}$ and $\mathfrak{n}^+ v = 0$. Equivalently, if C_λ is the corresponding 1-dimensional representation of $\mathfrak{b}$, there is some λ such that*

$$Hom_{\mathfrak{b}}(C_\lambda, M) \neq 0. \tag{301}$$

v is called a highest weight *vector of M and λ is called a* highest weight.

Lemma 12.2. *If M is a finite-dimensional irreducible $\mathfrak{g}$-module, then M has a unique highest weight and a unique, up to scale, highest weight vector.*

Proof. Suppose $v \in M_\lambda, w \in M_\mu$ are two highest weight vectors. The triangular decomposition induces a decomposition $U(\mathfrak{g}) = U(\mathfrak{n}^{-1})U(\mathfrak{h})U(\mathfrak{n}^+)$ which implies that $M(v) = U(\mathfrak{g})v = U(\mathfrak{n}^{-1})v$. Hence the weights of $M(v)$ lie in the set $\lambda - \Delta^+$, and the same is true of $M(w)$, hence $\lambda - \mu$ and $\mu - \lambda$ are both in Δ^+, but this is possible iff $\lambda = \mu$; moreover, $M(v)$ consists of $\text{span}(v)$ and then weight spaces with strictly lower weight than λ (with respect to the ordering determined by Δ^+), so if $\lambda = \mu$ then $w \in \text{span}(v)$. □

Hence we have a map which assigns to an irreducible finite-dimensional representation a highest weight. We want to know which weights occur as highest weights.

Lemma 12.3. *A highest weight λ must lie in P^+ (the dominant weights).*

This follows because $\lambda(H_i) = (\lambda, \alpha_i^\vee) \geq 0$ for a highest weight λ by inspection of the corresponding $\mathfrak{sl}_2$-triple.

The big theorem is the following.

Theorem 12.4. *The map assigning to an irreducible finite-dimensional representation its highest weight in P^+ is a bijection.*

To prove surjectivity we work as follows.

Definition Let $\lambda \in \mathfrak{h}^*$. The *Verma module with highest weight λ* is

$$M(\lambda) = U(\mathfrak{g}) \otimes_{U(\mathfrak{b})} C_\lambda. \tag{302}$$

Theorem 12.5. *If V is an irreducible finite-dimensional representation of $\mathfrak{g}$ with highest weight λ, then V is a quotient of $M(\lambda)$.*

Proof. An application of the tensor-hom adjunction (which can be thought of as Frobenius reciprocity here) gives

$$\text{Hom}_{\mathfrak{b}}(C_\lambda, V) \cong \text{Hom}_{\mathfrak{g}}(M_\lambda, V). \tag{303}$$

The LHS is nonzero, so the RHS is also nonzero. But the image of any nonzero morphism into an irreducible representation is surjective. □

The decomposition $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{b}$ gives $U(\mathfrak{g}) \cong U(\mathfrak{n}^-) \otimes U(\mathfrak{b})$, hence $M(\lambda) \cong U(\mathfrak{n}^-) \otimes C_\lambda$. The action of $h \in \mathfrak{h}$ on this module is

$$\begin{aligned} h(X \otimes v) &= \text{ad}(h)(X) \otimes v + Xh \otimes v & (304)\\ &= \text{ad}(h)(X) \otimes v + X \otimes hv & (305)\\ &= \text{ad}(h)(X) \otimes v + \lambda(h) X \otimes v. & (306) \end{aligned}$$

If $\alpha \in \Delta^+$ and $Y_\alpha \in \mathfrak{g}_{-\alpha}$, by picking an ordering we can construct a PBW basis $Y_{\beta_1}...Y_{\beta_k}$ of monomials with $\beta_1 \leq ... \leq \beta_k$ in $U(\mathfrak{n}^-)$. The action of $h \in \mathfrak{h}$ on this basis is

$$\begin{aligned} \text{ad}(h)(Y_{\beta_1}...Y_{\beta_k}) &= \sum Y_{\beta_1}...\text{ad}(h)(Y_{\beta_i})...Y_{\beta_k} & (307)\\ &= -(\beta_1 + ... + \beta_k)(h) Y_{\beta_1}...Y_{\beta_k}. & (308) \end{aligned}$$

So $U(\mathfrak{n}^-)$, and therefore $M(\lambda)$, is a weight module, and the set of weights is $\lambda - Q^+$ where Q^+ is the span of the positive roots. Moreover,

$$\dim M(\lambda)_{\lambda-\gamma} = K(\gamma) \tag{309}$$

does not depend on λ; it is the *Kostant partition function* counting the number of ways to write γ as a sum of positive roots. This is in particular finite, and in addition $\dim M(\lambda)_\lambda = 1$; moreover $M(\lambda)$ is generated by this weight space. It follows that if N is a proper submodule then $N_\lambda = 0$. Hence the sum of all proper submodules is a proper submodule, from which it follows that $M(\lambda)$ has a unique maximal proper submodule $N(\lambda)$, hence a unique simple quotient

$$L(\lambda) = M(\lambda)/N(\lambda). \tag{310}$$

We would like to know when $L(\lambda)$ is finite-dimensional. So far we know that if this is the case then $\lambda \in P^+$.

Definition A weight vector w is *singular* if $\mathfrak{n}^+ w = 0$.

Lemma 12.6. *Let $w \in M(\lambda)_\mu$ be a singular vector, with $\lambda \neq \mu$. Then $w \in N(\lambda)$.*

Proof. The set of weights of $U(\mathfrak{g})w$ lies in $\mu - Q^+$, which does not contain λ, hence $U(\mathfrak{g})w$ is a proper submodule. □

Write $\lambda(H_i) = (\lambda, \alpha_i^\vee) = m \in \mathbb{Z}_{\geq 0}$.

Lemma 12.7. *Let $v_\lambda \in C_\lambda$ be nonzero. Then $Y_i^{m+1} \otimes v_\lambda \in M(\lambda)_{-\lambda-(m+1)\alpha_i}$ is singular.*

Proof. If $j \neq i$ then X_j commutes with Y_i^{m+1} and the statement is clear. If $j = i$ then this computation reduces to a computation involving the $\mathfrak{sl}_2$-triple (H_i, X_i, Y_i) which we've done already. □

Corollary 12.8. *If $\lambda \in P^+$ then $X_1, ..., X_n, Y_1, ..., Y_n$ act locally nilpotently on $L(\lambda)$.*

