Instituto Superior Técnico
Departamento de Matemática

LECTURE NOTES ON LIE GROUPS AND LIE ALGEBRAS

Contents

0. Introduction

These are lecture notes for a graduate course on Lie Groups and Lie Algebras taught at IST Lisbon in the Fall semester of 2017/2018 and again in 2018/2019. It is assumed that the reader is familiar with basic Differential Geometry (vector fields, differential forms, immersions and the Frobenius theorem in particular), basic point set topology including the fundamental group and covering spaces as well as basic algebra (linear algebra, tensor product, exact sequences). Basic algebraic and differential topology will also be invoked at very isolated instances.

Date: January 25, 2019.

1. Basic definitions

Definition 1.1. *A* **Lie group** *is a smooth manifold*[1] G *together with an element* $e \in G$ *and a multiplication map* $\mu\colon G \times G \to G$ *which has* e *as a unit, is associative and has inverses. Moreover the map* μ *and the inverse map* $\iota\colon G \to G$ *are required to be smooth.*

These objects are used to describe continuous symmetries. They were first studied (in a local form) by Sophus Lie c. 1870 with a view to exploiting symmetries for the solution of differential equations. For more on this see [Ol].

Exercise 1.2. *Show that the requirement that* ι *be smooth can be omitted from the definition. That is, if* μ *is smooth and gives* G *a group structure then the inverse map is automatically smooth.*

Example 1.3. *(i) Any countable (discrete) group is a Lie group.*

(ii) $(\mathbb{R}^n, +)$ *is an abelian Lie group.*

(iii) The general linear groups $\mathrm{GL}(n;\mathbb{R})$ *(resp.* $\mathrm{GL}(n;\mathbb{C})$*) of invertible* $n \times n$ *matrices with real (resp. complex) coefficients are Lie groups. Note that these sets of matrices are open sets in the Euclidean spaces of all square matrices and hence have natural manifold structures. Moreover the usual formulas for matrix multiplication and inversion show that these operations are smooth. These groups are not abelian unless* $n = 1$.

(iv) The connected (or path) component of the identity in a Lie Group is again a Lie group (since the components are open sets). Applying this to the previous example we obtain the Lie group $\mathrm{GL}(n;\mathbb{R})^+ = \{A \in \mathrm{GL}(n;\mathbb{R})\colon \det A > 0\}$.

(v) The orthogonal groups

$$O(n) = \{A \in \mathrm{GL}(n;\mathbb{R})\colon A^tA = \mathrm{Id}\}$$

are Lie groups. To check that $O(n)$ *is a submanifold of* $\mathrm{GL}(n;\mathbb{R}) \subset \mathbb{R}^{n^2}$, *let* $\mathrm{Sym}_n(\mathbb{R})$ *denote the vector space of symmetric* $n \times n$ *matrices and consider the quadratic map*

$$\Phi\colon M_{n\times n}(\mathbb{R}) \to \mathrm{Sym}_n(\mathbb{R}) \cong \mathbb{R}^{\frac{n(n+1)}{2}}$$

defined by the expression $\Phi(A) = A^tA$, *so that* $O(n) = \Phi^{-1}(\mathrm{Id})$. *In order to check that* $O(n)$ *is a submanifold of* $\mathrm{GL}(n;\mathbb{R})$ *(or equivalently of* $\mathbb{R}^{n^2}$*) we must check that* Id *is a regular value of* Φ. *Since*

$$\Phi(A+H) = (A+H)^t(A+H) = \Phi(A) + A^tH + H^tA + H^tH$$

we see that

$$D\Phi(A)H = A^tH + H^tA$$

Since $A \in \mathrm{GL}(n;\mathbb{R})$ *is invertible, the matrix* AH *is arbitrary and hence so is its symmetrization* $D\Phi(A)H$. *We conclude that* $D\Phi(A)$ *is surjective for each* $A \in \mathrm{GL}(n;\mathbb{R})$ *and hence* $O(n)$ *is a submanifold of dimension* $\frac{n(n-1)}{2}$. *Since the multiplication and inverse maps on* $O(n)$ *are restrictions of smooth maps they are smooth and so* $O(n)$ *is a Lie group.*

[1] Manifolds are assumed to be Hausdorff and second countable.

One can check that $O(n)$ has exactly two components. The connected component of the identity consists of the rotations and is called the special orthogonal group $SO(n) = O(n) \cap GL^+(n)$. *When $n = 2$ we also write S^1 instead of $SO(2)$ since*

$$SO(2) = \left\{ \left[\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right] : \theta \in \mathbb{R} \right\}$$

is homeomorphic to the circle.

(vi) *The* unitary groups

$$U(n) = \{A \in \mathrm{GL}(n; \mathbb{C}) \colon A^*A = \mathrm{Id}\}$$

are again Lie groups of dimension n^2. The special unitary groups

$$SU(n) = \{A \in U(n) \colon \det(A) = 1\}$$

are also Lie groups. Note that $SO(2) \cong U(1)$.

(vii) *Cartesian products of Lie groups are again Lie groups. In particular the n-fold product $S^1 \times \cdots \times S^1$ is a Lie group called the n-*torus.

The groups of symmetries of many types of mathematical objects turn out to be Lie groups. For instance this is the case for the symmetries of any kind of algebraic structure on a real vector space (as we will soon see) and also for the groups of symmetries of Riemannian and complex manifolds [Ko].

Exercise 1.4. *(i) Check that $GL^+(n; \mathbb{R})$ and $SO(n)$ are connected.* Hint: Use Gauss elimination and polar decomposition.

(ii) Prove the statements about $U(n)$ and $SU(n)$ in the example above.

The other basic object of study in this course is of a purely algebraic nature.

Definition 1.5. *A* Lie algebra *over a field $\mathbb{K}$ is a vector space V over $\mathbb{K}$ together with a bilinear map $[\ ,\ ] \colon V \times V \to V$ (called the* Lie bracket*) satisfying*

- *Anti-symmetry: $[v, w] = -[w, v]$ for all $v, w \in V$*
- *Jacobi identity: $[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0$ for all $u, v, w \in V$.*

Example 1.6. *(i) If V is any vector space we can set the bracket to be the 0 map. This Lie algebras are called* abelian.

(ii) If A is an associative algebra over the field $\mathbb{K}$ and we set $[\ ,\ ]$ to be the commutator bracket

$$[a, b] = ab - ba$$

then it is immediate to check that $[\ ,\ ]$ gives A the structure of a Lie algebra. When $A = M_n(\mathbb{K})$ (or, more invariantly, $\mathrm{End}(V)$ *the algebra of endomorphisms of V) this lie algebra is denoted $\mathfrak{gl}(n, \mathbb{K})$ (or $\mathfrak{gl}(V)$).*

(iii) Suppose A is an algebra over $\mathbb{K}$, not necessarily associative or even unital (so A just has a bilinear product $A \times A \to A$ denoted by juxtaposition). A derivation *of A is a linear map $d \colon A \to A$ satisfying the Leibniz rule*

$$d(ab) = (da)b + a(db)$$

The set $\mathrm{Der}(A)$ *of derivations is clearly a linear subspace of* $\mathrm{End}(A)$*. It is closed under the commutator bracket in* $\mathrm{End}(A)$*:*

$$\begin{aligned}
[d_1, d_2](ab) &= d_1(d_2(a)b + ad_2(b)) - d_2(d_1(a)b + ad_1(b)) \\
&= (d_1d_2)(a)b + d_2(a)d_1(b) + d_1(a)d_2(b) + a(d_1d_2)(b) - \\
&\qquad (d_2d_1)(a)b - d_1(a)d_2(b) - d_2(a)d_1(b) - a(d_2d_1)(b) \\
&= [d_1, d_2](a)b + a[d_1, d_2](b)
\end{aligned}$$

so the commutator bracket gives $\mathrm{Der}(A)$ *the structure of a Lie algebra*

(iv) If M *is a manifold, the vector space* $\mathcal{X}(M)$ *of vector fields on* M *together with the Lie bracket of vector fields is a Lie algebra.*

Exercise 1.7. *Show that* $\mathcal{X}(M)$ *is the algebra of derivations of the* $\mathbb{R}$*-algebra* $C^\infty(M)$ *so example (iv) above is actually a special case of (iii).*

Remark 1.8. *The Jacobi identity may be interpreted as saying that a Lie algebra* L *acts on itself via derivations. Namely a vector space* L *together with an anti-symmetric (bilinear) product is a Lie algebra iff for each* x *in* L*, the endomorphism of* L *given by* $y \mapsto [x, y]$ *is a derivation.*

We will now explain the basic connection between the concepts of Lie group and Lie algebra. Given a Lie group G and an element $g \in G$ we write

$$L_g \colon G \to G \qquad R_g \colon G \to G$$

for the maps given by left and right multiplication by g respectively. So $L_g(h) = gh$ and $R_g(h) = hg$. Clearly L_g is a diffeomorphism (with inverse $L_{g^{-1}}$) and $g \mapsto L_g$ gives a left, simply transitive, action of G on itself by diffeomorphisms. Similarly $g \to R_g$ gives a right action.

Definition 1.9. *Let* G *be a Lie group. A vector field* $X \in \mathcal{X}(G)$ *is* left invariant *(resp.* right invariant*) if* $dL_g(X_h) = X_{gh}$ *(resp.* $dR_g(X_h) = X_{hg}$*) for all* $g, h \in G$*.*

In order to check whether a vector field is left invariant, it is sufficient to check the special case $X_g = dLg(X_e)$ as the general case then follows from the chain rule. From now on we will stick to left invariant vector fields and stop pointing out the analogous results for right invariant vector fields.

Lemma 1.10. *Let* G *be a Lie group.*

(a) A (not necessarily continuous) left invariant vector field on G *is smooth and complete.*
(b) If $X, Y \in \mathcal{X}(G)$ *are left invariant so is* $[X, Y]$*.*
(c) The left invariant vector fields on G *are closed under the Lie bracket of vector fields.*
(d) The left invariant vector fields form a vector subspace of $\mathcal{X}(G)$ *canonically isomorphic to* T_eG *(and hence of dimension* $\dim G$*).*

Proof. (a) Let X be a (not necessarily continuous) left invariant vector field and consider the diagram

$$\begin{tikzcd} G \arrow[r, "s"] \arrow[dr, "r"'] & TG \times TG \arrow[r, "d\mu"] \arrow[d] & TG \arrow[d] \\ & G \times G \arrow[r] & G \end{tikzcd}$$

where $s(g) = (0, X_e)$ and $r(g) = (g, e)$ (and the vertical maps are the canonical projections). Left invariance implies that $X = d\mu \circ s$ and it follows that X is smooth, as it is the composite of smooth maps.

(b) Let $\gamma \colon]a, b[\to G$ be an integral curve for the left invariant vector field X, so that

$$\frac{d\gamma}{dt} = X_{\gamma(t)}$$

Then $L_g \circ \gamma$ is also an integral curve for X. Thus, if $\gamma \colon]-2\epsilon, 2\epsilon[\to G$ is an integral curve for X with $\gamma(0) = e$, we may extend it to $]-2\epsilon, 3\epsilon[$ by setting

$$\gamma(t) = L_{\gamma(\epsilon)}\gamma(t - \epsilon), \qquad \text{for } 2\epsilon \le t < 3\epsilon$$

The expression on the right of the equality is defined and smooth for $-\epsilon < t < 3\epsilon$ and it is an integral curve taking the value $\gamma(\epsilon)$ at time ϵ. Therefore it agrees with $\gamma(t)$ in the common domain of definition $]-\epsilon, 2\epsilon[$. We thus have a smooth integral curve defined on $]-2\epsilon, 3\epsilon[$. Continuing in this way we can extend γ to the whole of $\mathbb{R}$. The remainder of the proof is left as an exercise.

(c) A vector field is left invariant iff it is L_g related to itself for each $g \in G$ (i.e. $X = dL_g X$). By the properties of the Lie bracket of vector fields, if X, Y are L_g-related to themselves, so is their Lie bracket.

(d) The evaluation at e map $\mathcal{X}(G) \to T_eG$ is linear. It is surjective because given $v \in T_eG$ we can define a left invariant vector field on G by the formula $X_g = dL_g(v)$ as we saw in (a). Evaluation is clearly injective because the value of a left invariant field at g is determined by its value at e.

□

Definition 1.11. *The space of left invariant vector fields on a Lie group G with the Lie bracket is denoted* $\mathrm{Lie}(G)$ *or* $\mathfrak{g}$ *and called the Lie algebra associated to G.*

Remark 1.12. *Traditionally the Lie algebra of a Lie group is denoted by the same letter with a german script, i.e. $\mathfrak{h}$ for the Lie algebra of the Lie group H, etc. It is also customary to use german letters from Lie algebras even when they are not being derived from any Lie group.*

Example 1.13. *(1) Let $G = \mathbb{R}^n$. Then left multiplication by v is translation by v, so the diffeomorphism L_v induces the canonical isomorphism between the tangent spaces at different points. With respect to these standard identifications, the left invariant vector fields are therefore the constant vector fields. It follows that the Lie bracket vanishes identically, i.e. $\mathfrak{g}$ is the abelian Lie algebra of dimension n.*

(2) *Consider the standard coordinates* $a, b, c, d\colon \mathrm{GL}(2, \mathbb{R}) \to \mathbb{R}$ *given by the matrix entries. Identifying the tangent spaces at all points with* $M_{2\times 2}(\mathbb{R})$ *in the usual way, let us write down the expression for the left invariant vector field* X *determined by the element* $\frac{\partial}{\partial a} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \in T_e \mathrm{GL}(2, \mathbb{R})$.

We have

$$X_{\begin{bmatrix} a & b \\ c & d \end{bmatrix}} = dL_{\begin{bmatrix} a & b \\ c & d \end{bmatrix}} \frac{\partial}{\partial a} = a\frac{\partial}{\partial a} + c\frac{\partial}{\partial c} = \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix}$$

(3) *Let* $G = \mathrm{GL}(n; \mathbb{R})$. *Then* T_eG *can be identified in the usual way with* $M_{n\times n}(\mathbb{R}) = \mathfrak{gl}(n; \mathbb{R})$. *Let's check that evaluation at the identity matrix* e *takes the Lie bracket of vector fields to the commutator bracket of matrices and thus identifies the left invariant vector fields in* $\mathcal{X}(GL(n; \mathbb{R}))$ *with* $\mathfrak{gl}(n; \mathbb{R})$.

Given left invariant vector fields X *and* Y *on* G, *we need to check that* $[X, Y]_e = [X_e, Y_e]$. *Consider the usual coordinates* $x_{ij}\colon \mathrm{GL}(n, \mathbb{R}) \to \mathbb{R}$ *which compute the various entries of the matrix. We can obtain the* ij*-th component of a vector field* X *at* e *by computing the value of the function* $X \cdot x_{ij}$ *at the point* e. *Now*

$$([X, Y] \cdot x_{ij})_e = X_e \cdot (Y \cdot x_{ij}) - Y_e \cdot (X \cdot x_{ij})$$

Given $g \in G$ *we have*

$$(Y \cdot x_{ij})(g) = Y_g \cdot x_{ij} = dL_g(Y_e) \cdot x_{ij} = Y_e \cdot (x_{ij} \circ L_g)$$

and

$$(x_{ij} \circ L_g)(h) = (gh)_{ij} = \sum_{k=1}^{n} g_{ik}h_{kj} = \sum_{k=1}^{n} x_{ik}(g)x_{kj}(h)$$

hence

$$(Y \cdot x_{ij})(g) = \sum_{k=1}^{n} x_{ik}(g)(Y_e \cdot x_{kj})$$

and

$$X_e \cdot (Y \cdot x_{ij}) = \sum_{k=1}^{n} (X_e \cdot x_{ik})(Y_e \cdot x_{kj})$$

Subtracting the above expression with X *and* Y *switched we obtain that* $([X, Y]\cdot x_{ij})_e$ *is given by the* ij*-th entry of the commutator of the matrices* X_e *and* Y_e *as required.*

Definition 1.14. *Let* G, H *be Lie groups. A map* $\phi\colon G \to H$ *is a* Lie group homomorphism *if it is a group homomorphism and a smooth map. Given Lie algebras* $\mathfrak{g}, \mathfrak{h}$, *a linear map* $f\colon \mathfrak{g} \to \mathfrak{h}$ *is a* Lie algebra homomorphism *if for all* $v, w \in \mathfrak{g}$ *we have* $[f(v), f(w)] = f([v, w])$. *Given a Lie group homomorphism* $\phi\colon G \to H$ *we define*

$$\phi_*\colon \mathfrak{g} \to \mathfrak{h}$$

by stipulating that $\phi_*(X)$ *is the unique left invariant vector field on* H *whose value at* $e \in H$ *is* $d\phi_e(X_e)$.

Note that under the identification of the Lie algebras associated with Lie groups with their tangent spaces at the identity, ϕ_* is just the derivative of ϕ at the identity, $d\phi_e$.

Proposition 1.15. *If ϕ is a Lie group homomorphism then ϕ_* is a Lie algebra homomorphism (meaning a map of vector spaces which preserves the Lie bracket).*

Proof. Let's first check that $\phi_*(X)$ is ϕ-related to X:

$$d\phi_g(X_g) = d\phi_g(dL_g)_e(X_e) = (dL_{\phi(g)})_e(d\phi_e)X_e = (dL_{\phi(g)})_e(\phi_*(X))_e = (\phi_*(X))_{\phi(g)}$$

where in the second equality we used $\phi \circ L_g = L_{\phi(g)} \circ \phi$ (which holds because ϕ is a homomorphism). Since the property of being ϕ-related is closed under the Lie bracket, it follows that $[\phi_*(X), \phi_*(Y)]$ is ϕ-related to $[X, Y]$. In particular $[\phi_*(X), \phi_*(Y)]_e = d\phi([X, Y]_e)$. As $[\phi_*(X), \phi_*(Y)]$ is left invariant, we conclude that $[\phi_*(X), \phi_*(Y)] = \phi_*([X, Y])$. □

Example 1.16.
(1) *The map $\phi\colon \mathbb{R}^n \to (S^1)^n$ defined by*

$$\phi(x_1, \dots, x_n) = (e^{2\pi i x_1}, \dots, e^{2\pi i x_n})$$

is clearly a Lie group homomorphism and ϕ_* is an isomorphism between the two (abelian) Lie algebras.

(2) *Given $A \in M_{n\times n}(\mathbb{R})$ the map $\phi\colon \mathbb{R} \to \mathrm{GL}(n; \mathbb{R})$ defined by*

$$A \mapsto e^{tA}$$

is a Lie group homomorphism. In general, Lie group homomorphisms $\mathbb{R} \to G$ are called 1-parameter subgroups *(even if they are not injective and hence not necessarily subgroups). The map induced on the level of Lie algebras is determined by $\phi_*(1) = A$.*

(3) $\det\colon \mathrm{GL}(n; \mathbb{R}) \to \mathrm{GL}(1; \mathbb{R})$ *is a Lie group homomorphism. The map induced on the level of Lie algebras is the trace map* $\mathrm{tr}\colon M_{n\times n}(\mathbb{R}) \to \mathbb{R}$. *The kernel of* $\det$ *is also a Lie group (exercise) called the* special linear group $\mathrm{SL}(n, \mathbb{R})$.

Definition 1.17. *Lie group homomorphisms $\phi\colon G \to GL(n; \mathbb{R})$ are called (real)* representations *of G. They correspond to linear actions $G \times \mathbb{R}^n \to \mathbb{R}^n$. Similarly a Lie algebra homomorphism $\phi\colon \mathfrak{g} \to \mathfrak{gl}(n; \mathbb{R})$ is called a real representation of $\mathfrak{g}$. Replacing $\mathbb{R}$ with $\mathbb{C}$ we obtain the notion of complex representations.*

If a representation is injective one says that the representation is *faithful*. In that case we can think of the elements of G (or $\mathfrak{g}$) as matrices by identifying the elements of G (or $\mathfrak{g}$) with their images under ϕ. We'll soon see this is not always possible for Lie groups.

At this point we have constructed a functor

$$D\colon \text{ Lie groups } \to \text{Lie algebras}$$

which sends a Lie group to its associated Lie algebra and a Lie group homomorphism ϕ to ϕ_*. Under the identification of the Lie algebra with the tangent space at the identity, D is nothing other than the functor "tangent at e" which justifies the notation - D for derivative. One of Lie's great achievements was his understanding that the functor D comes pretty close to being an equivalence, so that the theory of Lie groups, which are fairly complicated objects, can to a large extent be reduced to the linear algebraic problem of understanding Lie algebras and maps between them. In particular, the infinitesimal information encoded in the Lie algebra (we'll see later how the Lie bracket is related to the product of the Lie group) determines the Lie group G to a large extent. Of course, it doesn't completely determine G as D is clearly blind to the existence of non-trivial connected components and, for instance, the homomorphism ϕ in Example 1.16(1) is not an isomorphism even though ϕ_* is. Still it comes pretty close as we will see. Our next objective is to understand the basic properties of this correspondence between Lie groups and Lie algebras.

2. Lie's Theorems

Definition 2.1. *A* Lie subalgebra *of a Lie algebra* $\mathfrak{g}$ *is a vector subspace* $\mathfrak{h} \subset \mathfrak{g}$ *which is closed under the Lie bracket of* $\mathfrak{g}$. *A* Lie subgroup *of a Lie group* G *is a Lie group homomorphism* $\phi\colon H \to G$ *which is an injective immersion.*

Example 2.2. *(1)* $\phi\colon \mathbb{R} \to S^1 \times S^1$ *defined by* $\phi(t) = (e^{it}, e^{i\sqrt{2}t})$ *is a Lie subgroup but not an embedding.*

(2) Since $\operatorname{tr}(XY) = \operatorname{tr}(YX)$ *we see that*

$$\mathfrak{sl}(n;\mathbb{R}) = \{X \in \mathfrak{gl}(n;\mathbb{R})\colon \operatorname{tr} X = 0\}$$

is a Lie subalgebra of $\mathfrak{gl}(n;\mathbb{R})$ *(and similarly for complex, or indeed any other field, coefficients). These Lie algebras are called the* special linear Lie algebras*. They are the Lie algebras of the Lie groups appearing in Exameplo*

Clearly if $\phi\colon G \to H$ is a Lie subgroup, then $\phi_*\colon \mathfrak{g} \to \mathfrak{h}$ identifies $\mathfrak{g}$ with the subalgebra $\phi_*(\mathfrak{g}) \subset \mathfrak{h}$. Conversely we have the following basic result that gives a one-to-one correspondence between connected Lie subgroups of a Lie group and Lie subalgebras of its Lie algebra.

Theorem 2.3. *Let* G *be a Lie group and* $\mathfrak{k} \subset \mathfrak{g}$ *be a Lie subalgebra. Then there exists a unique connected Lie subgroup* $\phi\colon H \to G$ *such that* $\phi_*(\mathfrak{h}) = \mathfrak{k}$.

Uniqueness in the previous statement means that if $\phi'\colon H' \to G$ is another connected Lie subgroup satisfying $\phi_*(\mathfrak{h}') = \mathfrak{k}$, then there exists a Lie group homomorphism $\psi\colon H \to H$, which is also a diffeomorphism (i.e. an isomorphism of Lie groups) such that the following diagram commutes

$$\begin{tikzcd} H \arrow[r, "\phi"] \arrow[d, "\psi"] & G \\ H' \arrow[ur, "\phi'"] & \end{tikzcd}$$

In order to prove the uniqueness part of the Theorem we will use the following result.

Lemma 2.4. *Let* G *be a connected topological group and* U *a neighborhood of* e. *Then* U *generates* G, *i.e.* $\cup_{n=1}^{\infty} U^n = G$ *(where* U^n *denotes the set of* n*-fold products in* G *of elements in* U*).*

Proof. Let $V = U \cap U^{-1} \subset U$. This is still a neighborhood of e because the inverse map is a homeomorphism from G to itself. The subset $S = \cup_{n=1}^{\infty} V^n \subset G$ is clearly a subgroup because $V = V^{-1}$. It is also an open subset of G because it can be written as a union of copies of V (for instance $V^2 = \cup_{g \in V} gV$). G can be decomposed into left cosets of S as

$$G = \coprod_{[h] \in G/S} hS$$

Since the left cosets hS are open, the complement of $S = eS$ in G is a union of open subsets of G and hence is open. Therefore $S \subset G$ is both open and closed. As G is connected, it follows that $S = G$. □

Proof of Theorem 2.3: Consider the distribution $D \subset TG$ defined by $D_g = dL_g(\mathfrak{k})$. This distribution is globally generated by a basis of $\mathfrak{k}$ therefore it is smooth. Since $\mathfrak{k}$ is closed under the Lie bracket, by the Frobenius Theorem the distribution D is integrable and defines a regular foliation of G (i.e. a partition of G into leaves = maximal connected integral submanifolds of D, all of dimension $\dim \mathfrak{k}$).

Let $\phi\colon H \to G$ be the leaf through e. Given any $g \in \phi(H)$, $L_{g^{-1}} \circ \phi$ also contains e in its image, and by left invariance of D, is also an integral submanifold. By maximality of ϕ we have $L_{g^{-1}}(\phi(H)) \subset \phi(H)$. Thus, given $g, g' \in \phi(H)$ we have $g^{-1}g' \in \phi(H)$ and $\phi(H)$ is a subgroup of G.

Since leaves of foliations are initial submanifolds the product map on $\phi(H)$ factors through ϕ as a smooth map $\nu\colon H \times H \to H$

$$\begin{tikzcd} H \times H \arrow[r, "\phi\times\phi"] \arrow[d, dashed, "\nu"] & G \times G \arrow[d, "\mu"] \\ H \arrow[r, "\phi"] & G \end{tikzcd}$$

Since ϕ is injective, it is clear that $\phi^{-1}(e)$ is a unit for the multiplication ν, ν is associative and has inverses. One can also see that the inverse map is smooth in the same way that we saw that ν is smooth (although this is in fact not necessary by Exercise 1.2). We conclude that with the Lie group structure on H determined by ν, ϕ is a Lie subgroup of G. By construction we have $\phi_*(\mathfrak{h}) = \mathfrak{k}$. This concludes the proof of existence.

Suppose now that $\phi'\colon H' \to G$ is another connected Lie subgroup with $d\phi'(T_eH') = \mathfrak{k}_e$. Then ϕ' is also an integral submanifold through e of the distribution $D \subset TG$. By maximality we must have $\phi'(H') \subset \phi(H)$ and since ϕ is an initial submanifold, there is a smooth map $\psi\colon H' \to H$ such that the diagram

$$\begin{tikzcd} H \arrow[r, "\phi"] & G \\ H' \arrow[u, dashed, "\psi"] \arrow[ur, "\phi'"] & \end{tikzcd}$$

commutes. As ϕ' is injective, ψ is an injective group homomorphism. As $d\phi_e$ and $d\phi'_e$ are both isomorphisms onto $\mathfrak{k}_e$, it follows that $d\psi_e$ is a bijection. By the inverse function theorem this implies that ψ is a local diffeomorphism at e and hence the image of ψ contains a neighborhood of e in H. By Lemma 2.4, ψ is surjective and hence bijective. Since ψ is a Lie group homomorphism, the fact that is a local diffeomorphism at e implies it is a local diffeomorphism at each $h \in H'$. It is therefore a diffeomorphism, which completes the proof. $\square$

We can now see that the functor D is injective on morphisms from connected Lie groups:

Corollary 2.5. *Let G be a connected Lie group and $\phi, \psi\colon G \to H$ be such that $\phi_* = \psi_*$. Then $\phi = \psi$.*

Proof. Consider the Lie subgroups $\mathrm{id} \times \phi, \mathrm{id} \times \psi\colon G \to G \times H$ given by

$$g \mapsto (g, \phi(g)), \qquad g \mapsto (g, \psi(g))$$

Both these subgroups correspond to the same Lie subalgebra $\mathrm{graph}(\phi_*) = \mathrm{graph}(\psi_*) \subset \mathfrak{g} \times \mathfrak{h}$. Since G is connected, by Theorem 2.3 there exists a Lie group automorphism $\varphi\colon G \to G$ such that $\mathrm{id} \times \phi = (\mathrm{id} \times \psi) \circ \varphi$. Composing the homomorphisms $\mathrm{id} \times \phi, \mathrm{id} \times \psi$ with $\pi_1\colon G \times H \to G$, we see that φ must be the identity Hence $\phi(g) = \psi(g)$ for all $g \in G$ as required. $\square$

The correspondence between maps and graphs used in the previous proof suggests a method for integrating maps of Lie algebras. Suppose G, H are Lie groups and $f\colon \mathfrak{g} \to \mathfrak{h}$ is a map of Lie algebras. Then f determines a Lie subalgebra of the product Lie algebra $\mathfrak{g} \times \mathfrak{h}$

$$\mathrm{graph}(f) = \{(v, f(v)) \in \mathfrak{g} \times \mathfrak{h} \colon v \in \mathfrak{g}\}$$

which by Theorem 2.3 corresponds to a connected Lie subgroup of $G \times H$. Can this be used to integrate f, i.e. to find $\psi\colon G \to H$ such that $\psi_* = f$? Let

$$\phi\colon K \to G \times H$$

be the connected Lie group integrating $\mathrm{graph}(f)$. It's easy to check that if the integrating map ψ exists, its graph in $G \times H$ gives the subgroup ϕ. The subgroup ϕ is a graph iff $\pi_1 \circ \phi\colon K \to G$ is a diffeomorphism (and hence a Lie group isomorphism). In that case

$$\pi_2 \circ \phi \circ (\pi_1 \circ \phi)^{-1}$$

is the required Lie group homomorphism integrating f. Although we do know that $\pi_1 \circ \phi$ is a local diffeomorphism, because it induces an isomorphism

$$\mathfrak{k} \xrightarrow{\cong} \mathrm{graph}(f) \xrightarrow{\pi_{1*}} \mathfrak{g}$$

it need not be a diffeomorphism as the following example shows.

Example 2.6. *Identify in the usual way the Lie algebras of $G = S^1$ and $H = \mathbb{R}$ with $\mathbb{R}$ and consider the map of Lie algebras $f = \mathrm{id}_{\mathbb{R}}$. Then $\mathrm{graph}(f) = \Delta(\mathbb{R}) \subset \mathbb{R}^2 \cong \mathfrak{g} \times \mathfrak{h}$ is a diagonal copy of $\mathbb{R}$ inside $\mathbb{R}^2$. The corresponding Lie subgroup is $\mathbb{R} \hookrightarrow S^1 \times \mathbb{R}$ defined by*

$$t \mapsto (e^{it}, t)$$

and we see that $\pi_1 \circ \phi$ is not injective.

Proposition 2.7. *Let $\phi\colon G \to H$ be a Lie group homomorphism. Then $\phi_*\colon \mathfrak{g} \to \mathfrak{h}$ is an isomorphism iff ϕ is a covering map.*

Proof. $\Leftarrow$: Suppose that $\phi_* = d\phi_e$ is not surjective. By left invariance, the rank of $d\phi_g$ is independent of $g \in G$. The constant rank theorem then guarantees that the image of ϕ is contained in the image in H of a smooth map from a manifold of dimension less than $\dim H$, and hence has measure zero. Since a covering map is an open map, ϕ can not be a covering map. On the other hand, If $d\phi_e$ is not injective then there is a non-trivial one parameter subgroup $\psi\colon \mathbb{R} \to G$ such that $d\phi_e \circ d\psi_0 = 0$. From Corollary 2.5 it follows that $\phi\psi$ is constant, so $\psi(\mathbb{R})$ is contained in the kernel of ϕ. Since ψ is an immersion we see that $\ker(\phi)$ is not a discrete subspace of G and hence ϕ is not a covering map.

$\Rightarrow$: By the inverse function theorem and left invariance, ϕ is a local diffeomorphism at all $g \in G$. The result now follows from point set topology: A homomorphism $\phi\colon G \to H$ of topological groups which is a local homeomorphism is a covering map. In order to see this, first note that $D = \ker\phi$ is a discrete normal subgroup of G. Consider the map $d\colon G \times G \to G$ defined by $d(x, y) = xy^{-1}$ and pick a neighborhood V of e in G such that

- $d(V \times V) \cap D = \{e\}$ (i.e.($x, y \in V$ and $xy^{-1} \in D) \Rightarrow x = y$)
- $\phi_{|V}$ is a homeomorphism.

We'll check that

$$\phi^{-1}(\phi(V)) = \coprod_{x\in D} xV \tag{1}$$

so that V is a trivialising neighborhood of e for ϕ. Left invariance then guarantees that ϕ is a covering map.

Clearly $\phi(xV) = \phi(V)$ and $\phi_{|xV}$ is a local homeomorphism. Suppose that $xV \cap x'V \neq \emptyset$. Then $xv = x'w$ for some $v, w \in V$ and therefore $wv^{-1} \in D$. But in that case $v = w$ and hence $x = x'$. We conclude that the open sets $\{xV : x \in D\}$ are pairwise disjoint. Finally, if $\phi(g) \in \phi(V)$ then $\phi(g) = \phi(v)$ for some $v \in V$. Therefore $gv^{-1} = x$ for some $x \in D$ and hence $g \in xV$. ☐

Corollary 2.8 (Lie's second theorem)**.** *Let G be a simply connected Lie group, H a Lie group and $f\colon \mathfrak{g} \to \mathfrak{h}$ a Lie algebra homomorphism. Then there exists a unique Lie group homomorphism $\phi\colon G \to H$ such that $\phi_* = f$.*

Proof. Let $\psi\colon K \to G \times H$ be the Lie subgroup corresponding to the Lie subalgebra $\mathrm{graph}(f) \subset \mathfrak{g} \times \mathfrak{h}$. The map $\pi_1\psi\colon K \to G$ is a covering space by Proposition 2.7. Since G is simply connected, $\pi_1\psi$ is a homeomorphism and hence a diffeomorphism. The required Lie group homomorphism is $\phi = \pi_2\psi(\pi_1\psi)^{-1}$. Uniqueness is guaranteed by Corollary 2.5. ☐

The previous result says that the Lie correspondence D is a bijection on morphisms from a simply connected source. It is important to note that every Lie group has a "simply connected version":

Proposition 2.9. *Let G be a connected Lie group. Then the universal covering space $\widetilde{G}$ admits the structure of a Lie group in such a way that $\pi\colon \widetilde{G} \to G$ a Lie group homomorphism.*

Proof. Let $\pi\colon \widetilde{G} \to G$ be a universal covering map and $\tilde{e}$ be an arbitrary point in $\pi^{-1}(e)$. Writing $\mu\colon G \times G \to G$ for the multiplication on G, the lifting theorem gives a unique map $\tilde{\mu}$ sending $(\tilde{e}, \tilde{e})$ to $\tilde{e}$.

$$\begin{tikzcd} \widetilde{G} \times \widetilde{G} \arrow[r, dashed, "\tilde{\mu}"] \arrow[d, "\pi\times\pi"] & \widetilde{G} \arrow[d, "\pi"] \\ G \times G \arrow[r, "\mu"] & G \end{tikzcd}$$

and one easily checks (using the uniqueness of lifts) that this gives $\tilde{G}$ the structure of a Lie group with unit $\tilde{e}$ (Exercise). ☐

Example 2.10.
(1) *The map $\pi\colon \mathbb{R}^n \to (S^1)^n$ defined by $\pi(x_1, \dots, x_n) = (e^{2\pi i x_1}, \dots, e^{2\pi i x_n})$ is the universal covering map.*
(2) *Let S^3 denote the Lie group of unit quaternions. Then $\pi\colon S^3 \to SO(3)$ defined by*
$$\pi(q)(v) = qv\overline{q}$$
(where $v \in \mathbb{R}^3$ is regarded as a purely imaginary quaternion) is the universal covering map. The kernel consists of $\{\pm 1\}$ from which we see that $\pi_1(SO(3)) \cong \mathbb{Z}/2$.
(3) *The map $\pi\colon S^3 \times S^3 \to SO(4)$ defined by*
$$\pi(q_1, q_2)(v) = q_1 v \overline{q}_2$$
is the universal (double) covering with kernel $\{\pm(1, 1)\}$.

Remark 2.11. *For $n \geq 3$, the fundamental group of $SO(n)$ is $\mathbb{Z}/2$. The universal double covers are called the Spin groups* $\mathrm{Spin}(n)$. *We will see an algebraic construction later (see* [BtD] *for instance).*

Example 2.12 (A non-linear Lie group)**.** *Let $\widetilde{\mathrm{SL}(2;\mathbb{R})}$ be the universal cover of the Lie group $\mathrm{SL}(2;\mathbb{R}) = \{A \in \mathrm{GL}(2;\mathbb{R})\colon \det A = 1\}$.*

You will see in the exercises that $\mathrm{SL}(2,\mathbb{R})$ is diffeomorphic to $S^1 \times \mathbb{R}^2$ so that, in particular, $\pi_1(\mathrm{SL}(2,\mathbb{R})) \cong \mathbb{Z}$. You will also see that any representation $f\colon \mathfrak{sl}(2;\mathbb{R}) \to \mathfrak{gl}(n;\mathbb{R})$ can be integrated to a representation $\mathrm{SL}(2;\mathbb{R}) \to \mathrm{GL}(n;\mathbb{R})$ (even though Corollary 2.8 does not apply).

Now suppose we have a representation
$$\widetilde{\mathrm{SL}(2;\mathbb{R})} \xrightarrow{\phi} \mathrm{GL}(n;\mathbb{R})$$
for some n. Since $\pi\colon \widetilde{\mathrm{SL}(2;\mathbb{R})} \to \mathrm{SL}(2;\mathbb{R})$ determines an isomorphism of Lie algebras, we can consider the Lie algebra homomorphism $d\phi \circ (d\pi)^{-1}\colon \mathfrak{sl}(2;\mathbb{R}) \to \mathfrak{gl}(n;\mathbb{R})$. By the exercise mentioned above, this Lie algebra homomorphism gives rise to a Lie group homomorphism $\psi\colon \mathrm{SL}(2;\mathbb{R}) \to \mathrm{GL}(n,\mathbb{R})$. But $\psi\pi$ induces the same Lie algebra homomorphism as ϕ. Therefore, by Corollary 2.5 we have $\psi\pi = \phi$. But $\ker \pi = \pi_1(\mathrm{SL}(2;\mathbb{R}))$ is non-trivial

and we have $\phi(\ker \pi) = e$ so ϕ is not faithful. We conclude that $\widetilde{SL(2;\mathbb{R})}$ does not have any faithful representations, i.e. it can not be realized as a group of matrices.