Proof. The X_i already act locally nilpotently on $M(\lambda)$, since they eventually escape the set of weights. To show that the Y_i also act locally nilpotently, let $v \in L(\lambda)_\lambda$ be nonzero. We showed that $Y_i^{m_i+1} v = 0$ where $m_i = \lambda(H_i)$. Any other $w \in L(\lambda)$ is of the form Yv where $Y \in U(\mathfrak{n}^-)$. We have

$$Y_i Y v = \text{ad}(Y_i)(Y)v + Y Y_i v. \tag{311}$$

Since $\mathfrak{n}^-$ is nilpotent, $\text{ad}(Y_i)$ acts locally nilpotently on $U(\mathfrak{n}^-)$. The adjoint action on the left and right multiplication commute, so we can write

$$Y_i^a Y v = \sum \binom{a}{b} \text{ad}(Y_i)^{a-b}(Y) Y_i^b v \tag{312}$$

and we conclude that this vanishes for sufficiently large a. □

This lets us write down the exponentials

$$g_i = \exp X_i \exp(-Y_i) \exp X_i \in \mathrm{End}(L(\lambda)) \tag{313}$$

and

$$s_i = \exp \mathrm{ad}(X_i) \exp(-\mathrm{ad}(Y_i)) \exp \mathrm{ad}(X_i) = \mathrm{Ad}(g_i) \in \mathrm{End}(\mathfrak{g}) \tag{314}$$

as before. If $v \in L(\lambda)_\mu$ we compute that

$$\begin{aligned} hg_iv &= g_ig_i^{-1}hg_iv && (315)\\ &= g_is_i^{-1}(h)v && (316)\\ &= \mu(s_i^{-1}(h))g_iv && (317)\end{aligned}$$

hence that $g_iL(\lambda)_\mu = L_{r_i(\mu)}$, so the set of weights of $L(\lambda)$ is W-invariant. Since P^+ is a fundamental domain for the action of W on P, the set of weights $P(L(\lambda))$ satisfies

$$P(L(\lambda)) = \bigcup_{w \in W} w(P(L(\lambda)) \cap P^+). \tag{318}$$

Lemma 12.9. $(\lambda - Q^+) \cap P^+$ *is finite.*

Proof. let ρ be half the sum of the positive roots before. Every element of P^+ satisfies $(-, \rho) \geq 0$. If $\lambda - \sum n_i\alpha_i \in \lambda - Q^+$ has this property, then it satisfies

$$(\lambda, \rho) \geq \sum n_i \frac{(\alpha_i, \alpha_i)}{2}. \tag{319}$$

But there are only finitely many possibilities here since $n_i \in \mathbb{Z}_{\geq 0}$. □

Corollary 12.10. *If* $\lambda \in P^+$*, then* $L(\lambda)$ *is finite-dimensional.*

12.1 Fundamental weights

Every weight in P^+ is a non-negative integer linear combination of the fundamental weights $\omega_1, ..., \omega_n$.

Definition $L_i = L(\omega_i)$ are the *fundamental representations* of $\mathfrak{g}$.

Lemma 12.11. *Any irreducible finite-dimensional representation of* $\mathfrak{g}$ *is a subrepresentation of a tensor product of fundamental representations.*

Proof. Suppose $\omega = \sum m_i\omega_i \in P^+$. The tensor product

$$\bigotimes_i L_i^{\otimes m_i} \tag{320}$$

contains a vector

$$w = \bigotimes_i v_i^{\otimes m_i} \tag{321}$$

of weight ω, hence contains $L(\omega)$ as a subrepresentation. $\square$

Example (Type A_n) Let $\mathfrak{g} = \mathfrak{sl}_{n+1}$ with fundamental weights $\omega_i = \overline{\varepsilon_1 + ... + \varepsilon_i}$. With V the defining representation of $\mathfrak{g}$, it turns out that

$$L_k \cong \Lambda^k(V). \tag{322}$$

We can verify this by finding a heighest weight vector in $\Lambda^k(V)$. Let $v_1, ..., v_n \in V$ have weights $\overline{\varepsilon_1}, ..., \overline{\varepsilon_n}$. Then $v_1 \wedge ... \wedge v_k \in \Lambda^k(V)$ has the desired weight, and it is not hard to show that $\Lambda^k(V)$ is irreducible. These representations have the property that they are *minuscule*: the set of weights $P(L(\lambda))$ is smallest possible, namely the orbit $W\lambda$ of the action of W on λ.

We now know that every representation of $\mathfrak{sl}_{n+1}$ is contained in a tensor product of copies of the exterior powers of V. In fact they are contained in a tensor product of copies of V. This is the story of Schur-Weyl duality.

The adjoint representation of $\mathfrak{sl}_{n+1}$ is naturally a subrepresentation of $V \otimes V^*$, and $V^* \cong \Lambda^n(V)$.

Example (Type C_n) Let $\mathfrak{g} = \mathfrak{sp}(2n)$ with defining representation $V \cong \mathbb{C}^{2n}$. Fix a basis $v_1, ..., v_{2n}$ and skew-symmetric form ω such that $\omega(v_i, v_{i+n}) = 1$ and $\omega(v_i, v_{j+n}) = 0$ if $i \neq j$. The simple roots are $\varepsilon_i - \varepsilon_{i+1}, 1 \le i \le n-1$ and $2\varepsilon_n$. The fundamental weights are $\omega_i = \varepsilon_1 + ... + \varepsilon_i$. The basis above has weights $\varepsilon_1, ..., \varepsilon_n, -\varepsilon_1, ..., -\varepsilon_n$, hence we have fundamental representations

$$L_k \subsetneq \Lambda^k(V), 1 \le k \le n. \tag{323}$$

If we take higher exterior powers we get isomorphisms

$$\Lambda^k(V) \cong \Lambda^{2n-k}(V) \tag{324}$$

so we get nothing new. When $k = 2$ we have

$$\Lambda^2(V) \cong \mathbb{C}w \oplus L_2 \tag{325}$$

where $w = \sum v_i \wedge v_{i+n}$ is the bivector dual to the skew-symmetric form, hence is preserved by the action of $\mathfrak{g}$. Wedging with w gives $\mathfrak{g}$-invariant maps

$$\Lambda^{k-2}(V) \ni \gamma \mapsto \gamma \wedge w \in \Lambda^k(V) \tag{326}$$

and these are injections for $k \le n$. The corresponding quotients are L_k. Said another way, contracting with ω gives $\mathfrak{g}$-invariant maps $\Lambda^k(V) \to \Lambda^{k-2}(V)$, and the corresponding kernels are L_k.

The adjoint representation lies in $V \otimes V$ since V is self-dual. It has highest weight $2\varepsilon_1$, and in fact is $S^2(V)$.

Example (Type D_n) Let $\mathfrak{g} = \mathfrak{so}(2n)$ with defining representation $V \cong \mathbb{C}^{2n}$. The simple roots are $\varepsilon_i - \varepsilon_{i+1}, 1 \le i \le n-1$ together with $\varepsilon_{n-1} + \varepsilon_n$. The corresponding fundamental weights are $\omega_i = \varepsilon_1 + ... + \varepsilon_i, 1 \le i \le n-2$ and two exceptional weights

$$\omega_{n-1} = \frac{\varepsilon_1 + ... + \varepsilon_{n-1} - \varepsilon_n}{2}, \omega_n = \frac{\varepsilon_1 + ... + \varepsilon_{n-1} + \varepsilon_n}{2}. \tag{327}$$

The corresponding irreducible representations cannot be realized as subrepresentations of tensor powers of V. This reflects the fact that all such subrepresentations are representations of $\mathrm{SO}(2n)$, which is not simply connected, and there are some extra representations of the universal cover $\mathrm{Spin}(2n)$, the *spin group*, on which the fundamental group acts nontrivially. L_{n-1} and L_n are *spinor representations*.

Here the adjoint representation is $\Lambda^2(V)$ with highest weight $\varepsilon_1 + \varepsilon_2$.

Example (Type B_n) Let $\mathfrak{g} = \mathfrak{so}(2n+1)$ with defining representation $V \cong \mathbb{C}^{2n+1}$. The fundamental weights are $\omega_i = \varepsilon_1 + ... + \varepsilon_i, 1 \le i \le n-1$ and one exceptional weight

$$\omega_n = \frac{\varepsilon_1 + ... + \varepsilon_n}{2} \tag{328}$$

corresponding to one spinor representation. The orbit of the Weyl group action on ω_n consists of weights of the form $\frac{\pm\varepsilon_1 \pm ... \pm \varepsilon_n}{2}$. These are the vertices of a cube containing no integral weights. There are $\dim L_n = 2^n$ of them. We will construct this representation using Clifford algebras.

12.2 Clifford algebras

We work over a base field of characteristic zero, which can be taken to be $\mathbb{C}$, and which is algebraically closed unless otherwise stated.

Definition Let V be a finite-dimensional vector space and b be a symmetric bilinear form on V. The *Clifford algebra* $\mathrm{Cl}(V, b)$ is the quotient of the tensor algebra $T(V)$ by the relations

$$v \otimes w + w \otimes v = b(v, w). \tag{329}$$

We denote Clifford multipliation by $\cdot$.

Example Suppose $V = \mathbb{R}^2$ with basis i, j and b satisfies $b(i, j) = 0, b(i, i) = b(j, j) = -2$. The corresponding relations are $ij + ji = 0, i^2 = -1, j^2 = -1$, and so $\mathrm{Cl}(V, b)$ is the quaternions.

If b is nondegenerate, then all Clifford algebras of the same dimension are isomorphic. In fact the following is true.

Proposition 12.12. *If b is nondegenerate and $\dim V = 2n$, then $Cl(V,b) \cong M_{2^n}$ (the ring of $2^n \times 2^n$ matrices).*

Proof. Write $V = V_+ \oplus V_-$ as the direct sum of two maximally isotropic subspaces (so $b(V_+, V_+) = b(V_-, V_-) = 0$). We will construct a representation of $\text{Cl}(V,b)$ on $\Lambda(V_+)$. Any $v \in V_+$ acts by the wedge product

$$v(u) = v \vee u. \tag{330}$$

Any $w \in V_-$ acts inductively via $w(1) = 0$ and

$$w(v \vee u) = b(w,v)u' - v \wedge w(u'). \tag{331}$$

We can think of this action as an action by differential operators; the Clifford algebra is analogous to the Weyl algebra here. In particular the elements of V_- act by derivations and the elements of V_+ act by multiplication. By first acting by derivations and then acting by multiplications we can show that $\Lambda(V_+)$ is an irreducible representation. It is also faithful, so we get a morphism

$$\text{Cl}(V,b) \to \text{End}(\Lambda(V_+)) \tag{332}$$

and either by a dimension count or by the Jacobson density theorem it follows that this map is an isomorphism. □

A philosophical remark. For V as above it is clear that $\text{SO}(2n) \cong \text{SO}(V,b)$ acts on $\text{Cl}(V,b)$. In general, suppose R is a finite-dimensional algebra, M is a simple R-module, and G a group acting by automorphisms on R. Then we can twist M by $g \in G$ to get a new module gM. But if R has a unique simple module then $gM \cong M$. Let $\Phi_g \in \text{Aut}(M)$ be the corresponding isomorphism, hence

$$g(r)m = \Phi_g r \Phi_g^{-1} m. \tag{333}$$

By Schur's lemma, Φ_g is unique up to scalars, hence we obtain a projective representation $\Phi : G \to \text{PGL}(M)$. In the case above we obtain a representation

$$\text{SO}(2n) \to \text{PSL}(\Lambda(V_+)) \tag{334}$$

which lifts to simply connected covers

$$\text{Spin}(2n) \to \text{SL}(\Lambda(V_+)). \tag{335}$$

These are spin representations.

Here is a second construction. This time we will embed $\mathfrak{g} \cong \mathfrak{so}(2n)$ into $\text{Cl}(2n)$. Looking at $V \subset \text{Cl}(2n)$ we claim we can take $\mathfrak{g}$ to be the span $[V,V]$ of the commutator brackets $[v,w]$ with $v,w \in V$. First, note that $[v,w] = vw - wv = 2vw + b(v,w)$. We compute that for $x \in V$ we have

$$\begin{aligned}
\mathrm{ad}(vw)x &= vwx - xvw && (336)\\
&= vwx + vxw + b(v,x)w && (337)\\
&= vb(w,x) + b(v,x)w \in V. && (338)
\end{aligned}$$

Similarly, we compute that $[[V,V],[V,V]] \subseteq [V,V]$. Hence $[V,V]$ is a Lie subalgebra of $\mathrm{Cl}(2n)$ which we can explicitly verify must be $\mathfrak{so}(2n)$.

The action of $[V,V]$ only changes degree by a multiple of 2, hence as a representation of $[V,V]$, the representation $\Lambda(V_+)$ breaks up into a sum $\Lambda_0(V_+) \oplus \Lambda_1(V_+)$ of its even and odd parts. These are the two spin representations we found earlier using highest weights, and we can verify this by introducing suitable coordinates everywhere.

Now, $V \oplus [V,V]$ is a Lie algebra containing $\mathfrak{so}(2n)$ such that V has nontrivial commutator with $[V,V]$. In fact it must be $\mathfrak{so}(2n+1)$, and so we have also constructed the spinor representation $\Lambda(V_+)$ of $\mathfrak{so}(2n+1)$.

13 Back to Lie groups

13.1 Centers

Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$ with triangular decomposition $\mathfrak{n}^+ \oplus \mathfrak{h} \oplus \mathfrak{n}^-$. Let P be its weight lattice and Q its root lattice, which live in $\mathfrak{h}^*$. From these we can write down dual lattices

$$Q^* = \{h \in \mathfrak{h} | \alpha_i(h) \in \mathbb{Z} \forall i\} \tag{339}$$

$$P^* = \{h \in \mathfrak{h} | \omega_i(h) \in \mathbb{Z} \forall i\} \tag{340}$$

which sit inside $\mathfrak{h}$. The groups P/Q and Q^*/P^* are noncanonically isomorphic.