To complete our initial discussion of the Lie correspondence we will now state a fundamental theorem.

Theorem 2.13 (Lie's third theorem, or Cartan-Lie Theorem)**.** *Let $\mathfrak{g}$ be a finite dimensional Lie algebra over $\mathbb{R}$. There exists a unique (up to isomorphism) simply connected Lie group with Lie algebra isomorphic to $\mathfrak{g}$.*

The proof of this Theorem will be given later, when we have sufficiently developed the structure theory of Lie groups and algebras. You can also see [DK] for a direct proof. For now we will just point out that this follows from the following (difficult) algebraic result.

Theorem 2.14 (Ado)**.** *Every finite dimensional real Lie algebra $\mathfrak{g}$ has a faithful representation.*

Proof. See for instance [Kn]. □

Proof of Theorem 2.13. Let $f\colon \mathfrak{g} \to \mathfrak{gl}(n;\mathbb{R})$ be a faithful representation. By Theorem 2.3, there exists a Lie subgroup $\phi\colon G \to \mathrm{GL}(n;\mathbb{R})$ with $\mathrm{Lie}(G) \cong f(\mathfrak{g})$. The universal cover of G provides the required Lie group. If G' is another simply connected Lie group and $h\colon \mathfrak{g}' \to \mathfrak{g}$ is a Lie algebra isomorphism, Theorem 2.8 will give us a map $\psi\colon G' \to G$ with $\psi_* = f$. By Proposition 2.7 ψ is a covering space, but since G is simply connected, ψ must be an isomorphism. □

Theorem 2.13 states that the Lie correspondence D is essentially surjective, even when restricted to the category of simply connected Lie groups while Corollary 2.8 implies that D is fully faithful on this subcategory. This implies the following basic result.

Corollary 2.15. *The Lie correspondence D is an equivalence between the categories of simply connected Lie groups and finite dimensional real Lie algebras.*

Let G be a connected topological group and $\pi\colon G \to H$ be a homomorphism of topological groups which is a covering map. Then $\ker \pi$ is a discrete, central subgroup of G (it is central because for any $g \in \ker \pi$ the conjugation map $c_g\colon G \to G$ given by $c_g(x) = gxg^{-1}$ satisfies $c_g(e) = e$ and covers the identity morphism of H hence, by the lifting theorem, must be the identity map of G). Thus all connected Lie groups with Lie algebra $\mathfrak{g}$ can be expressed as the quotient of the unique simply connected Lie group with Lie algebra $\mathfrak{g}$ by a discrete central subgroup. Conversely it is easy to check that a quotient by a discrete central subgroup will give a covering map and there is a natural group structure on the quotient. Theorem 2.13 therefore implies that the classification of connected Lie groups up to isomorphism amounts to

- the classification of finite dimensional real Lie algebras.
- For each such Lie algebra, understanding the discrete subgroups of the center of the simply connected model[2].

[2]But note that different central subgroups may correspond to isomorphic quotient groups. For instance there are uncountably many discrete subgroups of $\mathbb{R}$ but only two possible quotients up to isomorphism.

If one is interested in the classification of all Lie groups one would further need to understand extensions of connected Lie groups by arbitrary countable discrete subgroups.

3. The exponential map

Definition 3.1. *Let G be a Lie group and $X \in \mathfrak{g}$. We will write*

$$\exp_X \colon \mathbb{R} \to G$$

for the unique Lie group homomorphism such that $d\exp_X(\frac{d}{dt}_{|0}) = X_e$. The exponential map $\exp\colon \mathfrak{g} \to G$ *is defined by the formula*

$$\exp(X) = \exp_X(1)$$

Example 3.2. *If $X \in \mathfrak{gl}(n;\mathbb{R})$ then the map $\mathbb{R} \to \mathrm{GL}(n;\mathbb{R})$ given by*

$$t \mapsto e^{tX}$$

(where e^A denotes the matrix exponential of A) is a Lie group homomorphism and

$$\frac{d}{dt}\left(e^{tX}\right)_{|t=0} = X$$

so we conclude that $\exp_X(t) = e^{tX}$ and hence

$$\exp(X) = e^X$$

(which justifies the name of the exponential map).

Note that, since $\exp_X$ is a group homomorphism we have

$$\begin{aligned}
\frac{d}{dt}\left(\exp_X(t)\right) &= d\exp_X\left(\frac{d}{dt}_{|t}\right) \\
&= d\exp_X\left(dL_t \frac{d}{dt}_{|0}\right) \\
&= dL_{\exp_X(t)}\left(d\exp_X\left(\frac{d}{dt}_{|0}\right)\right) \\
&= dL_{\exp_X(t)}(X) = X_{\exp_X(t)}
\end{aligned}$$

hence $t \mapsto \exp_X(t)$ is nothing other than the integral curve of the left invariant vector field X which takes the value e at $t = 0$. By left invariance,

$$t \mapsto g \cdot \exp_X(t)$$

is the integral curve of X which passes through g at $t = 0$ and hence

$$(g,t) \mapsto g \cdot \exp_X(t)$$

is the flow of the vector field X. Namely we have

$$\phi_X^t(g) = g\exp_X(t) \qquad \exp(X) = \phi_X^1(e) \tag{2}$$

Proposition 3.3. *(i)* $\exp(tX) = \exp_X(t)$
(ii) $\exp(-tX) = \exp(tX)^{-1}$

(iii) $\exp((t_1 + t_2)X) = \exp(t_1 X)\exp(t_2 X)$

(iv) $\exp\colon \mathfrak{g} \to G$ *is smooth and* $(d\exp_X)_0 = \mathrm{Id}$ *(under the canonical identification of* $\mathfrak{g}$ *with* T_eG*)*

(v) $\exp$ *is natural, i.e, given* $\phi\colon G \to H$ *a Lie group homomorphism then the following diagram commutes:*

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "d\phi_e"] \arrow[d, "\exp"] & \mathfrak{h} \arrow[d, "\exp"] \\ G \arrow[r, "\phi"] & H \end{tikzcd}$$

(vi) Let $\phi\colon A \hookrightarrow G$ *be a Lie subgroup. Then* $X \in d\phi_e(\mathfrak{a})$ *iff* $\exp(tX) \in A$ *for all* $t \in \mathbb{R}$ *(or even for all* t *in some interval with non-empty interior).*

Proof. (i) We'll check more generally that $\exp_{tX}(s) = \exp_X(st)$ for all s and t. The statement we want is the case when $s = 1$. On the one hand we have

$$\frac{d}{ds}(\exp_{tX}(s)) = (tX)_{|\exp_{tX}(s)}$$

on the other, by the chain rule,

$$\frac{d}{ds}(\exp_X(st)) = tX_{\exp_X(st)}$$

so both expressions give integral curves for the vector field tX. SInce they agree at $s = 0$ they must be equal.

(ii) Applying (i) we have $\exp(-tX) = \exp_X(-t) = (\exp_X(t))^{-1}$ (as $\exp_X$ is a group homomorphism).

(iii) This is just the statement that $\exp_X$ is a group homomorphism.

(iv) exp is smooth by (2) (smooth dependence of differential equations on parameters). Alternatively, we can deduce smoothness from smooth dependence on initial conditions of an auxiliary vector field: consider the vector field $Z \in \mathcal{X}(G \times \mathfrak{g})$ given by $Z(g, X) = (X_g, 0)$, which is smooth by an argument analogous to that in Lemma 1.10(i) (Exercise). Since

$$t \mapsto (g\exp(tX), X)$$

is an integral curve of Z which takes the value (g, X) at time 0, the expression for the flow of Z is

$$\phi_Z^t(g, X) = (g\exp(tX), X)$$

and this is smooth in t, g, X (by smooth dependence on initial conditions). Since $\exp(X) = \pi_1(\phi_X^1(e, X))$ we see that exp is smooth.

Finally

$$d\exp(X_e) = \frac{d}{dt}(\exp(tX))_{|t=0} = X_e.$$

(v) This follows from Corollary 2.5 since $t \mapsto \exp(td\phi(X_e))$ and $\phi(\exp(tX_e))$ are both 1-parameter subgroups of H (meaning Lie group homomorphism $\mathbb{R} \to H$) with the same derivative at $t = 0$.

(vi) $\Rightarrow$ follows immediately from (v). For the converse note that

$$\frac{d}{dt}\left(\exp(tX)\right) = dL_{\exp(tX)}(X)$$

If $\exp(tX) \in \phi(A)$ for $t \in]a, b[$ then, for such t we have $dL_{\exp(tX)}(X) \in T_{\exp(tX)}(\phi(A)) = dL_{\exp(tX)}(d\phi_e(\mathfrak{a}))$ hence $X \in d\phi_e(\mathfrak{a})$.

□

Statement (iv) above implies that exp is a diffeomorphism between a neighborhood of 0 in $\mathfrak{g}$ and a neighborhood of e in G. It therefore provides *canonical coordinates*[3] for G near the identity

$$\log\colon G \to \mathfrak{g}$$

(and by translation, near any point $g \in G$). There is a famous universal formula expressing the product of the Lie group in these canonical coordinates. This formula makes explicit the way in which the product on the Lie group is determined in a neighborhood of e by the Lie bracket on its Lie algebra.

Theorem 3.4 (Baker-Campbell-Hausdorff formula)**.**

$$\log(\exp(X)\exp(Y)) = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}\left([X,[X,Y]] + [Y,[Y,X]]\right) + \dots$$

where ... are higher order terms[4] *which can be expressed in terms of linear combinations of iterated Lie brackets of X and Y.*

Proof. See [H] or [DK]. □

Note that a map of Lie groups $p\colon G \to H$ induces an isomorphism of Lie algebras iff it is a covering map and then p induces a diffeomorphism between neighborhoods of the identity, preserving the products defined on those neighborhoods, which is consistent with the statement above.

Let us just check the validity of the second order term in the case of a matrix group. In this case, $\exp, \log$ are the usual exp and log of matrices. We can write

$$e^X e^Y = e^{A_0(X,Y)+A_1(X,Y)+A_2(X,Y)+\dots},$$

where $A_n(X,Y)$ are the homogeneous terms of degree n in the power series expansion of the product in exponential coordinates. Since $\exp(0) = e$ and $d(\exp)_0 = \mathrm{Id}$ we must have $A_0(X,Y) = 0$ and $A_1(X,Y) = X + Y$. Then

$$(I+X+\frac{1}{2!}X^2+\dots)(I+Y+\frac{1}{2!}Y^2+\dots) = I+X+Y+A_2(X,Y)+\frac{1}{2}(X+Y+A_2(X,Y))^2+\dots$$

and equating the second order terms we get

$$\frac{1}{2}X^2 + XY + \frac{1}{2}Y^2 = A_2(X,Y) + \frac{1}{2}(X^2 + XY + YX + Y^2)$$

[3]One can check that these are the same as the canonical coordinates of Riemannian geometry if G can be given a bi-invariant metric. For instance, if G is compact.

[4]There is an explicit formula due to Dynkin. See [DK].

and hence

$$A_2(X,Y) = \tfrac{1}{2}(XY - YX) = \tfrac{1}{2}[X,Y]$$

The computation becomes more and more difficult as we proceed. Even in degree 3, it is not immediately obvious that the expression for $A_3(X,Y)$ we obtain by expanding on both sides can be expressed in terms of iterated brackets. It is a good exercise to check that it does agree with the formula given above.

We will now use the exponential map to prove a very useful theorem concerning Lie subgroups.

Theorem 3.5 (Cartan's closed subgroup theorem)**.** *Let G be a Lie group and $A \subset G$ a subset which is a subgroup and a closed subset. Then A is an embedded Lie subgroup.*

Before we look at the proof, let's see some applications.

Example 3.6. *(i) Let G be a Lie group. Then its center*

$$Z(G) = \{g \in G \colon gh = hg \quad \forall h \in G\}$$

is a Lie group. Indeed, it is a subgroup and clearly a closed subset as the equalities in the definition are closed relations.

(ii) Let V be a vector space over $\mathbb{R}$ and $B : V \times V \to \mathbb{R}$ be a bilinear form. Then,

$$\mathrm{Aut}_B(V) = \{\varphi \in GL(V) \colon B(\varphi(v), \varphi(w)) = B(v,w)\}$$

is a Lie group as it is clearly a closed subgroup of $\mathrm{GL}(V)$. The same is true if we replace $\mathbb{R}$ by $\mathbb{C}$ or $\mathbb{H}$.

When B is the standard inner product on $\mathbb{R}^n$, we obtain the orthogonal group $O(n)$, when B is the standard Hermitian inner product on $\mathbb{C}^n$ we obtain the unitary group $U(n)$. Another important example is the case when B is a non-degenerate skew symmetric bilinear form on $\mathbb{R}^{2n}$. In a suitable basis B can be written as

$$B(v,w) = v^T J w \qquad \text{with } J = \begin{bmatrix} 0 & -I_{n\times n} \\ I_{n\times n} & 0 \end{bmatrix}$$

The group of linear isomorphisms preserving this bilinear form is called the symplectic linear group

$$\mathrm{Sp}(2n,\mathbb{R}) = \{A \in GL(\mathbb{R}^{2n}) \colon A^T J A = J\}$$

When B is the non-degenerate symmetric bilinear form on $\mathbb{R}^n$ with signature $(k, n-k)$, so that

$$B((x_1,\dots,x_n),(x_1,\dots,x_n)) = x_1^2 + \dots x_k^2 - x_{k+1}^2 - \dots - x_n^2$$

the resulting Lie group is denoted $O(k, n-k)$. Of special importance is the group of symmetries of Minkowski space $O(3,1)$.

(iii) Let $\phi : G \to H$ be a Lie group homomorphism. Then $\ker\phi$ is a Lie subggroup of G. Using the exponential map one checks that its Lie algebra is $\ker\phi_* \colon \mathfrak{g} \to \mathfrak{h}$: by Proposition 3.3 (vi) we have

$$X \in \mathrm{Lie}(\ker\phi) \Leftrightarrow \exp(tX) \in \ker\phi \text{ for all } t \in \mathbb{R} \Leftrightarrow \phi(\exp(tX)) = e \text{ for all } t \in \mathbb{R}$$

By naturality of exp *the last condition is equivalent to*

$$\exp(td\phi_e(X)) \in \{e\} \text{ for all } t \in \mathbb{R}$$

and again by Proposition 3.3 (vi) this is equivalent to

$$d\phi_e(X) \in \{0\} \Leftrightarrow X \in \ker \phi_*$$

(iv) *If $B : V \times V \to V$ is a bilinear map (i.e. a product, not necessarily unital, associative or commutative) then*

$$\mathrm{Aut}_B(V) = \{\varphi \in \mathrm{GL}(V) \colon B(\varphi(v), \varphi(w)) = \varphi(B(v, w))\}$$

is a Lie subgroup of $\mathrm{GL}(V)$*. In particular, if $V = \mathfrak{g}$ and $B = [\,,\,]$ we have that the group of automorphisms of a Lie algebra* $\mathrm{Aut}(\mathfrak{g})$ *is a Lie group. Corollary 2.15 then gives us a canonical Lie group structure on the automorphism group of a simply connected Lie group. Using the exponential map it is not hard to see that the resulting topology on the group of automorphisms of the Lie group is induced by the compact-open topology on the space of continuous maps from the Lie group to itself. Using this and Cartan's closed subgroup theorem, it then follows that the group of automorphisms of an arbitrary Lie group is a Lie group (Exercise).*

The following result identifies the Lie algebra of the Lie groups of automorphisms of a product and explains the reason for the ubiquity of derivations: they are the infinitesimal automorphisms.

Proposition 3.7. *Let V be a vector space and $B\colon V \times V \to V$ a bilinear map. The Lie algebra of* $\mathrm{Aut}_B(V)$ *is the Lie algebra of B-derivations.*

$$\mathrm{Der}_B(V) = \{\delta \in \mathfrak{gl}(V) \colon \delta(B(v, w)) = B(\delta v, w) + B(v, \delta w)\}$$

Proof. Let's first check that a tangent vector to $\mathrm{Aut}_B(V)$ at e is a derivation. Let $\varphi : (-\varepsilon, \varepsilon) \to \mathrm{Aut}_B(V)$ be a smooth curve with $\varphi(0) = e$ and let $X = \varphi'(0) \in \mathfrak{gl}(V)$. By assumption we have $\varphi(t)(B(v, w)) = B(\varphi(t)v, \varphi(t)w)$. Differentiating with respect to t we obtain

$$\left(\frac{d}{dt}\varphi(t)(B(v,w))\right)_{|t=0} = B\left(\frac{d}{dt}(\varphi(t))_{|t=0}v, \varphi(0)w\right) + B\left(\varphi(0)v, \frac{d}{dt}(\varphi(t))_{|t=0}w\right)$$

and hence $X(B(v, w)) = B(X(v), w) + B(v, X(w))$ so that X is a derivation.

Now suppose $\delta\colon V \to V$ is a B-derivation and let us check that

$$e^{t\delta} = \mathrm{id} + t\delta + \frac{1}{2!}(t\delta)^2 + \dots$$

is an isomorphism of B. That will show $\delta \in \mathrm{Lie}(\mathrm{Aut}_B(V))$ as $\frac{d}{dt}(e^{t\delta})_{|t=0} = \delta$. We can regard B as a linear map $B\colon V \otimes V \to V$ and, in this terms, we need to check whether $B \circ (e^{t\delta} \otimes e^{t\delta})$ equals $e^{t\delta} \circ B$.

We have

$$e^{t\delta} \circ B = \sum_{n=0}^{\infty} \frac{t^n}{n!}\delta^n \circ B$$

Since δ is a derivation we have $\delta \circ B = B \circ (\delta \otimes \mathrm{id} + \mathrm{id} \otimes \delta)$ and therefore the expression above is equal to

$$\sum_{n=0}^{\infty} \frac{t^n}{n!} B \circ (\delta \otimes \mathrm{id} + \mathrm{id} \otimes \delta)^n = B \circ \sum_{n=0}^{\infty} \frac{t^n}{n!} (\delta \otimes \mathrm{id} + \mathrm{id} \otimes \delta)^n = B \circ e^{t(\delta \otimes \mathrm{id} + \mathrm{id} \otimes \delta)}$$

Since the endomorphisms $\delta \otimes \mathrm{id}$ and $\mathrm{id} \otimes \delta$ of $V \otimes V$ commute, we see that this equals

$$B \circ (e^{t(\delta \otimes \mathrm{id})} \circ e^{t(\mathrm{id} \otimes \delta)}) = B \circ (e^{t\delta} \otimes \mathrm{id}) \circ (\mathrm{id} \otimes e^{t\delta}) = B \circ (e^{t\delta} \otimes e^{t\delta})$$

as required. □

We can repeat the computation in the proof of the previous proposition replacing B with a bilinear form $b \colon V \times V \to \mathbb{R}$ (or some other field of scalars). The derivations of the scalar product b are now the linear maps

$$\mathrm{Der}_b(V) = \{\delta \in \mathrm{End}(V) \colon b(\delta v, w) + b(v, \delta w) = 0 \text{ for all } v, w \in V\}$$

and just as above we see this is the Lie algebra of $\mathrm{Aut}_b(V)$ (exercise). When b is the skew-symmetric non-degenerate bilinear form on $\mathbb{R}^{2n}$ we obtain the *real symplectic Lie algebra*

$$\mathfrak{sp}(2n; \mathbb{R}) = \{A \in \mathfrak{gl}(\mathbb{R}^{2n}) \colon A^T J + JA = 0\}$$

Having seen some applications we will now prove Theorem 3.5. We will use the following Lemma.