Theorem 13.1. *Let G be a complex connected Lie group with Lie algebra $\mathfrak{g}$.*

1. *G is an algebraic subgroup of some $GL(V)$.*
2. *If G is simply connected, then $Z(G) \cong Q^*/P^*$.*
3. *Connected complex Lie groups with Lie algebra $\mathfrak{g}$ are naturally in bijection with subgroups of Q^*/P^*.*

Note that the first statement fails over $\mathbb{R}$ since, for example, $\widetilde{SL_2}(\mathbb{R})$ has no faithful finite-dimensional representations.

First we will need a lemma.

Lemma 13.2. *If G has a faithful representation V, then G is an algebraic subgroup of $GL(V)$.*

Proof. We first consider the case that V is irreducible.

Since $\mathfrak{g}$ is semisimple, any representation $G \to \mathrm{GL}(V)$ lands in $\mathrm{SL}(V)$. Consider the normalizer

$$N_{\mathrm{SL}(V)}(\mathfrak{g}) = \{g \in \mathrm{SL}(V) | \mathrm{Ad}_g(\mathfrak{g}) = \mathfrak{g}\}. \tag{341}$$

Since this is an algebraic condition, we get an algebraic subgroup. Its connected component of the identity $N^0_{\mathrm{SL}(V)}$ is also an algebraic subgroup. We want to show that it coincides with G. It suffices to show that the Lie algebras coincide, or equivalently that $N_{\mathfrak{sl}(V)}(\mathfrak{g}) = \mathfrak{g}$ ($\mathfrak{g}$ coincides with its normalizer in $\mathfrak{sl}(V)$). Since $\mathfrak{g}$ is semisimple, it acts semisimply on $\mathfrak{sl}(V)$, hence we can write

$$\mathfrak{sl}(V) \cong \mathfrak{g} \oplus \mathfrak{m} \tag{342}$$

where $[\mathfrak{g}, \mathfrak{m}] \subseteq \mathfrak{m}$. Since $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$, if we can find $X \in \mathfrak{sl}(V)$ not in $\mathfrak{g}$ such that $[X, \mathfrak{g}] \subseteq \mathfrak{g}$, then we can find X satisfying the stronger condition $[X, \mathfrak{g}] = 0$. But then X is a $\mathfrak{g}$-endomorphism of V, so by Schur's lemma is multiplication by a scalar, and since $X \in \mathfrak{sl}(V)$ it has trace zero and hence is zero.

In the general case we apply the above to each irreducible summand. □

Lemma 13.3. *Let $Y \subsetneq X$ be a proper inclusion of Zariski closed subsets of $\mathbb{C}^N$.*

1. *If X is smooth and connected then $X \setminus Y$ is connected.*
2. *The map $\pi_1(X \setminus Y) \to \pi_1(X)$ is surjective.*

Proof. It suffices to show that if $\gamma : [0, 1] \to X$ is a path such that $\gamma(0), \gamma(1) \notin Y$ then we can perturb γ such that it entirely avoids Y. This proves both of the above statements.

We can show that it is possible to choose γ to be piecewise analytic. Then γ intersects Y at at most finitely many points, and we can avoid these one at a time. □

Lemma 13.4. *In $GL(V)$ consider the Borel subgroup B of upper triangular matrices and the subgroup N^- of lower triangular matrices with 1s on the diagonal. The multiplication map*

$$N^- \times B \to GL(V) \tag{343}$$

is injective with Zariski open image.

Proof. If all principal minors of an element of $\mathrm{GL}(V)$ are nonzero then we can perform Gaussian elimination, which gives a unique element of N^- and a unique element of B. Otherwise there are no such elements. □

Proposition 13.5. *Let $\tilde{G}$ be a simply connected Lie group with Lie algebra $\mathfrak{g}$ and let $V = L_1 \oplus ... \oplus L_n$ be the sum of its fundamental representations. Then $\rho : \tilde{G} \to GL(V)$ is an embedding, and in particular V is a faithful representation.*

Proof. Let $G = \rho(\tilde{G})$. Since all the generators corresponding to simple roots of $\mathfrak{g}$ act non-trivially on V, we know that V is a faithful representation of $\mathfrak{g}$, hence G also has Lie algebra $\mathfrak{g}$ and it suffices to show that $\pi_1(G)$ is trivial.

With respect to the triangular decomposition of $\mathfrak{g}$, choose a basis of V such that $\mathfrak{n}^+$ is upper-triangular, $\mathfrak{n}^-$ is lower-triangular, and $\mathfrak{h}$ is diagonal. Let $B = G \cap B_{\text{GL}(V)}$ and $N^- = G \cap N^-_{\text{GL}(V)}$. Let $U \in \text{GL}(V)$ be the image of the multiplication map $N^- \times B \to \text{GL}(V)$, and let $G' = U \cap G$. Since G' is G minus a Zariski closed set, G' is connected.

We claim that if $g \in G'$ then $g = xy$ where $x \in N^-, y \in B$ (this is slightly stronger than what we get from the fact that $g \in U$). In a sufficiently small neighborhood of the identity this follows because we can write such neighborhoods in the form $\exp \mathfrak{n}^- \times \exp \mathfrak{b}$. Analyticity lets us conclude that this is true everywhere.

We now have a surjection $\pi_1(G') \to \pi_1(G)$, so to show that the latter is trivial it suffices to show that the former is trivial. But $G' = \exp \mathfrak{n}^- \times \exp \mathfrak{h} \times \exp \mathfrak{n}^+$ where $\exp \mathfrak{n}^-$ and $\exp \mathfrak{n}^+$ are contractible, so letting $T = \exp \mathfrak{h}$ it follows that $\pi_1(T) \cong \pi_1(G')$, so there is a surjection $\pi_1(T) \to \pi_1(G)$.

The sl_2-triples (H_i, X_i, Y_i) in $\mathfrak{g}$ give a multiplication map

$$\text{SL}_2(\mathbb{C})^n \to G \tag{344}$$

which realizes T as the product $T_1 \times ... \times T_n$ of the exponentials of $H_1, ..., H_n$. The inclusion $T \to G$ factors through the above map, hence the induced map on π_1 is surjective, but $\text{SL}_2(\mathbb{C})^n$ is simply connected and the conclusion follows.

Above we constructed a decomposition $G' = N^- T N^+$ where $N^- = \exp \mathfrak{n}^-, T = \exp \mathfrak{h}, N^+ = \exp \mathfrak{n}^+$. What is the kernel of $\exp : \mathfrak{h} \to T$? Looking at the weight decomposition, this is the set of $h \in \mathfrak{h}$ such that $e^{\mu(h)} = 1$ for all weights μ of V. Since V contains all fundamental weights, this is equivalently

$$\{h \in h | \omega_i(h) \in 2\pi i \mathbb{Z}\} \tag{345}$$

hence is $2\pi i$ times the lattice P^*. On the other hand, by Schur's lemma $Z(G) \subseteq T$, hence $Z(G) = \ker(\text{Ad}(T))$. Hence we have a diagram of short exact sequences

$$\begin{array}{ccccccccc} 0 & \longrightarrow & 2\pi i P^* & \longrightarrow & \mathfrak{h} & \xrightarrow{\exp} & T & \longrightarrow & 0 \\ & & \downarrow & & \downarrow \text{id} & & \downarrow & & \\ 0 & \longrightarrow & ? & \longrightarrow & \mathfrak{h} & \xrightarrow{\exp \text{ad}} & \text{Ad}(T) & \longrightarrow & 0 \end{array} \tag{346}$$

where the mystery group is $2\pi i Q^*$ by the same analysis as above, and this gives $Z(G) \cong Q^*/P^*$.

P can be interpreted as the lattice of characters of T, and Q can be interpreted as the lattice of characters of $\text{Ad}(T)$. Moreover, any subgroup in $Z(G)$ corresponds to a sublattice between P^* and Q^*, and we can use this to determine how many and which connected complex Lie groups have a given semisimple Lie algebra. $\square$

13.2 Characters

Definition Let M be a finite-dimensional representation of a semisimple Lie algebra $\mathfrak{g}$. Its (formal) *character* is the formal expression

$$\mathrm{ch}(M) = \sum_{\mu \in P(M)} (\dim M_\mu) e^\mu \tag{347}$$

in the group ring $R = \mathbb{Z}[P]$ of P, written multiplicatively (so $e^{\mu+\nu} = e^\mu e^\nu$).

We can think about formal characters as functions on T by interpreting e^μ to be the function $\exp(h) \mapsto e^{\mu(h)}$.

Definition Let M be a finite-dimensional representation of a Lie group G. Its *character* is the function

$$\chi_M(g) = \mathrm{tr}_M(g) \tag{348}$$

on G.

If G has semisimple Lie algebra $\mathfrak{g}$, then the two definitions agree when restricted to $T = \exp \mathfrak{h}$. Note that characters are conjugation-invariant and that GTG^{-1} is dense in G, so the restriction of a character to T uniquely determines it.

The (formal) character ch is additive and multiplicative with respect to direct sum and tensor product. It is also W-invariant. Hence it defines a homomorphism from the representation ring of $\mathfrak{g}$ to the invariant subring R^W.

Example Let $\mathfrak{g} = \mathfrak{sl}_2$, so $P = \mathbb{Z}\omega$ where $\omega = \frac{\alpha}{2}$. Write R as $\mathbb{Z}[z]$. The Weyl group action is $z \mapsto z^{-1}$, so all characters must be symmetric Laurent polynomials. In fact they have the form $z^n + z^{n-2} + ... + z^{-(n-2)} + z^{-n}$.

Theorem 13.6. *If $\mathfrak{g}$ is a semisimple Lie algebra the characters $ch(L(\lambda))$ form a basis of R^W.*

Proof. First observe that the sums $\frac{1}{|\mathrm{Stab}(\lambda)|} \sum_{w \in W} e^{w(\lambda)}$ form a basis of R^W. Consider the partial order $\mu \le \lambda \Leftrightarrow \lambda - \mu \in Q^+$. Then

$$\mathrm{ch}(L(\lambda)) = \frac{1}{|\mathrm{Stab}(\lambda)|} \sum_{w \in W} e^{w(\lambda)} + \sum_{\mu \in P^+, \mu < \lambda, w \in W} \frac{c_\mu}{|\mathrm{Stab}(\mu)|} e^{w(\mu)} \tag{349}$$

so the matrix describing these characters in terms of the basis is triangular with entries 1 on the diagonal, hence is invertible. □

Theorem 13.7. *(Weyl character formula) For $\lambda \in P^+$, the character of $L(\lambda)$ is given by*

$$ch(L(\lambda)) = \frac{\sum_{w \in W} \mathrm{sgn}(w) e^{w(\lambda+\rho)}}{\sum_{w \in W} \mathrm{sgn}(w) e^{w(\rho)}} \tag{350}$$

where $\mathrm{sgn}(w)$ *is alternatively either* $(-1)^{\ell(w)}$ *or* $\det(w)$ *and* $\rho = \omega_1 + ... + \omega_n$ *as before.*

Note that characters are W-invariant but the Weyl character formula expresses a charater as a quotient of two W-skew-invariant expressions. If we let

$$D_\mu = \sum_{w \in W} \operatorname{sgn}(w) e^{w(\mu)} \tag{351}$$

then we can show that, as μ ranges over the interior P^{++} of P^+, then D_μ forms a basis of the W-skew-invariants R^W_{alt}. Multiplication by D_ρ gives a map

$$D_\rho : R^W \to R^W_{\text{alt}} \tag{352}$$

and with respect to the basis S_μ of R^W we constructed earlier we have $S_\mu D_\rho = D_{\mu+\rho} +$ lower terms, hence multiplication by D_ρ gives an isomorphism. In particular, $\frac{D_{\lambda+\rho}}{D_\rho} \in R^W$, so the expression in the Weyl character formula lands in the correct space.

Example let $\mathfrak{g} = \mathfrak{sl}_n$. Then R can be identified with the quotient of the algebra of Laurent polynomials in n variables $z_1, ..., z_n$ by the relation $z_1 ... z_n = 1$. The invariant subring R^W consists of the symmetric polynomials. Here

$$\rho = \omega_1 + ... + \omega_{n-1} = (n-1)\overline{\varepsilon_1} + (n-2)\overline{\varepsilon_2} + ... + \overline{\varepsilon_{n-1}} \tag{353}$$

so we can write

$$e^\rho = z_1^{n-1} z_2^{n-2} ... z_{n-1}. \tag{354}$$

The elements of P^+ have the form $a_1\overline{\varepsilon_1} + ... + a_{n-1}\overline{\varepsilon_{n-1}}$ where $a_1 \geq ... \geq a_{n-1} \geq 0$. If $\mu = b_1\overline{\varepsilon_1} + ... + b_n\overline{\varepsilon_n}$, then D_μ can be expressed as a determinant

$$D_\mu = \det(z_i^{b_j})|_{i,j=1}^n. \tag{355}$$

In particular, D_ρ is the Vandermonde determinant $\prod_{i<j}(z_i - z_j)$. So here the Weyl character formula tells us that the character of $L(\lambda)$ is a ratio of two determinants. These ratios define the Schur polynomials.