Lemma 3.8. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. Let A be an abstract subgroup. Suppose that there is a subspace $\mathfrak{a} \subset \mathfrak{g}$ and a neighbourhood U of 0 in $\mathfrak{g}$ such that*

- $\exp : U \to V = \exp(U) \subset G$ *is a diffeomorphism.*
- $A \cap V = \exp(\mathfrak{a} \cap U)$.

Then A is an embedded Lie group and $\mathfrak{a}$ is its Lie algebra.

Proof. The open sets $\{L_a(W \cap A) \colon a \in A, W \subset V\}$ are a basis for the subspace topology on A (because $\{L_g(W) \colon W \subset V\}$ is a basis for the topology on G). On each of this basis elements, we have a chart

$$L_a(W \cap A) \xrightarrow{(\exp_{|U})^{-1} \circ L_{a^{-1}}} U \cap \mathfrak{a}.$$

whose image is some open subset of $U \cap \mathfrak{a}$. The changes of coordinates for these charts are of the form

$$L_{a'} \circ \exp_{|U} \circ (\exp_{|U})^{-1} \circ L_{a^{-1}}$$

for $a, a' \in A$, defined on some open subset of $U \cap \mathfrak{a}$.

These are restrictions to $\mathfrak{a}$ of diffeomorphisms between open subsets of $\mathfrak{g}$ and hence they are diffeomorphisms between open subsets of $\mathfrak{a}$. These charts give us an atlas for A with respect to the induced topology so $A \hookrightarrow G$ is an embedded submanifold (in the local coordinates described above, the inclusion is a linear map - the inclusion of an open subset of $\mathfrak{a}$ into $\mathfrak{g}$). The restriction of the product and the inverse to A are smooth, hence A is an embedded Lie group.

Since $d\exp_0 = \mathrm{Id}$ we see that $T_e A = \mathfrak{a}$ and therefore $\mathfrak{a}$ is a Lie subalgebra of $\mathfrak{g}$. □

Proof of Theorem 3.5. Given a closed subgroup A define

$$\mathfrak{a} = \{X \in \mathfrak{g} \colon \exp(tX) \in A \text{ for all } t \in \mathbb{R}\}.$$

Clearly $\mathfrak{a}$ is a cone in $\mathfrak{g}$. Let's see that it is actually a subspace. For $X, Y \in \mathfrak{g}$ and $t \in \mathbb{R}$

$$\exp(tX)\exp(tY) = \exp(t(X+Y) + O(t^2)) \tag{3}$$

This follows from the fact that the maps

$$\mathfrak{g} \times \mathfrak{g} \xrightarrow{\exp \times \exp} G \times G \xrightarrow{\mu} G$$

$$\mathfrak{g} \times \mathfrak{g} \xrightarrow{+} \mathfrak{g} \xrightarrow{\exp} G$$

both have $[I|I]$ as a derivative at $(0,0)$ and hence have contact of order ≥ 1. This implies equation (3) with $O(t^2)$ outside exp but since exp is a diffeomorphism at 0, that is equivalent to (3).

Given $X, Y \in \mathfrak{a}$

$$\begin{aligned}\left(\exp\left(\frac{X}{n}\right)\exp\left(\frac{Y}{n}\right)\right)^n &= \exp\left(\frac{X}{n} + \frac{Y}{n} + O\left(\frac{1}{n^2}\right)\right)^n \\ &= \exp\left(\frac{1}{n}\left(X + Y + O\left(\frac{1}{n}\right)\right)\right)^n \\ &= \exp\left(X + Y + O\left(\frac{1}{n}\right)\right) \in A.\end{aligned}$$

Since A is closed,

$$\lim_{n\to\infty} \exp\left(X + Y + \mathcal{O}\left(\frac{1}{n}\right)\right) = \exp(X+Y) \in A.$$

Now if $X, Y \in \mathfrak{a}$ then, by definition $tX, tY \in \mathfrak{a}$ and hence, from the above, $\exp(t(X+Y)) \in A$ for all $t \in \mathbb{R}$. It follows that $\mathfrak{a}$ is closed under sum and multiplication by scalar so it is a subspace of $\mathfrak{g}$.

We now need to check the hypotheses of Lemma 3.8 for A and the subspace $\mathfrak{a}$ above. Assume they do not hold. Then, we can pick a sequence $Z_k \in G$ so that

- $Z_k \to e$ (and $Z_k \neq e$)
- $Z_k \in A$.
- $\exp^{-1}(Z_k) \notin \mathfrak{a}$.

Pick a complement $\mathfrak{b}$ to $\mathfrak{a}$ in $\mathfrak{g}$ (so that $\mathfrak{a} \oplus \mathfrak{b} = \mathfrak{g}$) and consider the map

$$\phi : (\mathfrak{a} \oplus \mathfrak{b}) \cap U \to G$$

defined by

$$\phi(X, Y) = \exp(X)\exp(Y).$$

Its derivative at $(0,0)$ is the isomorphism $\mathfrak{a} \times \mathfrak{b} \xrightarrow{+} \mathfrak{g}$. Since ϕ is a local diffeomorphism we can write for k large enough $Z_k = \exp(X_k)\exp(Y_k)$ with $X_k \in \mathfrak{a}$, $Y_k \in \mathfrak{b}$ and $X_k, Y_k \to 0$. Since $\exp(X_k) \in A$ and $Z_k = \exp(X_k)\exp(Y_k) \in A$ we must have $\exp(Y_k) \in A$. Moreover

$Y_k \neq 0$, otherwise we would have $\exp^{-1}(Z_k) = X_k \in \mathfrak{a}$. Pick some norm on $\mathfrak{b}$ and a subsequence so that $\frac{Y_k}{\|Y_k\|} \to Y \in \mathfrak{b}$. Then

$$\underbrace{\left\lfloor \frac{1}{\|Y_k\|} \right\rfloor}_{n_k} Y_k \to Y$$

(the distance to Y is bounded by the sum of $\|Y_k\|$ and the angle between $\frac{Y_k}{\|Y_k\|}$ and Y which both tend to 0). Since exp is continuous

$$\exp(n_k Y_k) = \exp(Y_k)^{n_k} \to \exp(Y)$$

and since A is closed, it follows that $\exp(Y) \in A$. Fixing $t \neq 0$ and running the same argument with $n_k = \lfloor \frac{t}{n_k} \rfloor \in \mathbb{Z}$ we see that $\exp(tY) \in A$ for all $t \in \mathbb{R}$ and hence $Y \in \mathfrak{a}$, which is a contradiction. □

4. The adjoint actions

Definition 4.1. *Let M be a smooth manifold and G be a Lie group. A* left action *of G on M is a smooth map $\phi : G \times M \to M$ satisfying:*

(i) $\phi(e, m) = m$, $\forall m \in M$.
(ii) $\phi(g, \phi(h, m)) = \phi(gh, m)$, $\forall g, h \in G$, $\forall m \in M$.

The action of $g \in G$ on $m \in$ is usually denoted by $g \cdot m$ or even gm. We also have the notion of a *right action.* This is a smooth map $M \times G \to M$ sending (m, g) to $m \cdot g$ and satisfying

$$m \cdot e = m \qquad (m \cdot g) \cdot h = m \cdot (gh) \text{ for all } g, h \in G,\ m \in M$$

For instance a Lie group G acts smoothly on itself on the left by left multiplication and on the right by right multiplication:

$$g \cdot x = gx \qquad x \cdot g = xg \qquad \text{for } g, x \in G$$

We need to set some terminology regarding actions. An action is **effective** if $g \cdot m = m$ for all m implies $g = e$ (i.e. all non-identity elements of G move some point of M). The **orbit** of a point m is the subset of M

$$G \cdot m = \{g \cdot m \colon g \in G\}$$

An action is **transitive** if there is only one orbit, i.e. if any point is accessible to any other via the action of G. Given a point $m \in M$, the **isotropy subgroup** or **stabilizer subgroup** of m is the subgroup

$$G_m = \{g \in G \colon g \cdot m = m\}$$

An action is **free** if $G_m = \{e\}$ for all $m \in M$ (i.e. all non-identity elements of G move every point of M). This is often confused with the notion of an effective action, but it is stronger as the following example shows.

Example 4.2. *The action*

$$\begin{aligned} \mathbb{Z}/2 \times \mathbb{R} &\to \mathbb{R} \\ (\pm 1, x) &\to \pm x \end{aligned}$$

is effective but not free as the isotropy group of 0 *is* $\mathbb{Z}/2$.

A **fixed point** of an action is an element $m \in M$ such that $g \cdot m = m$ for all $g \in G$, equivalently it is an element of M whose isotropy group is G.

Note that for each $g \in G$, the map

$$\begin{aligned} a_g : M &\to M \\ m &\to g \cdot m \end{aligned}$$

is a diffeomorphism with inverse $a_{g^{-1}}$. The axioms (i), (ii) in the definition of an action precisely say that

$$\begin{aligned} G &\to \mathrm{Diff}(M) \\ g &\mapsto a_g \end{aligned}$$

is a group homomorphism. So an action is a kind of "nonlinear representation", a way of regarding elements in G as diffeomorphisms of a manifold. The parallel is even closer since $\mathrm{Diff}(M)$ can be given the structure of an infinite dimensional Lie group for which the map $g \mapsto a_g$ is smooth. We will not get into this, but it is easy to see that the map is at least a homomorphism of topological groups when we give $\mathrm{Diff}(M)$ the standard Whitney topology where convergence means uniform convergence of a map and all its derivatives on compact subsets of M.

Example 4.3. *(i) A smooth action of the additive group* $\mathbb{R}$, $a : \mathbb{R} \times M \to M$, *is called a* flow *on* M. *It is exactly the same as a complete vector field on* M. *Indeed, given such an action we can define* $X \in \mathcal{X}(M)$ *by the formula*

$$X_m = \frac{d}{dt}(t \cdot m)_{|t=0}.$$

It is easy to check that X *is smooth (exercise). We have*

$$\begin{aligned} \frac{d}{dt}(a(t,m)) &= \frac{d}{ds}(a(t+s,m))_{|s=0} \\ &= \frac{d}{ds}(s \cdot (t \cdot m))_{s=0} \\ &= X_{t \cdot m} \end{aligned}$$

so $t \to t \cdot m$ *is an integral curve for* X. *It follows that* X *is complete. Conversely, if* X *is a complete vector field, we can define a smooth action of* $\mathbb{R}$ *by the formula* $t \cdot m := \phi^t_X(m)$.

(ii) A representation $\rho : G \to GL(V)$ *gives an action*

$$\begin{aligned} G \times V &\to V \\ (g, v) &\to \rho(g)v. \end{aligned}$$

Conversely, given such an action with the property that for each g

$$v \to g \cdot v$$

is linear, then the mapping defined by

$$\begin{aligned} \rho : G &\to GL(V) \\ g &\to (v \to g \cdot v) \end{aligned}$$

is a representation (it is an exercise to check that the resulting map $G \to \mathrm{GL}(V)$ *is smooth).*

(iii) The **conjugation action** *of a Lie group on itself is*

$$\begin{aligned} G \times G &\to G \\ (g, h) &\to c_g(h) = ghg^{-1} \end{aligned}$$

Note that this is an action by automorphisms of the Lie group G*, since*

$$\begin{aligned} c_g(e) &= e \\ c_g(h_1 h_2) &= c_g(h_1) c_g(h_2) \end{aligned}$$

The conjugation action is effective iff $ghg^{-1} = h$ *for all* h *implies that* $g = e$*. Since*

$$ghg^{-1} = h \Leftrightarrow gh = hg$$

we see that this happens if and only if the center of G *is trivial, i.e.* $Z(G) = \{e\}$*. The orbits of the conjugation action are, by definition, the conjugation classes of* G*. The set* $\{e\}$ *is always a conjugation class therefore the action is transitive if and only if the group* G *is the trivial group. Similarly the action is free only if* G *is the trivial group, as* e *is a fixed point. The isotropy group of an element* $h \in G$ *is*

$$G_h = \{g \in G \colon c_g(h) = h \Leftrightarrow gh = hg\} = Z_G(h)$$

the centralizer of the element h *in* G*.*

(iv) Given a smooth representation $G \times M \to M$*, the* **isotropy representation** *at* m *is defined by*

$$\begin{aligned} G_m &\to \mathrm{GL}(T_m M) \\ g &\to d(g\cdot)_m, \end{aligned}$$

This is a "linearization of the action of the isotropy group near m*" as we will make precise later. It is easy to check that the isotropy group acts smoothly on the tangent space (exercise).*

Definition 4.4. *The* adjoint representation *of* G *is the isotropy representation of the conjugation action of* G *on itself at* e*.*

$$\begin{aligned} G &\xrightarrow{\mathrm{Ad}} \mathrm{Aut}(\mathfrak{g}) \subset \mathrm{GL}(\mathfrak{g}) \\ g &\mapsto (X \mapsto dc_g(X)). \end{aligned}$$

Ad *is a Lie group homomorphism and so its derivative gives a representation of the Lie algebra which is denoted by*

$$\mathrm{ad}\colon \mathfrak{g} \to \mathrm{Der}(\mathfrak{g}) \subset \mathrm{End}(\mathfrak{g})$$

and is also called the adjoint representation *of the Lie algebra* $\mathfrak{g}$.

Note that when G is connected, $Ad(g) = \mathrm{id} \Leftrightarrow c_g = \mathrm{id} \Leftrightarrow g \in Z(G)$. Therefore, when G has trivial center, Ad is a faithful representation and G is canonically a matrix group. The naturality property of exp (Proposition 3.3 (v)) implies the following very useful commutation relations:

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\mathrm{Ad}(g)"] \arrow[d, "\exp"] & \mathfrak{g} \arrow[d, "\exp"] \\ G \arrow[r, "c_g"] & G \end{tikzcd} \qquad \begin{tikzcd} \mathfrak{g} \arrow[r, "\mathrm{ad}"] \arrow[d, "\exp"] & \mathrm{Der}(\mathfrak{g}) \arrow[d, "\exp"] \\ G \arrow[r, "\mathrm{Ad}"] & \mathrm{Aut}(\mathfrak{g}) \end{tikzcd}$$

Example 4.5. *Let $G = \mathrm{GL}(n, \mathbb{R})$ so that we have*

$$c_g(A) = gAg^{-1}$$

Since conjugation by g is a linear map from $M_{n\times n}(\mathbb{R})$ to $M_{n\times n}(\mathbb{R})$, the automorphism of G given by c_g is the restriction of a linear map to an open set. With the usual identifications, the derivative of c_g at any point is itself, hence $Ad(g)A = gAg^{-1}$, for $A \in \mathfrak{gl}(n, \mathbb{R}) = M_{n\times n}(\mathbb{R})$. In order to compute ad *we can compute*

$$\begin{aligned} ad(X)(A) &= \frac{d}{dt}\left(Ad(\exp tX)A\right)_{|t=0} \\ &= \frac{d}{dt}\left(\exp(tX)A\exp(-tX)\right)_{|t=0} \\ &= XA\exp(0X) + \exp(0X)A(-X) \\ &= XA - AX \\ &= [X, A] \end{aligned}$$

The formula we obtained for ad in the previous example is actually completely general.

Proposition 4.6. *Let G be a Lie group with Lie algebra $\mathfrak{g}$ and let $X \in \mathfrak{g}$. Then, $ad(X)(Y) = [X, Y]$.*

Proof. Since $c_g = R_{g^{-1}}L_g$, applying the chain rule we have

$$\mathrm{ad}(X)Y = \frac{d}{dt}\left(Ad(\exp tX)Y\right)_{|t=0} = \frac{d}{dt}\left(dR_{\exp(-tX)}dL_{\exp(tX)}Y\right)_{|t=0} = \frac{d}{dt}\left(dR_{\exp(-tX)}Y_{\exp(-tX)}\right)_{|t=0}$$

Recalling formula (2) for the flow of the vector field X the last expression is, by definition, the Lie derivative of Y along X hence

$$\mathrm{ad}(X)(Y) = (\mathcal{L}_X Y)_e = [X, Y].$$

□

Motivated by the previous Proposition we make the following definition.

Definition 4.7. *Let $\mathfrak{g}$ be a Lie algebra over an arbitrary field.*

$$\mathrm{ad} : \mathfrak{g} \to \mathrm{Der}(\mathfrak{g})$$

by the formula $\mathrm{ad}(X)(Y) = [X, Y]$ *(recall that the Jacobi identity guarantees that* $\mathrm{ad}(X)$ *is a derivation, and also that* ad *is a map of Lie algebras for the standard commutator bracket on derivations). This map* ad *is still called the* adjoint representation *of* $\mathfrak{g}$*. The derivations of the form* $\mathrm{ad}(X)$ *are called* inner derivations*.*

The center *of a Lie algebra is*

$$Z(\mathfrak{g}) = \{X \in \mathfrak{g} \colon [X, Y] = 0 \textit{ for all } Y \in \mathfrak{g}\}$$

Note that the kernel of the adjoint representation of $\mathfrak{g}$ is exactly the center of $\mathfrak{g}$. Thus, if $Z(\mathfrak{g}) = \{0\}$ then the adjoint representation is faithful and $\mathfrak{g}$ is (isomorphic to) a matrix Lie algebra.

Proposition 4.8. *Let G be a connected Lie group.*

(i) $Lie(Z(G)) = Z(\mathfrak{g})$ and in particular G is abelian iff $\mathfrak{g}$ is abelian.

(ii) Let $A \hookrightarrow G$ be a connected Lie subgroup. Then, A is normal (i. e., $gAg^{-1} = A$ for all $g \in G$) iff $\mathfrak{a}$ is an ideal *of $\mathfrak{g}$ (i. e., $\mathfrak{a}$ is a subspace of $\mathfrak{g}$ such that for all $X \in \mathfrak{g}, Y \in \mathfrak{a}$, $[X, Y] \in \mathfrak{a}$).*

Proof. (i) Since G is connected, $Z(G) = \ker \mathrm{Ad}$, so this is a special case of Example 3.6(iii). We have $\mathrm{Lie}(Z(G)) = \mathrm{Lie}(\ker \mathrm{Ad}) = \ker \mathrm{ad} = Z(\mathfrak{g})$.

(ii) A is normal if and only if $c_g(A) = A$ for all $g \in G$. Since A is connected this is equivalent to $\mathrm{Ad}(g)(\mathfrak{a}) = \mathfrak{a}$ for all $g \in G$. In particular, for all $X \in \mathfrak{g}$ and $t \in \mathbb{R}$ we have $Ad(\exp(tX))\mathfrak{a} = \mathfrak{a}$ Differentiating at $t = 0$ we obtain

$$\mathrm{ad}(X)(\mathfrak{a}) \subset \mathfrak{a} \text{ for all } X \in \mathfrak{g}$$

which precisely means that $\mathfrak{a}$ is an ideal.