Definition D_ρ is the *Weyl denominator.*

The following generalizes our observation above about the Vandermonde determinant.

Theorem 13.8. *(Weyl denominator formula) We have*

$$D_\rho = \prod_{\alpha \in \Delta^+} (e^{\alpha/2} - e^{-\alpha/2}) \tag{356}$$

$$= e^\rho \prod_{\alpha \in \Delta^+} (1 - e^{-\alpha}). \tag{357}$$

Proof. Let $R = \prod_{\alpha \in \Delta^+} (e^{\alpha/2} - e^{-\alpha/2})$. If r_i is a simple reflection then it permutes the positive roots not equal to α_i and sends α_i to its negative. Hence $r_i R = -R$, from which it follows that R is W-skew-invariant. Hence, after making sure that the constants match, we have

$$R = D_\rho + \sum_{\mu \in P^{++}, \mu < \rho} c_\mu D_\mu. \tag{358}$$

But in fact if $\mu < \rho$ then, writing $\mu = \sum a_i \omega_i$, at least one a_i is less than or equal to 0, so $\mu \notin P^{++}$. The conclusion follows. □

We now turn to the proof of the Weyl character formula itself.

Proof. We first compute the characters of Verma modules. We know, from the structure of $U(\mathfrak{n}^-)$, that these have the form

$$\begin{aligned}
\mathrm{ch}(M(\lambda)) &= \sum (\dim M(\lambda)_\mu) e^\mu & (359)\\
&= e^\lambda \prod_{\alpha \in \Delta^+} \left(1 + e^{-\alpha} + e^{-2\alpha} + ...\right) & (360)\\
&= e^\lambda \prod_{\alpha \in \Delta^+} \frac{1}{1 - e^{-\alpha}} & (361)\\
&= \frac{e^{\lambda + \rho}}{D_\rho} & (362)
\end{aligned}$$

where this character lies, not in R, but in a completion $\tilde{R}$ where we allow formal power series in the negative direction. More formally, we want to allow formal sums $\sum_{\mu \in P} a_\mu e^\mu$ where there exist $\lambda_1, ..., \lambda_k$ such that if $a_\mu \neq 0$ then $\mu \le \lambda_i$ for some i. We now see that the Weyl character formula expresses $\mathrm{ch}(L(\lambda))$ as an alternating sum of characters of Verma modules

$$\mathrm{ch}(L(\lambda)) = \sum_{w \in W} \mathrm{sgn}(w) M(w(\lambda + \rho) - \rho). \tag{363}$$

Note that some of these Verma modules are of the form $M(\lambda)$ for $\lambda \in P \setminus P^+$. We showed that even these Verma modules have unique simple quotients $L(\lambda)$. It turns out that the Casimir $\Omega \in Z(U(\mathfrak{g}))$ acts by scalars on all Verma modules. To show this we will compute the Casimir with respect to a very specific basis, namely an orthonormal basis u_i of $\mathfrak{h}$ and the elements X_α, Y_α for $\alpha \in \Delta^+$. The dual basis is u_i again and the elements Y_α, X_α, hence the Casimir is

$$\begin{aligned}
\Omega &= \sum_{\alpha \in \Delta^+} (X_\alpha Y_\alpha + Y_\alpha X_\alpha) + \sum_{i=1}^n u_i^2 & (364)\\
&= \sum_{\alpha \in \Delta^+} (2 Y_\alpha X_\alpha + H_\alpha) + \sum_{i=1}^n u_i^2. & (365)
\end{aligned}$$

If v_λ is a highest weight vector generating $M(\lambda)$ then, since $X_\alpha v_\lambda = 0$ for all $\alpha \in \Delta^+$, we compute that

$$\begin{aligned}
\Omega v_\lambda &= \sum_{\alpha\in\Delta^+} H_\alpha v_\lambda + \sum_{i=1}^{n} u_i^2 v_\lambda && (366)\\
&= \left(\sum_{\alpha\in\Delta^+} \lambda(H_\alpha) + \sum \lambda(u_i)^2\right) v_\lambda && (367)\\
&= (\lambda + 2\rho, \lambda) v_\lambda. && (368)
\end{aligned}$$

This constant $f(\lambda)$ is also conveniently written $(\lambda+\rho, \lambda+\rho) - (\rho, \rho)$.

For a given λ define

$$O_\lambda = \{\mu \in P | f(\lambda) = f(\mu)\}. \tag{369}$$

This set is finite since it is the intersection of a sphere and a lattice.

Recall that a module is said to have *finite length* if it admits a finite filtration with simple quotients.

Lemma 13.9. *Any Verma module has finite length.*

Proof. Let $N \subset M(\lambda)$ be a submodule. The set of weights $P(N)$ contains a maximal weight μ. Let N' be the sum of all submodules of N not containing a vector of weight μ. Then we have a short exact sequence

$$0 \to N' \to N \to L(\mu) \to 0. \tag{370}$$

But since the Casimir Ω acts by $f(\lambda)$ on $M(\lambda)$ we know that $\mu \in O_\lambda$, hence only finitely many simple modules can occur this way. □

Corollary 13.10. *The character of a Verma module has the form*

$$ch(M(\lambda)) = ch(L(\lambda)) + \sum_{\mu<\lambda, \mu\in O_\lambda} c_{\lambda\mu} ch(L(\mu)) \tag{371}$$

where

$$c_{\lambda\mu} = \begin{cases} 0 \text{ if } \mu \not\leq \lambda \\ 1 \text{ if } \mu = \lambda. \end{cases} \tag{372}$$

Here we use the fact that characters are additive in short exact sequences. By inverting the matrix $c_{\lambda\mu}$ we obtain the following.

Corollary 13.11. *The character of $L(\lambda)$ has the form*

$$ch(L(\lambda)) = ch(M(\lambda)) + \sum_{\mu<\lambda,\mu\in O_\lambda} b_{\lambda\mu} ch(M(\mu)) \tag{373}$$

where $b_{\lambda\mu} \in \mathbb{Z}$.

Multiplying both sides by D_ρ then gives

$$D_\rho \text{ch}(L(\lambda)) = e^{\lambda+\rho} + \sum_{\mu<\lambda,\mu\in O_\lambda} b_{\lambda\mu} e^{\mu+\rho}. \tag{374}$$

Since the LHS is a product of a W-invariant and W-skew-invariant element, the RHS must be W-skew-invariant, hence

$$D_\rho \text{ch}(L(\lambda)) = D_{\lambda+\rho} + \sum_{\mu\in P^+,\mu\in O_\lambda,\mu<\lambda} b'_{\lambda\mu} D_{\mu+\rho} \tag{375}$$

where $b'_{\lambda\mu} = \text{sgn}(w) b_{\lambda(w(\mu+\rho)-\rho)}$, but we don't need this because there is no μ such that $\mu \in P^+, \mu < \lambda$, and $(\mu+\rho, \mu+\rho) = (\lambda+\rho, \lambda+\rho)$. To see this, we subtract, giving

$$(\lambda + \mu + 2\rho, \lambda - \mu) = 0. \tag{376}$$

On the other hand, $\lambda - \mu = \sum n_i \alpha_i$ where $n_i \in \mathbb{Z}_{\geq 0}$ and at least one $n_i > 0$. So $\lambda - \mu \in P^+$. We also have $\lambda + \mu + 2\rho \in P^{++}$, which contradicts the above. Hence

$$\text{ch}(L(\lambda)) = \frac{D_{\lambda+\rho}}{D_\rho} \tag{377}$$

as desired. □

A more conceptual proof of the Weyl character formula proceeds via the BGG resolution. When $\lambda \in P^+$ this is a resolution of $L(\lambda)$ via sums of Verma modules. Writing $w \cdot \lambda = w(\lambda+\rho) - \rho$, this resolution takes the form

$$\cdots \to \bigoplus_{\lambda(w)=k} M(w\cdot\lambda) \to \cdots \to \bigoplus_{i=1}^{n} M(r_i \cdot \lambda) \to M(\lambda) \to L(\lambda) \to 0 \tag{378}$$

and writing down the alternating sum of its characters gives the Weyl character formula.

We can try to use the Weyl character formula to compute the dimension of $L(\lambda)$, but the numerator and denominator of the Weyl character formula both vanish to high order at $h = 0$ (where $h \in \mathfrak{h}$). We need to take a limit. It is convenient to take $h = \rho^\vee t$ and let $t \to 0$, which gives

$$
\begin{aligned}
\dim L(\lambda) &= \lim_{t\to 0} \frac{\sum_{w\in W} \operatorname{sgn}(w) e^{(w(\lambda+\rho),\rho t)}}{\sum_{w\in W} \operatorname{sgn}(w) e^{(w(\rho),\rho t)}} && (379)\\
&= \lim_{t\to 0} \frac{\sum_{w\in W} \operatorname{sgn}(w) e^{((\lambda+\rho)t, w\rho)}}{\sum_{w\in W} \operatorname{sgn}(w) e^{(w(\rho),\rho t)}} && (380)\\
&= \lim_{t\to 0} \frac{\prod_{\alpha\in\Delta^+} e^{(\alpha/2,\lambda+\rho)t} - e^{-(\alpha/2,\lambda+\rho)t}}{\prod_{\alpha\in\Delta^+} e^{(\alpha/2,\rho)t} - e^{-(\alpha/2,\rho)t}} && (381)
\end{aligned}
$$

by the Weyl denominator formula. This gives the Weyl dimension formula

$$
\dim L(\lambda) = \frac{\prod_{\alpha\in\Delta^+} (\lambda+\rho,\alpha)}{\prod_{\alpha\in\Delta^+} (\rho,\alpha)}. \tag{382}
$$

14 Compact groups

Some of the following material can be found in a book of Helgason.

Theorem 14.1. *Let K be a connected compact Lie groups. Then there is a biinvariant (left- and right- invariant) volume form on K which is unique if we impose the condition that the total volume $\int_K dg$ is equal to 1.*

Biinvariance is equivalent to the condition that

$$
\int_K \varphi(gh)\, dg = \int_K \varphi(hg)\, dg = \int_K \varphi(g)\, dg \tag{383}
$$

for any smooth function $\varphi : K \to \mathbb{R}$.

Proof. Let $\omega_e \in \Lambda^{\dim \mathfrak{k}}(\mathfrak{k}^*)$ where $\mathfrak{k}$ is the Lie algebra of K. By left or right translating ω_e we obtain left and right invariant forms $\omega_g^r, \omega_g^\ell$ on K (this does not require compactness), and the claim is that these coincide (this requires compactness). This is equivalent to the claim that the adjoint action on ω_e is trivial, but its image would be a closed connected subgroup of $\mathbb{R}^+$, which must be trivial. □

This gives the existence of Haar measure for compact Lie groups, which lets us take averages and makes the representation theory similar to the representation theory of finite groups. Below K is a connected compact Lie group, although connectedness is often unnecessary.

Theorem 14.2. *Let V be a finite-dimensional representation of K (over $\mathbb{R}$ or $\mathbb{C}$). Then V admits a K-invariant inner product (over $\mathbb{R}$ or $\mathbb{C}$).*

Proof. Begin with any inner product $\langle -, - \rangle$ on V and consider the averaged inner product

$$
\langle v, w\rangle_K = \int_K \langle gv, gw\rangle\, dg. \tag{384}
$$

□

Corollary 14.3. *Any finite-dimensional representation of K is completely reducible.*

Proof. If V is such a representation, we can first give it a K-invariant inner product and then take orthogonal complements of subrepresentations. □

Corollary 14.4. *The Lie algebra $\mathfrak{k}$ of K is reductive.*

Proof. By the above, the adjoint representation is completely reducible. □

The abelian compact groups are precisely the tori $\mathbb{R}^n/\Gamma$ where Γ is a lattice in $\mathbb{R}^n$ of rank n.

Corollary 14.5. *If K is simply connected, then $\mathfrak{k}$ is semisimple; moreover, the Killing form B on $\mathfrak{k}$ is negative definite.*

Proof. Since $\mathfrak{k}$ is reductive we can write $\mathfrak{k} \cong \mathfrak{k}' \oplus Z(\mathfrak{k})$. If K' is the simply connected Lie group with Lie algebra $\mathfrak{k}'$, then K must be a quotient of $K' \times \mathbb{R}^n$ ($n = \dim Z(\mathfrak{k})$) by a discrete subgroup Γ of the center. Γ must have the property that $\Gamma \cap \mathbb{R}^n$ has full rank or else the quotient will fail to be compact. But Γ is the fundamental group of K, which we assumed was trivial, hence $n = 0$ and $\mathfrak{k}$ is semisimple.

The adjoint representation of K lands in $\mathrm{SO}(\mathfrak{k})$ where we have chosen some invariant inner product on $\mathfrak{k}$. Hence if $X \in \mathfrak{k}$ it follows that $\mathrm{ad}(X)$ is skew-symmetric, hence has purely imaginary eigenvalues. Hence

$$B(X, X) = \mathrm{tr}(\mathrm{ad}(X)^2) < 0 \tag{385}$$

if $X \neq 0$ as desired. □

Definition A Lie algebra $\mathfrak{k}$ is *compact* if the Killing form B is negative definite.

A compact Lie algebra is in particular semisimple. Moreover, if K is a Lie group with compact Lie algebra, then $\mathrm{Ad}(K)$ is compact. So now we want to classify compact Lie algebras.