Conversely, suppose $\mathfrak{a}$ is an ideal. Then $\mathrm{ad}(X)\mathfrak{a} \subset \mathfrak{a}$ for all $X \in \mathfrak{g}$. Hence $\exp(\mathrm{ad}(tX))\mathfrak{a} \subset \mathfrak{a}$ for all $t \in \mathbb{R}, X \in \mathfrak{g}$. Since G is connected and therefore the image of exp generates G, the last condition is equivalent to $\mathrm{Ad}(g)\mathfrak{a} \subset \mathfrak{a}$ for all $g \in G$. Since A is connected this amounts to saying that $c_g(A) \subset A$ for all g, which precisely means that A is a normal subgroup.

□

5. Smooth actions

We start by giving a condition on actions which make the orbit spaces "reasonable".

Definition 5.1. *Let M, N be smooth manifolds. A continuous map $f : M \to N$ is* proper *if for all compact $K \subset N$, the set $f^{-1}(K)$ is compact.*

Although the definition above makes sense for general topological spaces, the definition of proper map for topological spaces is usually different (it agrees with the above for locally compact Hausdorff spaces). See [tD, 3.13, p.27] for a thorough discussion of this point. In practice the way this property of maps is used is via the following characterization, the proof of which is a simple point set topology exercise.

Exercise 5.2. *A map $f: M \to N$ is proper if and only if any sequence $x_k \in M$ such that $f(x_k)$ converges, has a convergent subsequence (in M).*

It is useful to keep in mind the following immediate consequences of properness.

Remark 5.3. *Let $f: M \to N$ be a proper map between manifolds*

(i) For all $y \in N$, the fiber over *y, $f^{-1}(y) \subset M$ is compact.*
(ii) f is closed, i. e., if $F \subset M$ is closed, then $f(F)$ is closed in N.

In fact the above two conditions characterize properness. This is again a simple point set topology exercise.

Exercise 5.4. *Check that if conditions $(i) + (ii)$ above hold then f proper.*

Definition 5.5. *A smooth action $\phi : G \times M \to M$ is* proper *if the map*

$$\begin{aligned}\Phi : G \times M &\to M \times M \\ (g, m) &\to (g \cdot m, m)\end{aligned}$$

is proper.

Note the following immediate consequences of properness of an action(cf. Remark 5.3):

- Isotropy groups are compact (as $G_m \times \{m\} = \Phi^{-1}(m, m)$).
- Orbits are closed because $G \cdot m = \Phi(G \times \{m\})$.

Example 5.6. *(i) If G is compact, then any smooth action of G is proper. Indeed given $(g_k, m_k) \in G \times M$ such that $(g_k \cdot m_k, m_k)$ converges in $M \times M$, we have that m_k converges and, since G is compact, g_k has a convergent siubsequence.*

(ii) The standard action of $GL(n, \mathbb{R})$ on $\mathbb{R}^n$ is not proper. Indeed, there are only two orbits of this action: $\{0\}$ and $\mathbb{R} \setminus \{0\}$ and the latter is not closed.

(iii) The $\mathbb{R}$-action on $\mathbb{R}^2$ given by

$$t \cdot (x, y) = (x + ty, y)$$

(corresponding to the vector field $y\frac{\partial}{\partial x}$) is not proper because the isotropy of the points $(x, 0)$ is $\mathbb{R}$, which is not compact.

(iv) Let $H \subset G$ be a closed subgroup. The action of H on G by left (or right) translation is proper. Indeed let $(h_k, g_k) \in H \times G$ be such that $(h_k g_k, g_k)$ converges to $(a, b) \in G \times G$. Then $g_k \to b$. Since $h_k g_k \to a$ it follows that $h_k \to ab^{-1} \in G$. Since H is closed, the limit ab^{-1} belongs to H and hence (h_k, g_k) converges in $H \times G$. In this case we don't even need to take a subsequence.

Here is the first manifestation of the "reasonableness" of proper actions.

Proposition 5.7. *If G acts properly on M, then the orbit space $G \backslash M$ is Hausdorff.*

Proof. Let $G \cdot x$ and $G \cdot y$ be distinct orbits. We need to find neighbourhoods U of $G \cdot x$ and V of $G \cdot y$ so that $U \cap V = \emptyset$. Note that a small neighborhood of $G \cdot x$ in the quotient topology is obtained by saturating an arbitrary small neighborhood W of x in M with respect to the action, i.e. by considering a small neighborhood W of $x \in M$ and taking the open set $\cup_{g \in G}\, gW$ with $g \in G$.

Suppose that we can not find neighborhoods U and V as above. Then we can find convergent sequences $x_k \to x$ and $y_k \to y$ in M and g_k, h_k in G such that $g_k x_k = h_k y_k$. Therefore $h_k^{-1} g_k x_k = y_k$ converges and we have

$$\Phi(h_k^{-1} g_k x_k, y_k) \to (y, y)$$

By properness $h_k^{-1} g_k$ has a convergent subsequence to some element $\ell \in G$. Continuity of the action implies that $\ell x = y$ and therefore $G \cdot x = G \cdot y$ which contradicts the assumption that $G \cdot x$ and $G \cdot y$ are distinct. □

Given a smooth action $G \times M \to M$, an element $X \in \mathfrak{g}$ defines a section $a(X) \colon M \to TM$ by the expression

$$a(X)_m = \frac{d}{dt}(\exp(tX)m)_{|t=0}$$

Exercise 5.8. *(i) Show that $a(X)$ is a smooth vector field whose integral curves going through m at $t = 0$ are $t \mapsto \exp(tX) \cdot m$*

(ii) Show that $[a(X), a(Y)] = -a([X, Y])$, i.e. $a \colon \mathfrak{g} \to \mathcal{X}(M)$ is an anti-homomorphism.

The previous exercises suggest the notion of infinitesimal action of a Lie algebra which we now define, and which will be used in the proof of the next theorem.

Definition 5.9. *An* infinitesimal action *of a real Lie algebra $\mathfrak{g}$ on a manifold M is a Lie algebra anti-homomorphism $\mathfrak{g} \to \mathcal{X}(M)$*

It will be an exercise in the next problem set to show that any infinitesimal action of $\mathfrak{g}$ with the property that $a(X)$ is complete for every X can be integrated to a smooth action of the simply connected Lie group corresponding to the Lie algebra $\mathfrak{g}$.

Theorem 5.10. *Let $G \times M \to M$ be a smooth, proper and free action. Then, $G\backslash M$ has a unique smooth structure with the property that the quotient map $\pi : M \to G\backslash M$ has local sections, i.e. for all $G \cdot x \in G\backslash M$, there exists a neighborhood U of $G \cdot x$ in $G\backslash M$ and a smooth map $s : U \to M$ such that $\pi \circ s = id_U$*

Note that a smooth map has local sections if and only if it is a submersion. This allows for an alternative statement of the previous theorem.

Proof. Let us first see that for each $m \in M$ the map

$$\begin{aligned} \mathfrak{g} &\to T_m M \\ X &\to a(X)_m \end{aligned}$$

is injective. If $a(X)_m = 0$ then m is a zero of the vector field $a(X)$ and hence the integral curve through m is constant. This means that $\exp(tX) \cdot m = m$ for all t. Since the action is free, $\exp(tX) = e$ for all $t \in \mathbb{R}$ and therefore $X = 0$.

From this is follows that the orbit maps $G \to M$ given by $g \mapsto g \cdot m$ are immersions. Indeed, writing L_g also for the diffeomorphism of M given by the action of g, the diagram

$$\begin{tikzcd} h \arrow[d, mapsto] & G \arrow[r, "L_g"] \arrow[d, hook] & G \arrow[d, hook] & h \arrow[d, mapsto] \\ h \cdot m & M \arrow[r, "L_g"] & M & h \cdot m \end{tikzcd}$$

shows that the differential of this map at $g \in G$ sends $dL_g(X)$ to $dL_g(a(X)_m)$ and is therefore an inclusion[5].

Now let $S \hookrightarrow M$ be a submanifold containing x such that

$$T_xS \oplus a(\mathfrak{g})_x = T_xM$$

and consider the action map

$$\phi\colon G \times S \to M$$

defined by the expression $\phi(g, s) = g \cdot s$. Since

$$d\phi_{|(e,x)}(X, Y) = a(X)_x + Y,$$

by the inverse function theorem ϕ is a local diffeomorphism at (e, x) and hence also at (e, y) for y in some open set $W \subset S$ containing x. By homogeneity, ϕ is a local diffeomorphism for all $(g, y) \in G \times W$.

Let us see that there is a neighbourhood $V \subset W$ of x in S such that $\phi_{|G\times V}$ is injective and hence a diffeomorphism. If not, we can pick sequences $x_k, y_k \in S$ with $x_k \neq y_k$ such that x_k, y_k converge to x and elements $g_k, h_k \in G$ such that

$$\phi(g_k, x_k) = \phi(h_k, y_k)$$

and hence $h_k^{-1}g_kx_k = y_k$. Since the action is proper, $h_k^{-1}g_k$ must have a convergent subsequence. Without loss of generality let's assume that $h_k^{-1}g_k$ converges to an element $\ell \in G$. Taking the limit on both sides of the equality $h_k^{-1}g_kx_k = y_k$ we obtain $\ell x = x$. As the action is free, it follows that $\ell = e$. However the equality

$$\phi(h_k^{-1}g_k, x_k) = \phi(e, y_k)$$

would (for large k) contradict the local injectivity of ϕ at (e, x). We conclude that $\phi : G \times V \to M$ is a diffeomorphism onto its image, which is an open G-invariant subset of M containing the orbit of x. Since ϕ takes the obvious G-action (given by left multiplication on the first factor) to the G-action on M, this diffeomorphism provides an equivalence between the standard G-action on $G \times S$ and the action of G on a neighborhood of the orbit $G \cdot x$.

[5]This way of deriving a property at all points of an orbit from the validity at a single point is called an homogeneity argument. It will be used often in the sequel and we will not be repeating this argument of conjugating by diffeomorphisms over and over again.

In particular, since $G\backslash(G \times V) = V$, we have that the composite

$$V \xrightarrow{\phi_{|\{e\}\times V}} M \xrightarrow{\pi} G\backslash M$$

is a homeomorphism onto its image, which is an open neigborhood of the orbit $G \cdot x$ in the orbit space.

We will use the inverses of these homeomorphisms for local charts on $G\backslash M$ (assuming without loss of generality that the open sets $V \subset S$ are diffeomorphic to open subsets in some euclidean space). Let $\psi : G \times W \to M$ be an action map giving an equivariant diffeomorphism onto a neighborhood of an orbit $G \cdot y$, and $W \xrightarrow{\psi_{|\{e\}\times W}} M \to G\backslash M$ be the associated local parametrization. Then the change of coordinates determined by the charts associated to ϕ and ψ is the composite

$$V \cap \psi(G \times W) \hookrightarrow \psi(G \times W) \xleftarrow{\psi} G \times W \xrightarrow{\pi_2} W$$

which is a composition of smooth maps and hence smooth. The projection map $\pi : M \to G\backslash M$ is smooth as, on the open set $\phi(G\times V)$ it is given, in the local coordinates determined by V, by the smooth map by $\pi_2 \circ \phi^{-1}$. Moreover π has local sections determined by the inclusions of the submanifolds V in M

$$\begin{tikzcd} & M \arrow[d, "\pi"] \\ V \arrow[ur, hook, "s"] \arrow[r, hook, "\cong"] & G\backslash M \end{tikzcd}$$

To see that the smooth structure constructed above is unique, consider the diagram

$$\begin{tikzcd} & M \arrow[d, "\pi"] \arrow[dr, "\pi"] & \\ U \arrow[ur, hook, "s"] \arrow[r, hook] & G\backslash M \arrow[r, "\mathrm{id}"] & G\backslash M \end{tikzcd}$$

where the two copies of the orbit space $G\backslash M$ are given smooth structures for which π has sections. The identity maps between the two copies of $G\backslash M$ are smooth because their restriction to the domains of local sections can be written as the composite of two smooth maps (the sections followed by the projections). It follows that the identity map is a diffeomorphism and hence the two smooth structures are equivalent. $\square$

Example 5.11. *An important example of the previous Theorem is the case of the quotient of a Lie group G by a closed subgroup H. The orbit space $H\backslash G$ is called a* homogeneous space *since it has a transitive action of G and so "all points in the space look the same".*

Note that

$$d\pi_{|e} \colon \mathfrak{g} \to T_H(H\backslash G)$$

has kernel $\mathfrak{h}$*. Indeed* $d\pi_e$ *is nothing other than the infinitesimal action map of the (right) action of* G *on* $H\backslash G$*, which clearly has kernel* $\mathfrak{h} \subset G$*. It follows that there is a canonical isomorphism*

$$T_H(H\backslash G) = \mathfrak{g}/\mathfrak{h}.$$

Note however that there is no canonical identification of $T_{Hg}(H\backslash G)$ *with* $\mathfrak{g}/\mathfrak{h}$ *for* $g \neq e$ *as there is no canonical way of identifying this space with* $T_H(H\backslash G)$*.*

Remark 5.12. *(a) Of course there is nothing special about left vs right actions. If* G *acts properly, smoothly and freely on* M *on the right then we have likewise a smooth structure on* M/G*.*

(b) Suppose H *is a closed subgroup of* G*. Then the inverse map* $\iota\colon G \to G$ *induces a bijection* $H\backslash G \to G/H$ *and it is easy to see that with respect to the smooth structures on the quotients provided by Theorem 5.10, this map is a diffeomorphism. One just needs to consider the following diagram*

$$\begin{tikzcd} & G \arrow[r, "\iota"] \arrow[d, "\pi"] & G \arrow[d, "\pi"] \\ V \arrow[ur] \arrow[r, hook] & H\backslash G \arrow[r, "\bar{\iota}"] & G/H \end{tikzcd}$$

where $\bar{\iota}(Hg) = g^{-1}H$*.*

When H *is a normal subgroup, the multiplication* $\mu\colon G \times G \to G$ *also induces a map on the quotients, and consideration of local sections again shows this is smooth:*

$$\begin{tikzcd} & G \times G \arrow[r, "\mu"] \arrow[d, "\pi\times\pi"] & G \arrow[d, "\pi"] \\ V \times W \arrow[ur] \arrow[r, hook] & G/H \times G/H \arrow[r, "\bar{\mu}"] & G/H \end{tikzcd}$$

giving G/H *a canonical Lie group structure.*

We can now use Example 5.11 to give the orbits of any smooth action a differential structure.

Proposition 5.13. *Let* $G \times M \to M$ *be a smooth (not necessarily proper) action. Given* $x \in M$*, the canonical map*

$$\begin{aligned} G/G_x &\xrightarrow{\phi} M \\ gG_x &\mapsto g \cdot x \end{aligned}$$

is an injective immersion and an initial submanifold.

Proof. The map ϕ is injective as

$$g \cdot x = g' \cdot x \Leftrightarrow g'^{-1}g \in G_x \Leftrightarrow g \in g'G_x \Leftrightarrow gG_x = g'G_x$$

It is smooth by the usual argument: the composition of ϕ with the projection from G is smooth by assumption and this is sufficient. In order to check that ϕ is an immersion it is enough, by homogeneity, to see that $d\phi$ is injective at the point G_x.

According to Example 5.11 we just need to check that the kernel of the infinitesimal action map of G on M is the Lie algebra $\mathfrak{g}_x$ of the isotropy group G_x. Given $X \in \mathfrak{g}$ we have

$$d\phi(X + \mathfrak{g}_x) = \frac{d}{dt}(\exp(tX) \cdot x)$$

and hence

$$d\phi(X + \mathfrak{g}_x) = 0 \Leftrightarrow a(X)_x = 0 \Leftrightarrow \exp(tX) \cdot x = x \text{ for all } t \Leftrightarrow X \in \mathfrak{g}_x$$

We leave the proof that ϕ is an initial submanifold for the hoework as a (not very easy) exercise. $\square$

It is worth pointing out that the smooth structure we just gave to the orbits of a smooth action does not depend on the choice of a point in the orbit. If x and $y = gx$ are points in the same orbit then their isotropy subgroups are conjugate in G:

$$h(gx) = gx \Leftrightarrow g^{-1}hgx = x \Leftrightarrow g^{-1}hg \in G_x \Leftrightarrow h \in gG_xg^{-1}$$

so

$$G_{gx} = gG_xg^{-1}$$

Moreover, clearly all the conjugate subgroups of G_x will occur as isotropy groups of some point in the orbit. For that reason one refers to a conjugacy class of closed subgroups of a Lie group G as an *orbit type*. There is a canonical bijection φ relating the submanifolds determined by the actions of G on x and y

$$\begin{tikzcd} G/G_x \arrow[rr, hook] \arrow[dr, "\varphi"] & & M \\ & G/G_y \arrow[ur, hook] & \end{tikzcd}$$

defined by the expression

$$\varphi(hG_x) = hg^{-1}G_x$$

To see this, write down the equation for when points in the homogeneous spaces map to the same point in M:

$$hx = kgx \Leftrightarrow hG_x = kgG_x \Leftrightarrow h^{-1}kg \in G_x \Leftrightarrow k \in hg^{-1}G_y$$

(where in the last equality we used that $G_y = gG_xg^{-1}$). The map φ is smooth because G/G_y is an initial submanifold and then φ^{-1} is also smooth for the same reason, so that φ is a diffeomorphism. The map φ is also clearly G-equivariant. Alternatively, note that φ is the map induced on the quotients by the diffeomorphism $R_{g^{-1}}$ of G.

Remark 5.14. *If G acts transitively on M, then the canonical map $G/G_x \to M$ is a diffeomorphism. Indeed, the inclusion of an orbit is always an immersion. If it were not also a submersion that would mean that $\dim G/G_x < \dim M$ which would preclude the surjectivity of the smooth map $G/G_x \to M$.*

Example 5.11 allows us to give many sets a canonical smooth structure: we just need to provide a transitive action of a Lie group on the set and check that the isotropy group is closed. Often these sets will have natural manifold structures, but then the action we give will almost certainly be smooth in which case Remark 5.14 says that the smooth structure on the set agrees with the smooth structure on the corresponding homogeneous space.

Example 5.15. *(i) The group $SO(n+1)$ acts on $S^n \subset \mathbb{R}^{n+1}$ smoothly (it is the restriction to a submanifold of the action of $\mathrm{GL}(n+1;\mathbb{R})$ on $\mathbb{R}^{n+1}$) and transitively. The isotropy subgroup of $(1,0,\dots,0)$ is the subgroup $SO(n) \subset SO(n+1)$ rotating the $0 \times \mathbb{R}^n \subset \mathbb{R}^{n+1}$. From Remark 5.14 it follows that*

$$S^n \cong SO(n+1)/SO(n)$$

(ii) The Stiefel manifolds *of k-frames in $\mathbb{R}^n$ (or $\mathbb{C}^n$, or $\mathbb{H}^n$) are*

$$\tilde{V}_{k,n} = \{A \in M_{n\times k} \colon \operatorname{rank} A = k\}.$$

$GL(n)$ acts smoothly and transitively on $\tilde{V}_{n,k}$ by matrix multiplication. The isotropy group of the standard k-frame

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

is the semi-direct product

$$GL(n-k) \ltimes M_{k\times(n-k)} = \left\{ \begin{bmatrix} 1 & \dots & 0 & \\ \vdots & \dots & \vdots & B \\ 0 & \dots & 1 & \\ 0 & 0 & 0 & \\ \vdots & \vdots & \vdots & A \\ 0 & 0 & 0 & \end{bmatrix} : A \in GL(n-k), B \in M_{k\times(n-k)} \right\}$$

Remark 5.14 again gives us a diffeomorphism

$$\tilde{V}_{k,n} = \mathrm{GL}(n)/(\mathrm{GL}(n-k) \ltimes M_{k\times(n-k)})$$

We can also consider the Stiefel manifolds of orthonormal k-frames

$$V_{k,n} = \{A \in M_{n\times k} \colon A^T A = I\}$$

These are submanifolds of the Euclidean space $M_{n\times k}$ by the argument in Example 1.3(v). The orthogonal group $O(n)$ acts transitively on $V_{k,n}$ with isotropy group

$$\big(\mathrm{GL}(n-k) \ltimes M_{k\times(n-k)}\big) \cap O(n) = O(n-k)$$

where $O(n-k)$ is the subgroup of $O(n)$ rotating $0\times\mathbb{R}^{n-k} \subset \mathbb{R}^n$. Indeed, an orthogonal matrix in $\mathrm{GL}(n-k) \ltimes M_{k\times(n-k)}$ must have $B = 0$ and then A must be orthogonal. We conclude from Remark 5.14 that

$$V_{k,n} \cong O(n)/O(n-k)$$

If we don't want to bother proving that $V_{k,n}$ is a smooth manifold, as $V_{k,n}$ has a transitive action of $O(n)$ with isotropy group a closed subgroup, Example 5.11 gives $V_{k,n}$ a canonical smooth manifold structure as a homogeneous space.