14.1 Real forms

Let $\mathfrak{k}$ be a real Lie algebra and let $\mathfrak{g} = \mathfrak{k} \otimes \mathbb{C}$ be its complexification. If $\mathfrak{k}$ is semisimple then so is $\mathfrak{g}$, so we can try to classify the former using the fact that we have classified the latter.

Definition If $\mathfrak{g} \cong \mathfrak{k} \otimes \mathbb{C}$ then $\mathfrak{k}$ is a *real form* of $\mathfrak{g}$.

Example There are two real Lie algebras with complexification $\mathfrak{sl}_2(\mathbb{C})$, namely $\mathfrak{sl}_2(\mathbb{R})$ and $\mathfrak{su}(2)$. The latter is compact but the former is not.

The general way to understand forms is via Galois cohomology, but in this case since $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$ is small we can be more explicit. If $\mathfrak{k}$ is a real form of $\mathfrak{g}$ then we can write

$$\mathfrak{g} \cong \mathfrak{k} \oplus i\mathfrak{k}. \tag{386}$$

Complex conjugation defines a map $\theta : \mathfrak{g} \to \mathfrak{g}$ which is skew-linear, squares to the identity, and is an endomorphism of $\mathfrak{g}$; call such a thing an *anti-involution*. The real form $\mathfrak{k}$ can be recovered as the fixed points of θ. Conversely, if θ is an anti-involution, then its fixed points are a real form of $\mathfrak{g}$. Hence if we find all anti-involutions we will find all real forms. Moreover, if θ is an anti-involution then $\varphi\theta\varphi^{-1}$ is also an anti-involution where $\varphi \in \mathrm{Aut}(\mathfrak{g})$, and the corresponding real forms are isomorphic. Hence we really want to classify anti-involutions up to conjugacy.

Example Consider $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$ again. The anti-involution giving by $\theta(H, X, Y) = H, X, Y$ has fixed subalgebra $\mathfrak{sl}_2(\mathbb{R})$, the *split form* of $\mathfrak{sl}_2(\mathbb{C})$. (In general the split form is defined by the anti-involution fixing the Chevalley basis.) But we can also consider

$$\sigma(H, X, Y) = -H, -Y, -X \tag{387}$$

and then the fixed subalgebra is $\mathfrak{su}(2)$.

If B is the Killing form and θ an anti-involution, we can define a Hermitian form

$$B_\theta(X, Y) = B(\theta X, Y) \tag{388}$$

and the Killing form on $\mathfrak{g}^\theta$ is negative definite (which is equivalent to $\mathfrak{g}^\theta$ compact) iff B_θ is negative definite.

Definition The *Cartan involution* is the anti-involution on a semisimple Lie algebra $\mathfrak{g}$ given on Chevalley generators by

$$\sigma(H_i) = -H_i, \sigma(X_i) = -Y_i, \sigma(Y_i) = -X_i. \tag{389}$$

Proposition 14.6. *$\mathfrak{g}^\sigma$ is a compact Lie algebra.*

Proof. It suffices to check that B_σ is negative definite. This is clear on the subspace generated by the Chevalley generators. If $s_i = \exp X_i \exp(-Y_i) \exp X_i$ is a simple reflection then $\sigma(s_i) = s_i$, so we conclude that

$$\sigma(H_\alpha) = -H_\alpha, \sigma(X_\alpha) = -Y_\alpha, \sigma(Y_\alpha) = -X_\alpha. \tag{390}$$

The spaces $\mathbb{C}X_\alpha \oplus \mathbb{C}Y_\alpha$ are orthogonal with respect to B_σ, so to check negative definiteness it suffices to check on these subspaces, and the conclusion follows. □

Theorem 14.7. *Any compact real form of $\mathfrak{g}$ is conjugate to $\mathfrak{g}^\sigma$.*

Lemma 14.8. *Let σ be the Cartan involution and let θ be some other anti-involution. Then there exists $\varphi \in \text{Aut}(\mathfrak{g})$ such that $\varphi\theta\varphi^{-1}$ and σ commute.*

Proof. Let $\psi = \theta\sigma$. Then ψ is an automorphism with inverse $\sigma\theta$. We compute that

$$B_\sigma(\psi X, Y) = B(\sigma\theta\sigma X, Y) = B(\sigma X, \theta\sigma Y) = B_\sigma(X, \psi Y) \tag{391}$$

hence that ψ is self-adjoint. By the spectral theorem it follows that ψ is diagonalizable with real eigenvalues. If $\zeta = \psi^2$, then ζ is diagonalizable with positive eigenvalues, so ζ^t is well-defined for any real t. In fact $\zeta^t = \exp Xt$ where $X \in \mathfrak{p} \cong i\mathfrak{k}$ (where $\mathfrak{k} = \mathfrak{g}^\sigma$).

Let $\varphi = \zeta^{1/4}$. Then $\sigma\varphi = \varphi^{-1}\sigma$ by the above, hence

$$\sigma\varphi\theta\varphi^{-1} = \varphi^{-1}\sigma\theta\varphi^{-1} = \sigma\theta\varphi^{-2} \tag{392}$$

and

$$\varphi\theta\varphi^{-1}\sigma = \varphi\theta\sigma\varphi = \varphi^2\theta\sigma \tag{393}$$

but $(\sigma\theta\varphi^{-2})^2 = 1$, hence it is equal to its inverse and the conclusion follows. □

Now we prove the theorem.

Proof. Let θ be an anti-involution such that $\mathfrak{g}^\theta$ is compact. By the above lemma, we may assume WLOG that θ and σ commute, so $\sigma\theta = \theta\sigma$. In particular, θ must preserve $\mathfrak{k} = \mathfrak{g}^\sigma$. Hence

$$\mathfrak{g}^\theta = \mathfrak{k}' \oplus \mathfrak{p}' \tag{394}$$

where $\mathfrak{k}' = \mathfrak{g}^\theta \cap \mathfrak{k}, \mathfrak{p}' = \mathfrak{g}^\theta \cap \mathfrak{p}$. Since the Killing form is negative definite on $\mathfrak{g}^\theta$ but positive definite on $\mathfrak{p} = i\mathfrak{k}$, we must have $\mathfrak{p}' = 0$, hence $\mathfrak{k}' = \mathfrak{k}$. □

Let G be the simply connected complex Lie group with Lie algebra $\mathfrak{g}$. We can lift the Cartan involution σ to an involution of G. Let $K = G^\sigma$.

Proposition 14.9. *K is compact and simply connected.*

Proof. It's clear that the Lie algebra of K is $\mathfrak{g}^\sigma = \mathfrak{k}$. Consider the polar decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ and define $P = \exp(\mathfrak{p})$. Then we claim that K is compact, that P is diffeomorphic to $\mathfrak{p}$, and that the induced multiplication map $K \times P \to G$ is a diffeomorphism. □