(iii) The Grassmann manifolds of k-planes in $\mathbb{R}^n$ can be described as sets

$$\mathrm{Gr}_{k,n} = \tilde{V}_{k,n}/\,\mathrm{GL}(k) \cong V_{k,n}/O(k)$$

where the groups by which we are quotienting act on the (orthonormal) frames by matrix multiplication. Note that the two quotient sets in the formula above are indeed identical, since any plane has an orthonormal frame and two such yield the same plane if and only if they differ by an orthogonal transformation. The set $\mathrm{Gr}_{k,n}$ has a natural smooth structure coming from the fact that it is a homogeneous space.

Indeed, the group $\mathrm{GL}(n)$ acts transitively on the set of k-planes and the isotropy of $\mathbb{R}^k \times 0 \subset \mathbb{R}^n$ is the subgroup

$$(\mathrm{GL}(k)\times\mathrm{GL}(n-k))\ltimes M_{k\times(n-k)} = \left\{ \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} : A \in \mathrm{GL}(k), D \in \mathrm{GL}(n-k), B \in M_{k\times(n-k)} \right\}$$

which is a closed subgroup of $GL(n)$. Alternatively, the action of $\mathrm{GL}(k)$ on $\tilde{V}_{k,n}$ is smooth (it is given by matrix multiplication) and free and one can check that it is also proper. Theorem 5.10 then gives $\mathrm{Gr}_{k,n}$ a smooth structure. Since the action of $\mathrm{GL}(n)$ on $\mathrm{Gr}_{k,n}$ comes from a smooth action on $\tilde{V}_{k,n}$, the action is smooth with respect to this smooth structure given by Theorem 5.10 and then Remark 5.14 implies that the smooth structure given by Theorem 5.10 is equivalent to the homogeneous space structure.

Alternatively we can regard $\mathrm{Gr}_{k,n}$ as the homogeneous space

$$\mathrm{Gr}_{k,n} = O(n)/(O(k)\times O(n-k))$$

or as the quotient by the smooth, proper, free action

$$\mathrm{Gr}_{k,n} = V_{k,n}/O(k)$$

It is an exercise in the next homework to check (in a more general situation) that these smooth structures agree with the ones discussed above.

Finally one can also give a smooth structure to $\mathrm{Gr}_{k,n}$ directly: any plane near a given k-plane $P \subset \mathbb{R}^n$ is the graph of a unique linear map $P \to P^\perp$ and this gives a chart around P taking values in $\mathbb{R}^{k(n-k)}$. It is a good exercise to write down the changes of coordinates for these charts using linear algebra (see the homework). From

Remark 5.14 it follows that this smooth structure is the same as those discussed above as soon as we check that the action of $\mathrm{GL}(n)$ *(or* $O(n)$*) on this manifold is smooth.*

(iv) *The set of complex structures on* $\mathbb{R}^{2n}$ *is*

$$\mathcal{J} = \{J\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}\colon J^2 = -\operatorname{Id}\}.$$

These endomorphisms J *give* $\mathbb{R}^{2n}$ *the structure of a complex vector space via*

$$(a+bi)\cdot v = av + bJv$$

Conversely, multiplication by i *for such a structure on* $\mathbb{R}^{2n}$ *gives a (real) endomorphism of* $\mathbb{R}^{2n}$ *squaring to* $-\operatorname{Id}$ *so* $\mathcal{J}$ *is precisely the set of complex vector space structures on* $\mathbb{R}^{2n}$.

Since all n*-dimensional complex vector spaces are isomorphic, given* $J_1, J_2 \in \mathcal{J}$ *there exists an isomorphism* $A\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ *(of real vector spaces) such that*

$$AJ_1 = J_2A \Leftrightarrow J_2 = AJ_1A^{-1}.$$

This shows that the action of $GL(2n,\mathbb{R})$ *on* $\mathcal{J}$ *given by* $A\cdot J = AJ_1A^{-1}$ *is transitive. The isotropy for this action on the standard complex structure*

$$\begin{bmatrix} 0 & -\operatorname{Id} \\ \operatorname{Id} & 0 \end{bmatrix}$$

is precisely the subgroup of complex linear maps, which is isomorphic to $\mathrm{GL}(n;\mathbb{C})$. *For instance, when* $n=1$, *the isotropy group is*

$$\left\{\begin{bmatrix} a & -b \\ b & a \end{bmatrix} : a^2+b^2 \neq 0\right\} \cong \mathrm{GL}(1,\mathbb{C})$$

We conclude that the set of complex structures on a real vector space has a canonical smooth structure as the homogeneous space

$$\mathcal{J} = \mathrm{GL}(2n,\mathbb{R})/\,\mathrm{GL}(n,\mathbb{C}).$$

Again we can show directly that the set $\mathcal{J}$ *is a submanifold of* $M_{2n\times 2n}(\mathbb{R})$ *and then it will follow from Remark 5.14 that this smooth structure will agree with the homogeneous one.*

There are many more examples in the spirit of the ones aboves. Essentially any of the familiar classification theorems of linear algebra (quadratic forms, hermitean inner products, ...) can be framed as a statement identifying a set of structures as a homogeneous space (and then Example 5.11 provides these spaces with a canonical smooth structure).

6. Invariant integration

Definition 6.1. *Let* G *be a Lie group. A* k*-form* $\omega \in \Omega^k(G)$ *is said to be* left invariant *if* $L_g^*\omega = \omega$ *for all* $g \in G$. *It is said to be* right invariant *if* $R_g^*\omega = \omega$ *for all* $g \in G$. *The vector space of left invariant* k*-forms on* G *is denoted by* $\Omega_l^k(G)$ *and the vector space of right invariant* k*-forms is denoted by* $\Omega_r^k(G)$.

Just as for invariant vector fields we have the following result, the proof of which is left as an exercise.

Proposition 6.2. *Evaluation at $e \in G$ gives vector space isomorphisms*

$$\Omega^k_l(G) \to \Lambda^k(\mathfrak{g}^*) \qquad \Omega^k_r(G) \to \Lambda^k(\mathfrak{g}^*)$$

where $\Lambda^k(\mathfrak{g}^*)$ denotes the vector space of k-multilinear alternating maps $\mathfrak{g} \times \cdots \times \mathfrak{g} \to \mathbb{R}$.

In particular, for $k = \dim G$, a volume form on $\mathfrak{g}$ gives a left invariant volume form $\omega \in \Omega^k_l(G)$, which is unique up to a scalar. If we fix an orientation of G, a left invariant form determining this orientation is unique up to a positive scalar.

Given a function $f\colon G \to \mathbb{R}$ (or more generally given a function with values in some finite dimensional vector space) we can define its integral with respect to the invariant volume form by

$$\int_G f \overset{\text{def}}{=} \int_G f\omega$$

where, on the right, G is given the orientation determined by ω. This integral is, by construction, left invariant, in the sense that the integral of f and $f \circ L_h$ are the same for all $h \in G$:

$$\int_G f \circ L_h = \int_G f \circ L_h \ \omega = \int_G f \circ L_h \ L_h^*\omega = \int_G L_h^*(f\omega) = \int_G f\omega = \int_G f$$

where in the second equality we used the left invariance of ω and, in the fourth, the fact that L_h^* is a diffeomorphism which preserves the orientation on G (by definition of the orientation) together with the change of variables formula.

The usual notation for the integral $\int_G f$ is

$$\int_G f(g)dg$$

With this notation, the previous computation would appear in the following way:

$$\int_G f(hg)dg = \int_G f(k)\overbrace{d(h^{-1}k)}^{=dk} = \int_G f(k)dk$$

where we have used the change of variable $k = hg \Leftrightarrow g = h^{-1}k$ and the equality over the brace is true because $\omega = dk$ is left invariant.

Remark 6.3. *If G is a locally compact (Hausdorff) topological group, there exists a unique up to multiplicative constant left invariant measure on G (satisfying certain regularity properties) called the* Haar measure*.*

We also need to understand the behaviour of the integral defined above under the diffeomorphisms given by right multiplication. Note that for any $g \in G$, the form $R_g^*\omega$ is still left invariant because $L_hR_g = R_gL_h$;

$$L_h^*R_g^*\omega = R_g^*L_h^*\omega = R_g^*\omega.$$

Since left invariant forms are determined up to a scalar multiple this means that there exists a function

$$\lambda\colon G \to \mathbb{R}^\times$$

such that

$$R_g^*\omega = \lambda(g)\omega \quad \text{for all } g \in G$$

One readily checks that λ is independent of the choice of the left invariant form ω appearing in its definition.

Proposition 6.4. *The function $\lambda\colon G \to \mathbb{R}^\times$ is a Lie group homomorphism.*

Proof. We have

$$R_{(gh)}^*\omega = (R_h \circ R_g)^*\omega = R_g^* R_h^*\omega = R_g^*(\lambda(h)\omega) = \lambda(h)\mathbb{R}_g^*\omega = \lambda(h)\lambda(g)\omega$$

therefore $\lambda(gh) = \lambda(g)\lambda(h)$. The proof that λ is smooth is left as an exercise. Alternatively,

$$\lambda(g)\omega_e = R_g^*\omega_g = R_g^*(L_{g^{-1}}^*\omega_e) = \mathrm{Ad}(g^{-1})^*\omega_e$$

so that multiplication by $\lambda(g)$ is the map induced on the volume elements $\Lambda^n(\mathfrak{g}^*)$ by the endomorphism $\mathrm{Ad}(g^{-1})$ of $\mathfrak{g}$. It follows that

$$\lambda(g) = \det \mathrm{Ad}(g^{-1})$$

□

The homomorphism λ can be seen as the obstruction to the existence of a bi-invariant (i.e. both left and right invariant) volume form on a Lie group G. That is the content of the following result.

Proposition 6.5. *Let $h \in G$ and $f\colon G \to \mathbb{R}$ be an integrable function. Then*

$$\int_G f \circ R_h = \frac{1}{|\lambda(h)|}\int_G f$$

Proof. We have

$$\int_G f \circ R_h\omega = \int_G f \circ R_h \frac{R_h^*\omega}{\lambda(h)} = \frac{1}{\lambda(h)}\int_G R_h^*(f\omega)$$

There are now two possiblities: either the diffeomorphism R_h preserves the orientation on G or it doesn't. Whether this happens or not is controlled by the sign of $\lambda(h)$, so applying the change of variables formula we have that the integral above is equal to

$$\frac{1}{\lambda(h)}\,\mathrm{sgn}(\lambda(h))\int_G f\omega = \frac{1}{|\lambda(h)|}\int_G f$$

□

Definition 6.6. *The Lie group homomorphism $G \to \mathbb{R}^+$ defined by $g \mapsto \frac{1}{|\lambda(g)|} = |\det(\mathrm{Ad}(g))|$ is called* modular function *of the Lie group G.*

The only compact subgroup of the multiplicative group $\mathbb{R}^+$ is the trivial subgroup, so if G is a compact Lie group, the modular function must be constant equal to 1, which is to say that *a left invariant volume form on a compact Lie group is automatically right invariant.* Thus a bi-invariant integral on a compact Lie group is completely determined by the choice of an orientation for G together with the standard normalisation

$$\int_G 1 = 1$$

which we will always assume from now on.

The existence of a bi-invariant integral on a compact Lie group has many important consequences as we will see throughout the course. We will point out a couple by way of example.

Proposition 6.7. *Let G be a compact Lie group and $\rho : G \to GL(V)$ be a representation. Then, there exists a G-invariant inner product on V, i.e. a positive definite symmetric bilinear form $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ such that*

$$\langle \rho(g)v, \rho(g)w\rangle = \langle v, w\rangle \quad \text{for all } g \in G, v, w \in V.$$

Proof. Take any inner product $b : V \times V \to \mathbb{R}$ and define

$$\langle v, w\rangle = \int_G b(\rho(g)v, \rho(g)w)$$

Then,

$$\langle \rho(h)v, \rho(h)w\rangle = \int_G b(\rho(g)\rho(h)v, \rho(g)\rho(h)w)dg = \int_G b(\rho(gh)v, \rho(gh)w)dg$$

Writing $gh = k$, right invariance tells us that $dg = dk$ and hence

$$\int_G b(\rho(gh)v, \rho(gh)w)dg = \int_G b(\rho(k)v, \rho(k)w)dk = \langle v, w\rangle$$

□

The previous basic observation ensures that the representation theory of compact Lie groups is "as simple as possible".

Definition 6.8. *A representation V of a Lie group G (or a Lie algebra $\mathfrak{g}$ or anything really) is* irreducible *if the only subspaces of V which are invariant under the action are V and 0.*

The irreducible representations can be thought of as the "atoms" out of which all representations are built. The easiest case in representation theory is when any representation can be written as a direct sum of irreducible representations. In that case one says that the representations are *completely reducible.* Then representation theory boils down to understanding the irreducible representations as any representation factors (uniquely it turns out) as a sum of such, similarly to how a natural number factors as a product of prime numbers.

This does not always happen as the following simple example shows:

Example 6.9. *Let G be the group of invertible upper triangular 2×2 real matrices.*

$$G = \left\{ \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \,:\, ac \neq 0 \right\}$$

The standard (defining) representation $V = \mathbb{R}^2$ is not completely reducible as the subspace

$$W = \{(x, 0) \colon x \in \mathbb{R}\} \subset V$$

is invariant and non-trivial but V cannot be written as the sum of W and a complementary invariant subspace.

However, Proposition 6.7 immediately implies that representations of compact Lie groups *are* completely reducible.

Proposition 6.10. *If G is a compact Lie group, then any representation $\rho : G \to GL(V)$ is a direct sum of irreducible representations.*

Proof. Let $\langle \ , \ \rangle$ be an invariant inner product on V. Given an invariant subspace $W \subset V$, so that $\rho(g)W \subseteq W$ for all g, then the orthogonal complement of W with respect to this inner product is also invariant: given $v \in W^\perp$ and $w \in W$ we have

$$\langle \rho(g)v, w\rangle = \langle v, \rho(g^{-1})w\rangle = 0$$

as $\rho(g^{-1})w \in W$. Hence V can be decomposed as the direct sum

$$V = W \oplus W^\perp$$

of invariant subspaces and, inductively, we see that V can be written as a direct sum of irreducible representations. ☐

Here is another important consequence of invariant integration.

Theorem 6.11 (Böchner Linearization Theorem)**.** *Let G be a compact Lie group acting smoothly on M and $x \in M$ be a fixed point. Given any neighborhood W of x, there exists an open invariant neighborhood U of x contained in W and a diffeomorphism*

$$\Psi : U \to V \subset T_xM$$

onto a neighborhood V of 0 in T_xM such that

- $\Psi(x) = 0$*;*
- $d\Psi_x : T_xM \to T_xM$ *is the identity;*
- $\Psi(g \cdot x) = g \cdot \Psi(x)$ *(where the action on the right is the isotropy action of G on T_xM).*

Proof. Let W be a neighborhood of x. Since G is compact, the set $U' = \bigcap_{g\in G} g \cdot W$ is open. This is clear in the case when G is finite and an easy point set topology exercise in general[6]. Clearly U' contains x and is contained in W as one of the sets in the intersection is W itself. Moreover, U' is G-invariant since $hU' = \cap_{g\in G} hgW = \cap_{k\in G} kW = U'$ for any $h \in G$.

[6]This is a manifestation of the general principle in point set topology that compact sets behave like finite sets.

Thus any open set containing a fixed point contains a G-invariant open neighborhood of the fixed point.

We may assume that W is contained in a coordinate chart centered at x and use that to pick a diffeomorphism $\phi\colon W \to T_xM$ such that $\phi(x) = 0$ and $d\phi_x = \mathrm{Id}$ (in the local coordinates on W and the corresponding coordinates on T_xM we can set ϕ to be the identity). This diffeomorphism satisfies all the required conditions except equivariance which would be

$$\phi(gy) = g\phi(y) \Leftrightarrow \phi(y) = g^{-1}\phi(gy).$$

We can fix this problem by averaging over the group G: let $U' = \bigcap_{g\in G} W$ and define

$$\psi\colon U' \to T_xM$$

by

$$\psi(y) = \int_G g^{-1}\phi(gy)dg.$$

Let us check that all conditions in the statement are satisfied:

- $\psi(x) = \int_G g^{-1} \overbrace{\phi(\underbrace{g\cdot x}_{x})}^{0} dg = \int_G g^{-1} 0 dg = 0$
- $d\psi_x(v) = \frac{d}{dt}(\psi(c(t)))_{|t=0} = \frac{d}{dt}\left(\int_G g^{-1}\phi(g\cdot c(t))dg\right)_{|t=0}$, where $c\colon]-\epsilon,\epsilon[\to M$ is a smooth curve with $c(0) = x$ and $\frac{dc}{dt}(0) = v$. By the Leibniz rule the latter expression is equal to
$$\int_G \frac{d}{dt}\left(g^{-1}\phi(g\cdot c(t))\right)_{|t=0} dg$$
Since G acts on T_xM linearly this is
$$\int_G g^{-1}\frac{d}{dt}\left(\phi(g\cdot c(t))\right)_{|t=0} dg = \int_G g^{-1} d\phi_{g\cdot c(0)}\left(\frac{d}{dt}(g\cdot c(t))_{|t=0}\right) dg$$
But $\frac{d}{dt}(g\cdot c(t))_{|t=0} = g\cdot v$ by definition of the isotropy representation on T_xM, and since $d\phi_x = \mathrm{Id}$ the above expression equals
$$\int_G g^{-1}\,\mathrm{Id}(g\cdot v)dg = \int_G v\ dg = v$$
Hence $d\psi_x = \mathrm{Id}$.
- Finally, since ψ was obtained by averaging, it is invariant: given $h \in G$,
$$\psi(hy) = \int_G g^{-1}\phi(ghy)dg = \int_G hk^{-1}\phi(ky)dk = h\int_G k^{-1}\phi(ky)dk = h\psi(y)$$

Since $d\psi_x = \mathrm{Id}$, there is a neighborhood $W' \subset U'$ of x such that $\psi_{|W'}$ is a diffeomorphism. Let $U = \bigcap_{g\in G} W'$ and set $\Psi = \psi_{|U}$. □

Remark 6.12. *The Böchner linearization Theorem may also be proved by invoking Riemannian geometry: averaging an arbitrary metric on M over G produces a G-invariant metric on X and then the Riemannian exponential map from T_xM to a neighborhood of x will produce a G-equivariant diffeomorphism defined in some neighborhood of 0 in T_xM.*

7. The slice theorem

Definition 7.1. *Let $a : G \times M \to M$ be a smooth action. A* slice *for the action at a point $x \in M$ is an embedded submanifold $S \subset M$ satisfying*

(i) $x \in S$;
(ii) S *is* G_x*-invariant;*
(iii) $T_x M = T_x S \oplus a(\mathfrak{g})_x$;
(iv) $T_y M = T_y S + a(\mathfrak{g})_y$, *for all* $y \in S$;
(v) *Given* $g \in G$, $y \in S$ *such that* $g \cdot y \in S$, *then* $g \in G_x$.

Let us see what these conditions mean.

$(i) + (iv)$ say that the action map $\varphi\colon G \times S \to M$, defined by $\phi(g, y) = gy$ is a G-equivariant submersion onto an open invariant neighborhood of the orbit Gx. Indeed (for the obvious action of G on $G \times S$) we have

$$\varphi(h(g, y)) = \varphi(hg, y) = hgy = h\varphi(g, y)$$

Condition (i) says that the orbit Gx is contained in the image of φ, while condition (iv) says that

$$d\varphi_{(e,y)} : \mathfrak{g} \times T_y S \to T_y M,$$

which is given by the expression

$$d\varphi_{(e,y)}(X, v) = a(X)_y + v, \tag{4}$$

is surjective. Thus φ is submersive at the points of the form (e, y) with $y \in S$, but then homogeneity implies it is submersive at all (g, y). Since φ is a submersion its image is open and since φ is equivariant, its image is a G-invariant set.

$(ii) + (v)$ identify the fibers of φ with the orbits of a free action of G_x on $G \times S$: indeed

$$\varphi(g, y) = \varphi(g', y') \Leftrightarrow g \cdot y = g' \cdot y' \Leftrightarrow {g'}^{-1} gy = y'$$

and by (v) this means that $g'^{-1} g \in G_x$. If we define an action of G_x on $G \times S$ by the formula

$$h \cdot (g, y) = (gh^{-1}, h \cdot y),$$

which makes sense by (ii), and is free because G_x is a subgroup of G, then

$$\varphi(g, y) = \varphi(g', y') \Leftrightarrow (g, y) \in G_x(g', y').$$

Now, G_x is a closed subgroup of G so the action is proper and therefore we have a smooth structure on $G_x \backslash (G \times S)$. The usual notation for this quotient (which is analogous to a tensor product) is

$$G \times_{G_x} S$$

and this is the notation we will use from now on. Getting back to our story, we see that the map φ passes to the quotient and induces a smooth *bijection* $\bar{\varphi}\colon G \times_{G_x} S \to$

$\varphi(G \times S) \subset M$:

$$\begin{tikzcd} G \times S \arrow[d, "\pi"] \arrow[dr, "\varphi"] & \\ G \times_{G_x} S \arrow[r, dashed, "\bar{\varphi}"'] & M \end{tikzcd}$$

The manifold $G \times_{G_x} S$ has a natural G-action given by the expression $g \cdot [(h, y)] = [(gh, y)]$ for which the projection π and hence $\bar{\varphi}$ are equivariant.

(*iii*) says that $\bar{\varphi}$ is a local diffeomorphism at $[(e, x)]$: Indeed, the proof of Theorem 5.10 identifies the tangent space to the quotient by a free proper Lie group $H \backslash N$ at an orbit Hz with the quotient $(T_z N)/(a(\mathfrak{h})_z)$ of the tangent space by the infinitesimal action. In our situation we have

$$T_{[(e,x)]} G \times_{G_x} S = (\mathfrak{g} \oplus T_x S)/a(\mathfrak{g}_x)_{(e,x)} \cong \mathfrak{g}/\mathfrak{g}_x \oplus T_x S$$

The expression (4) for $d\varphi_{(e,x)}$ then shows that $d\bar{\varphi}_{[(e,x)]}$ induces an isomorphism of this space with

$$T_x M = a(\mathfrak{g})_x \oplus T_x S$$

hence $\bar{\varphi}$ is a local diffeomorphism at $[(e, x)] = G_x(e, x)$.

Now by the inverse function theorem, $\bar{\varphi}$ is also a local diffeomorphism at $G_x(g, y)$ for y sufficiently close to x, and then by homogeneity, $\bar{\varphi}$ is a local diffeomorphism in a neighborhood of the G-orbit of $G_x(e, x)$. However we have already seen that $\bar{\varphi}$ is a bijection hence it will be a G-equivariant diffeomorphism from a neighborhood of $G_x(e, x)$ in $G \times_{G_x} S$ to a neighborhood of the orbit Gx in M.

The upshot is that a slice gives a local model for the action near the orbit which is completely determined by the isotropy action on the slice. If we can pick a small G_x-invariant neighborhood W of x in S, for instance, if G_x is compact then we get a simple model for the neighborhood

$$\bar{\varphi} : G \times_{G_x} W \xrightarrow{\cong} U \subset M,$$

In this case the neighborhood U is called a *tube* around the orbit. It is a G-equivariant tubular neighbourhood of the orbit.

To see a sample consequence of this formula for the action near an orbit in terms of the G_x action on a neighborhood W of x in the slice, note that $\bar{\varphi}$ identifies a neighbourhood of Gx in the topological space $G \backslash M$ (it may not be a manifold) with $G \backslash (G \times_{G_x} W) \cong G_x \backslash W$ which is the orbit space of the action of the isotropy group on W.

Example 7.2. *Consider the action $S^1 \times S^3 \to S^3 \subset \mathbb{C}^2$ given by the expression*

$$e^{i\theta}(z, w) = (e^{i\theta} z, e^{3i\theta} w)$$

and let $x = (0, 1) \in S^3$. The isotropy group of x is

$$G_x = \{e^{i\theta} \in S^1 \colon e^{i\theta}(0, 1) = (0, e^{3i\theta}) = (0, 1)\} = \{1, \omega, \omega^2\} \subset S^1, \qquad \textit{with } \omega = e^{2\pi i/3}$$

Let us check that

$$S = \{(z, \sqrt{1 - |z|^2}) \ : \ |z| < 1\} \subset S^3$$

is a slice:

(i) $(0,1) \in S$

(ii) S *is* G_x*-invariant:*

$$\omega \cdot (z, \sqrt{1-|z|^2}) = (\omega z, \omega^3\sqrt{1-|z|^2}) = (\omega z, \sqrt{1-|z|^2}) \in S \text{ as } |\omega| = 1$$

(iii) $T_{(0,1)}S^3 = T_{(0,1)}S \oplus a(\mathrm{Lie}(S^1))_{(0,1)}$*: We have that* $T_{(0,1)}S^3 = \mathbb{C} \oplus i\mathbb{R} \subset \mathbb{C}^2$ *and* $T_{(0,1)}S = \mathbb{C} \oplus 0$*. Considering the generator* $\frac{d}{d\theta}$ *of* $\mathrm{Lie}(S^1)$ *we have*

$$a\left(\frac{d}{d\theta}\right)_{(0,1)} = \frac{d}{dt}\left(e^{it}\cdot(0,1)\right)_{|t=0} = \frac{d}{dt}(0, e^{3it})_{|t=0} = (0, 3i)$$

so $a(\mathrm{Lie}(S^1))_{(0,1)} = 0 \oplus i\mathbb{R} \subset \mathbb{C} \oplus i\mathbb{R}$ *and the condition above is satisfied.*

(iv) $T_{(z,\sqrt{1-|z|^2})}S^3 = T_{(z,\sqrt{1-|z|^2})}S \oplus a(\mathrm{Lie}(S^1))_{(z,\sqrt{1-|z|^2})}$*: this is left as an exercise.*

(v) Suppose $e^{i\theta}\cdot(z,\sqrt{1-|z|^2}) = (w,\sqrt{1-|w|^2})$ *for some* z, w *with* $|z| < 1, |w| < 1$*. Then* $w = e^{i\theta}z$*, so* $|w|^2 = |z|^2$*. Hence* $e^{3i\theta}\sqrt{1-|z|^2} = \sqrt{1-|z|^2}$ *and therefore* $e^{3i\theta} = 1 \Leftrightarrow e^{i\theta} \in G_x$*.*

This gives us the following model for a neighborhood of the orbit. First note that the action of G_x *on the slice is equivariantly diffeomorphic (via the projection* $\pi_1\colon \mathbb{C}^2 \to \mathbb{C}$*) to the linear action by rotations of* $G_x \cong \mathbb{Z}/3$ *on the unit disk* $D \subset \mathbb{C}$*.*

Then a neighborhood of the orbit $S^1\cdot(0,1) \cong S^1/(\mathbb{Z}/3)$ *is* S^1 *equivariantly diffeomorphic to*

$$S^1 \times_{\mathbb{Z}/3} D$$

Here we have over each point in the central orbit $(S^1)/(\mathbb{Z}/3)$ *a unit disk. All orbits other than the central orbit intersect this disk at three distinct points at the same distance from the center of the disk and spaced by* $\frac{2\pi}{3}$ *angles. As these orbits approach the central orbit they wrap around it three times.*

This is in fact a model for any sufficiently small neighborhood of an S^1 *orbit with isotropy* $\mathbb{Z}/3$ *and isotropy representation* $\mathbb{C}_{std} \oplus \mathbb{R}_{triv}$ *on an* arbitrary 3*-manifold (see Remark 7.5 below.)*

Theorem 7.3 (Slice Theorem)**.** *Let* $a : G \times M \to M$ *be a smooth and proper action of a Lie group on a manifold. Then there exists a slice through every point* $x \in M$*.*

Proof. Let x be a point in M. Since the action is proper, the isotropy group G_x is compact and hence by the Böchner Linearization theorem 6.11 there exists a G_x-invariant neighbourhood U of x and a G_x-equivariant diffeomorphism

$$\psi : U \to V \subset T_xM$$

with $\psi(x) = 0$, $d\psi_x = \mathrm{Id}$. Pick a G_x-invariant metric on T_xM and let A be the orthogonal complement to $a(\mathfrak{g})_x \subset T_xM$ (with respect to the chosen G_x-invariant metric).

Note that $a(\mathfrak{g})_x \subset T_xM$ is G_x-invariant (for the isotropy action) as

$$\begin{aligned} g\cdot a(X)_x &= \frac{d}{dt}(g\cdot\exp(tX)\cdot x)_{|t=0} = \frac{d}{dt}(g\cdot\exp(tX)g^{-1}\cdot gx)_{|t=0} \\ &= \frac{d}{dt}(\exp(t\,\mathrm{Ad}(g)X)\cdot x)_{|t=0} = a(\mathrm{Ad}(g)X)_x \end{aligned}$$

therefore its orthogonal complement A is also G_x-invariant. Moreover, since $d\psi_x = \mathrm{Id}$, we have that

$$a(\mathfrak{g})_x = d\psi_x(T_x(G\cdot x)).$$

Define a proto-slice

$$\widetilde{S} = \psi^{-1}(A\cap V)$$

This is a submanifold of M (it's the image under a diffeomorphism of an open set in a plane), it is G_x-invariant (because $A \cap V$ is and ψ^{-1} is G_x-equivariant) and clearly $x = \psi^{-1}(0)$ is in $\widetilde{S}$. Since $d\psi_x = \mathrm{Id}$ we have $T_x\widetilde{S} = A$ and therefore $\widetilde{S}$ satisfies conditions $(i) - (iii)$ in Definition 7.1.

Condition (iv) says that the map $G\times\widetilde{S} \to M$ is submersive at (e, y). We know that this is true when $y = x$, as this is condition (iii). Since being submersive is an open condition, we can find a G_x-invariant neighbourhood $S' \subset \widetilde{S}$ of x in $\widetilde{S}$ such that $(i) - (iv)$ hold.

We are now going to show that S' contains a neighbourhood W of x such that condition (v) holds, i. e., such that whenever $g \in G, y \in W$ and $g\cdot y \in W$ we have $g \in G_x$. If this is not true then we can find sequences $g_k \in G$, $y_k \in S$ such that

- y_k converges to x
- $g_k \cdot y_k$ converges to x and $g_k \cdot y_k \in S'$
- $g_k \notin G_x$.

Since the action is proper there is a subsequence of g_k converging to an element $g \in G$. Let us assume for simplicity that g_k itself converges. Then,

$$x = \lim_{k\to\infty} g_k y_k = g\cdot x$$

so $g \in G_x$. Replacing g_k with $g_k g^{-1}$ and y_k with gy_k, we may assume $g_k \to e \in G$.

Let B be a complement to $\mathfrak{g}_x$ in $\mathfrak{g}$ and consider the local diffeomorphism at 0

$$B\times\mathfrak{g}_x \to G$$

defined by

$$(Z, H) \mapsto \exp(Z)\exp(H).$$

For sufficiently large k we can write $g_k = \exp(Z_k)\cdot\exp(H_k)$ for unique Z_k, H_k which are moreover converging to 0. Replacing g_k with $g_k \exp(-H_k)$ and y_k with $\exp(H_k)y_k$ we may further assume the H_k are 0.

Note that the map $B \to a(\mathfrak{g})_x \cong \mathfrak{g}/(\mathfrak{g}_x) \subset T_xM$ given by $Z \mapsto a(Z)_x$ is an isomorphism. Now consider the map

$$f\colon B\times S' \to M$$

defined by

$$(Z, y) \mapsto \exp(Z)\cdot y$$

On the one hand we have

$$f(Z_k, y_k) = f(0, \exp(Z_k)y_k) \tag{5}$$

with $Z_k \neq 0$. Note that (5) makes sense because by assumption $\exp(Z_k)y_k \in S'$. Moreover, since the action is continuous and Z_k is converging to 0, the sequence $\exp(Z_k)y_k$ is converging to x. On the other hand the derivative of f at $(x, 0)$ is given by the formula

$$df_{(0,x)}(Z, v) = df_{(0,x)}((Z, 0) + (0, v)) = a(Z)_x + v,$$

and is therefore an isomorphism by condition (iii). Hence f is a local diffeomorphism at $(0, x)$, which contradicts (5) for large k. $\square$

Remark 7.4. *In the case when G is compact, the slice Theorem can be proved using Riemannian geometry (cf. Remark 6.12). Picking a G-invariant metric on M, the Riemannian exponential in the orthogonal directions to an orbit provide a tube around the orbit. However, this will not work if G is not compact.*

Remark 7.5. *The construction of the slice shows that it is equivariantly diffeomorphic to a neighbourhood of 0 in the representation of G_x on $(T_xM)/a(\mathfrak{g})_x$ induced by the isotropy representation of G_x on T_xM. Therefore, a neighbourhood of $G \cdot x$ in M is G-equivariantly diffeomorphic to*

$$G \times_{G_x} (T_xM/\mathfrak{g}_x)$$

which is a (G-equivariant) vector bundle over the orbit G/G_x.

An important consequence of the slice theorem is that if y is close to x in M then the isotropy group G_y is subconjugate to G_x, meaning there exists $g \in G$ such that $gG_yg^1 \subset G_x$. Indeed, if y is close enough to x, then there exists $g \in G$ so that $g \cdot y$ is in the slice S through x. Then $G_y = g^{-1}G_{gy}g$ and G_{gy} is a subgroup of G_x by condition (v) in Definition 7.1. This says that the isotropy group in a smooth proper action varies "upper semi-continuously" with the point of the manifold: nearby isotropy groups are either the same or smaller.

Recall that two orbits are said to be of the same *type* if the conjugacy class formed by the isotropy groups of points in the orbits are the same (equivalently, if the orbits are G-equivariantly diffeomorphic to G/H for the same closed subgroup H). It can be shown that the union of all the orbits of a given type in a manifold M form a submanifold. Thus there is a canonical partition of M into submanifolds $M_{(H)}$, one for each conjugacy class (H) of closed subgroups of G. Moreover this decomposition can be shown to have very nice properties - it is a Whitney stratification of M. Amongst all the orbit types, there is one whose isotropy groups are the smallest. The corresponding orbits are called *principal orbits* for the action. They always form an open set in M. For instance in Example 7.2 the stratification is

$$S^3 = \left(S^3\right)_{(\{e\})} \coprod \left(S^3\right)_{(\mathbb{Z}/3)} = S^3 \setminus \{(0,1)\} \coprod \{(0,1)\}$$

and the principal orbits are free. See [DK] for all this and much more. For more on actions [Br] is highly recommended (but assumes the reader is familiar with Lie theory).

8. Review of linear algebra

We have already studied the correspondence between Lie groups and Lie algebras and studied some applications, notably to group actions. It is time to delve into the structure of these objects. We will start by taking an algebraic approach and examining Lie algebras as purely linear algebraic objects. Before we start we will review some needed Linear Algebra. There is a basic dichotomy in Linear Algebra between nilpotent and semisimple endomorphisms of a vector space which will play an important role in our study of the structure of Lie Algebras.

Definition 8.1. *Let V be a finite dimensional vector space and* $\mathrm{End}(V) = \mathrm{Hom}(V, V)$ *be the space of linear maps from V to itself.*

An element $x \in \mathrm{End}(V)$ is said to be nilpotent *if $x^k = 0$ for some $k \in \mathbb{N}$. The least k such that $x^k = 0$ is called the* degree of nilpotence *of x.*

An element $x \in \mathrm{End}(V)$ is said to be semisimple *if given a subspace $W \subset V$ invariant under x (i. e. such that $xW \subset W$) there is a subspace $W' \subset V$ with $W \oplus W' = V$ which is also invariant under x.*

Proposition 8.2. *An element $x \in \mathrm{End}(V)$ is nilpotent iff V has a basis with respect to which x is represented by a strictly upper triangular matrix.*

Proof. The condition is certainly sufficient as an $n \times n$ strictly upper triangular matrix is nilpotent of degree $\leq n$.

Conversely, if x is nilpotent we have $\det x = 0$ and so there is $v \in \ker x \setminus \{0\}$. Let $W \subset V$ be such that $V = \langle v \rangle \oplus W$. Then the matrix representing x in a basis adapted to this decomposition of V is of the form

$$x = \begin{bmatrix} 0 & * \\ 0 & \bar{x} \end{bmatrix}$$

with $\bar{x} : V/\langle v \rangle \to V/\langle v \rangle$ the endomorphism induced by x on the quotient (which is canonically isomorphic to W). Since $\bar{x}$ is again nilpotent with nilpotence degree less than or equal to that of x and $\dim W < \dim V$ we may inductively assume there is a basis for W with respect to which $\bar{x}$ is strictly upper triangular. □

Our next goal is to characterize semisimple endomorphisms of vector spaces over an algebraically closed field $\mathbb{K}$. First we need to review the Jordan normal form of such an endomorphism $x \colon V \to V$. Recall that V can be written as a direct sum $V = W_1 \oplus \cdots \oplus W_n$ of generalized eigenspaces

$$W_i = \{v \in V \colon (x - \lambda_i)^k v = 0 \text{ for some } k\}$$

where the λ_i are the eigenvalues of x and we are assuming that $\lambda_i \neq \lambda_j$.

Let $\pi_i \colon V \to W_i$ denote the projections onto the generalized eigenspaces. The fact that x is block diagonal with respect to the generalized eigenspace decomposition implies that $\pi_i x = x \pi_i$. But in fact, the projections π_i are polynomials with zero constant term on x, i. e. there exist $p_i(t) \in \mathbb{K}[t]$ such that $\pi_i = p_i(x)$. To see this let

$$q_i(t) = \prod_{j \neq i} (t - \lambda_j)^{k_j}$$

with k_j the size of the largest Jordan block corresponding to λ_j. Since the polynomials $q_1, \ldots, q_n$ are coprime there exist polynomials $h_i(t)$ such that

$$h_1(t)q_1(t) + \cdots + h_n(t)q_n(t) = 1$$

But then $h_i(x)q_i(x)$ is the projection from V to W_i. Indeed, if $v \in W_j$, with $j \neq i$ then $q_i(x)v = 0$, while, if $v \in W_i$, then since

$$h_1(x)q_1(x) + \cdots + h_n(x)q_n(x) = \mathrm{Id}_V$$

and $q_j(x)v = 0$ for $j \neq i$, we have

$$0 + \cdots + 0 + h_i(x)q_i(x)v + 0 + \cdots + 0 = v.$$

Proposition 8.3. *Let V be a vector space over an algebraically closed field $\mathbb{K}$. Then $x \in \mathrm{End}(V)$ is semisimple iff it is diagonalizable.*

Proof. Assume that x is semisimple and let $v \in V$ be an eigenvector for x (which exists because $\mathbb{K}$ is algebraically closed). Since x is semisimple we may write $V = \mathbb{K}v \oplus W$ with $xW \subset W$. As $x_{|W}$ has an eigenvector we may continue in this way until we obtain a basis of eigenvectors for V.

Conversely, suppose x is diagonalizable and let $V = W_1 \oplus \cdots \oplus W_n$ be the spectral decomposition of x. Suppose $W \subset V$ is an invariant subspace of x. Then $(W \cap W_1) \oplus \cdots \oplus (W \cap W_n) \subset W$. But, in fact it is equal to W: given $w \in W$ we saw that there is a polynomial $p_i(x)$ with no constant term so that $p_i(x)w = w_i$ the component of w along w_i. Since $xW \subset W$, $p_i(x)W \subset W$, so $w_i = p_i(x)w \in W \cap W_i$. Writing, $w = w_1 + \ldots + w_n$, we see that

$$w \in W \cap W_1 + \cdots + W \cap W_n.$$

Now, take for each i a complement W_i' to $W \cap W_i$ in W_i. Then,

$$W_1' \oplus \cdots \oplus W_n'$$

is an invariant complement to W. □

Exercise 8.4. *Let V be a finite dimensional real vector space. Show that $x \in \mathrm{End}(V)$ is semisimple iff $x \otimes \mathbb{C}$ is diagonalizable. One can show more generally that x is semisimple if and only if the minimal polynomial of x has no repeated irreducible factors (see* [HK]*).*

Theorem 8.5 (Jordan-Chevalley decomposition.)**.** *Let V be a vector space over an algebraically closed field $\mathbb{K}$ and $x \in \mathrm{End}(V)$. Then there are $x_s, x_n \in \mathrm{End}(V)$ with x_s semisimple and x_n nilpotent such that*

(i) $x = x_s + x_n$
(ii) x_s, x_n are polynomials with no constant terms in x. In particular, if $xW \subset W'$ with $W' \subset W$, then $x_sW, x_nW \subset W'$. Moreover, x_s and x_n commute and they also commute with any polynomial in x.

Given any other decomposition $x = x_s' + x_n'$ with x_s' semisimple and x_n' nilpotent such that x_s' and x_n' commute, then $x_s = x_s'$ and $x_n = x_n'$.

Proof. Let $p(x) = (x - \lambda_1)^{k_1} \dots (x - \lambda_n)^{k_n}$ be the minimal polynomial of x. By the chinese remainder theorem, there exists $q(x)$ such that $q(x) \equiv \lambda_i \mod (x-\lambda_i)^{k_i}$, and also $q(x) \equiv 0 \mod x$ if all the λ_i are different from 0. Letting $V = W_1 \oplus \cdots \oplus W_n$ be the decomposition of x into generalized eigenspaces this means that given $v_i \in W_i$, there exists a polynomial $r_i(x)$ such that

$$q(x)v_i = (\lambda_i + (x - \lambda_i)^{k_i} r_i(x))v_i = \lambda_i v_i + 0 = \lambda_i v_i$$

Hence $q(x)$ is diagonalized by the decomposition $V = W_1 \oplus \cdots \oplus W_n$. Take $x_s = q(x)$, $x_n = x - q(x)$ to obtain the required decomposition.

Now suppose $x = x'_s + x'_n$ is another decomposition of x into a sum of a semisimple and nilpotent elements such that x'_s, x'_n commute. Then, x'_s, x'_n commute with x and therefore with x_s and x_n. The element

$$x_s - x'_s = x_n - x'_n$$

is both semisimple (because x_s and x'_s are simultaneously diagonalizable) and nilpotent (as x_n, x'_n commute, if $x_n^l = 0$ and $x'^j_n = 0$ then by the binomial theorem we have $(x_n - x'_n)^{l+j} = 0$). Since the only diagonalizable element which is also nilpotent is 0 we have that $x'_s - x_s = 0$ and then if follows that $x_n - x'_n = 0$. □

9. Basic notions of Lie algebras

Definition 9.1. *Let L be a Lie algebra over a field $\mathbb{K}$. A subspace $I \subset L$ is said to be an* ideal *if $[x, I] \subset I$ for all $x \in L$, i. e. if $[x, y] \in I$ for all $x \in L$ and $y \in I$.*

For instance, the *center* of L,

$$Z(L) = \{x \in L \colon [x, L] = 0\}$$

is easily checked to be an ideal using the Jacobi identity: given $y \in Z(L)$ and $x, z \in L$ we have

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

Recall that $Z(L) = \ker(\mathrm{ad} \colon L \to \mathrm{End}(L))$ and therefore, when $Z(L) = 0$, L is a linear Lie algebra. In general there is a canonical short exact sequence of vector spaces

$$0 \to Z(L) \to L \xrightarrow{\mathrm{ad}} \mathrm{ad}(L) \to 0 \tag{6}$$

where the maps are in fact maps of Lie algebras. This is called an *extension* of the Lie algebra $\mathrm{ad}(L)$ by $Z(L)$.

Definition 9.2. *Let L be a Lie algebra. The subspace of L generated by all the brackets of elements of L, $[L, L] \subset L$ is called the* derived Lie algebra *of L.*

Again it follows immediately from the Jacobi identity that $[L, L]$ is an ideal of L.

Definition 9.3. *A Lie algebra L is said to be* simple *if its only ideals are 0 and L and L is not abelian.*

Note that an algebra is abelian iff $Z(L) = L$ if and only if $[L, L] = 0$. In particular if L is simple then $Z(L) = 0$ (hence L is linear) and $[L, L] = L$.

Let $f : L_1 \to L_2$ be a Lie algebra homomorphism. Then $\ker(f) = f^{-1}(0)$ is an ideal. More generally, if $I \subset L_2$ is an ideal then $f^{-1}(I)$ is an ideal. If I is an ideal, then the bracket $[\ ,\ ]$ is well defined on the quotient vector space

$$\begin{aligned}[\ ,\ ]\colon L/I \times L/I &\to L/I \\ (x + I, y + I) &\mapsto [x, y] + I,\end{aligned}$$

so L/I becomes a Lie algebra called the quotient Lie algebra and

$$\begin{aligned}\pi : L &\to L/I \\ x &\mapsto x + I\end{aligned}$$

is a Lie algebra homomorphism with kernel I. Thus ideals are exactly the kernels of Lie algebra homomorphisms. As usual we have the following isomorphism theorems, whose proofs are left as an exercise.

Theorem 9.4 (Isomorphism Theorems). *(1) Given a Lie algebra homomorphism*

$$f : L_1 \to L_2$$

there exists a unique homomorphism $\bar{f} : L_1/\ker f \to L_2$ *such that*

$$\begin{tikzcd} L_1 \arrow[r, "f"] \arrow[d, "\pi"] & L_2 \\ L_1/\ker f \arrow[ur, dashed, "\bar{f}"'] & \end{tikzcd} \qquad \bar{f}(x + \ker f) := f(X)$$

(2) Given ideals $I \subset J \subset L$ of the Lie algebra L, we have that J/I is an ideal of L/I and

$$(L/I)/(J/I) = L/J$$

(where = means that the isomorphism between the two quotients is natural). The isomorphism is defined by

$$(x + I) + (J/I) \mapsto x + J$$

(3) Given ideals $I \subset J \subset L$ of the Lie algebra L we have that $I \cap J$ is an ideal of J and

$$J/(I \cap J) = (I + J)/I$$

with the canonical isomorphism given by

$$x + I \cap J \mapsto x + I$$

If I is an ideal of L we have an extension of L/I by I

$$0 \to I \hookrightarrow L \xrightarrow{\pi} L/I \to 0$$

and in particular we have, for any Lie algebra L two canonical extensions (6) and

$$0 \to [L, L] \to L \to L/[L, L] \to 0$$

Note that $L/[L, L]$ is abelian, and in fact the *maximal abelian quotient* of L, i.e. an abelian quotient through which any map from L to an abelian Lie algebra factors.

Associated to any Lie algebra L we have the following two canonical descending sequences of ideals:

- The *derived series of* L (obtained by iteratively taking the derived Lie algebra):
$$L = L^{(0)} \supset L^{(1)} = [L, L] \supset L^{(2)} = [L^{(1)}, L^{(1)}] \supset \cdots$$
- The *descending or lower central series of* L
$$L = L^0 \supset L^1 = [L, L] \supset L^2 = [L, L^1] \supset L^3 = [L, L^2] \supset \dots$$

Note that $L^{(i)} \subset L^i$. One immediately checks that all the L^i and $L^{(i)}$ are ideals using the Jacobi identity. We now come to two basic definitions concerning the structure of a Lie algebra.

Definition 9.5. *A Lie algebra L is* solvable *if $L^{(i)} = 0$ for some i. A Lie algebra is* nilpotent *if $L^i = 0$ for some i.*

Note that if L is nilpotent then it is also solvable. Concretly, a Lie algebra is nilpotent if there exists n such that all brackets

$$[x_1, [x_2, [\dots, x_n]]]$$

of length n vanish. The basic examples of solvable and nilpotent Lie algebras are the following.

Example 9.6. *Let $\mathbb{K}$ be a field and $L = \mathfrak{gl}(n, \mathbb{K}) = \mathrm{End}(\mathbb{K}^n)$. The subspace $\mathfrak{t}(n)$ of upper triangular matrices form a solvable subalgebra of L and the subspace $\mathfrak{n}(n)$ of strictly upper triangular matrices form a nilpotent subalgebra of L. Indeed the bracket two upper triangular matrices which are zero below a certain diagonal has all entries zero below a "higher diagonal".*

Remark 9.7. *If L is nilpotent, the Campbell-Baker-Hausdorff formula gives the vector space L the structure of an (algebraic) group. When the field in question is $\mathbb{R}$ or $\mathbb{C}$, this is a Lie group with Lie algebra L which is the simply connected Lie group corresponding to L. The homogeneous spaces of the latter Lie groups are important examples in Riemannian geometry called* nilmanifolds.

Exercise 9.8. *Let $\mathbb{K}$ be a field. The Lie algebra $\mathfrak{sl}(2; \mathbb{K})$ of traceless 2×2 matrices over $\mathbb{K}$ is nilpotent is the characteristic of $\mathbb{K}$ is two and simple otherwise.*

Proposition 9.9. *Let L be a Lie algebra, then*

(i) If L is solvable (nilpotent) then any subalgebra and homomorphic image is solvable (nilpotent).
(ii) If $I \subset L$ is a solvable ideal and L/I is solvable, then L is solvable.
(iii) If $I, J \subset L$ are solvable ideals, then $I + J$ is solvable.

(iv) If $L/Z(L)$ is nilpotent, then L is nilpotent.
(v) If L is nilpotent and $L \neq 0$, then $Z(L) \neq 0$.

Proof. (1) This is clear from the definitions.

(2) Saying that $(L/I)^{(k)} = 0$ means that $L^{(k)} \subset I$. If $I^{(\ell)} = 0$ it follows that $(L^{(k)})^{(\ell)} = L^{(k+\ell)} \subset I^{(\ell)} = 0$ hence L is solvable.

(3) Consider the extension

$$0 \to I \to I + J \to (I + J)/I \cong J/(I \cap J) \to 0,$$

where we have used one of the isomorphism theorems. By (i), we know that $J/(I \cap J)$ is solvable as it is a homomorphic image of the solvable Lie algebra J. It follows from (ii) that $I + J$ is solvable.

(4) Note that the degree of nilpotence of L is n if and only if $L^{n-1} \neq 0$ and $L^{n-1} \subset Z(L)$. Therefore, if $L/Z(L)$ is nilpotent, then $(L/Z(L))^k = 0$ for some k and hence $L^{k+1} = 0$.

(5) if L is nilpotent and $L \neq 0$, then $Z(L) \neq 0$ as it contains the last nonzero term of the lower central series of L.

□

Note that (i) and (ii) in the previous Proposition can be rephrased as saying that given an extension

$$0 \to I \to L \to L/I \to 0$$

then L is solvable if and only if both I and L/I are solvable.

Because of (iii) every finite dimensional Lie algebra contains a maximal solvable ideal (the sum of all its solvable ideals)

Definition 9.10. *The* radical of L *is the maximal solvable ideal of* L *and is denoted* $\operatorname{Rad} L$. *A Lie algebra is said to be* semisimple *if* $\operatorname{Rad} L = 0$.

Recall that a Lie algebra L is simple if its only ideals are 0 and L and L is not abelian. If L is simple, then $\operatorname{Rad} L$ must be 0 or L. If it were L then L would be solvable and hence $[L, L] \subsetneq L$ would necessarily be 0, which would mean that L is abelian. We conclude that a simple Lie algebra is semisimple.

The radical of a Lie algebra gives us a new canonical realization of L as an extension

$$0 \to \operatorname{Rad} L \to L \to L/\operatorname{Rad} L \to 0$$

It is a simple exercise to check that $L/\operatorname{Rad} L$ is semisimple so this is expressing L canonically as an extension of a semisimple Lie algebra by a solvable subalgebra. We will see that this extension plays a fundamental role in the large scale structure of Lie algebras.

10. Engel's and Lie's Theorems

We now come to two of the basic Theorems about Lie algebras which characterize nilpotent and solvable subalgebras. The first one - Engel's Theorem - gives a simple linear algebraic criterion to check nilpotence of a Lie algebra.

Theorem 10.1 (Engel's Theorem)**.** *Let L be a Lie algebra (over an arbitrary field $\mathbb{K}$). If for each $x \in L$, the endomorphism $\mathrm{ad}(x) \in \mathrm{End}(L)$ is nilpotent, then L is nilpotent.*

Recall that L being nilpotent means that for some n, all n-fold brackets vanish. In particular we will have $\mathrm{ad}(x)^{n-1} = 0$ so the condition in Engel's Theorem is certainly necessary for nilpotence. The proof of the Theorem will follow easily from the following Linear algebraic statement.

Proposition 10.2. *Let $L \subset \mathfrak{gl}(V)$ be a Lie algebra such that each $x \in L$ is a nilpotent endomorphism. Then there exists $v \in V \setminus \{0\}$ such that*

$$xv = 0 \text{ for all } x \in L.$$

Proof. The proof will be by induction on the dimension of L. If $\dim L = 1$, then $L = \mathbb{K} \cdot x$. Since $x^n = 0$ for some n, $\det x = 0$ and therefore x has a 0-eigenvector v.

Now assume $\dim L > 1$ and that the proposition holds for lower dimensional Lie algebras. Let $K \subset L$ be a maximal proper subalgebra. By the induction hypothesis, there exists $v \in V \backslash \{0\}$ such that $xv = 0$ for all $x \in K$. We start by proving that K must be an ideal of L: Given $x \in L$, the endomorphism

$$\mathrm{ad}(x) \in \mathrm{End}(L) \subset \mathrm{End}(\mathrm{End}(V))$$

is nilpotent as $\mathrm{ad}(x) = L_x - R_x$ with L_x and R_x the left and right multiplication by x on $\mathrm{End}(L)$. Clearly L_x and R_x are nilpotent (because x is) and commute, so their difference $\mathrm{ad}(x)$ is also nilpotent.

Given $y \in K$, the induced map

$$\overline{\mathrm{ad}(y)} \colon L/K \to L/K$$

defined by $z + K \mapsto [y, z] + K$ is nilpotent so by induction hypothesis there exists $x + K \in L/K$ not equal to 0 such that $\overline{\mathrm{ad}(y)}(x + K) = 0$ for all $y \in K$. But this exactly means that $[y, x] \in K$ for all y in K. Therefore the normalizer of K in L

$$N_L(K) = \{x \in L \colon [y, x] \in K \text{ for all } y \in K\},$$

(the largest subalgebra of L containing K as an ideal) strictly contains K. Since K is a maximal proper subalgebra of L we conclude that $N_L(K) = L$ which is to say that K is an ideal of L.

Let's now see that the codimension of K in L must be one. Consider the extension

$$0 \to K \to L \to L/K \to 0.$$

Given $y \in L/K \setminus \{0\}$, then $K + \mathbb{K}y$ is a subalgebra of L containing K. Since K is maximal, we see that $L = K + \mathbb{K}y$. Now consider $W = \{v \in V : xv = 0 \text{ for all } x \in K\}$ and let us prove that W is invariant under y. We need to see check that for any$v \in W$

$$x(yv) = 0 \text{ for all } x \in K$$

Since K is an ideal we have $[x, y] \in K$ and it then follows that

$$x(yv) = [x, y]v + y(xv) = 0 + 0 = 0$$

We conclude that $yW \subset W$. As y acts nilpotently on V, $y_{|W}$ is also nilpotent so there exists $w \in W \setminus \{0\}$ with $yw = 0$. This element w is the required common 0 eigenvector for all the elements of L. $\square$

Definition 10.3. *Let V be a finite dimensional vector space. A* complete flag *in V is a chain $0 = V_0 \subset V_1 \subset \cdots \subset V_n = V$ of subspaces of V with $\dim V_i = i$.*

Proposition 10.2 can equivalently be formulated as follows

Proposition 10.4. *Let $L \subset \mathfrak{gl}(V)$ be a Lie subalgebra consisting of nilpotent endomorphisms of V. Then there exists a complete flag in V such that $LV_i \subset V_{i-1}$ for all i.*

Proof. We will show that this statement is equivalent to that of Proposition 10.2. If there is a complete flag with $LV_i \subset V_{i-1}$ then any vector $v \in V_1 \setminus \{0\}$ is a common 0 eigenvector for all the elements of L.

Conversely, given a common 0-eigenvector v for all the elements of L let $V_1 = \mathbb{K}v$. Consider the action of L on V/V_1 (which makes sense because V_1 is invariant under L). By the previous Proposition, the endomorphisms of V/V_1 determined by the elements L have a common 0 eigenvector $v_2 + \mathbb{K}v_1$. But

$$L(v_2 + \mathbb{K}v_1) = 0 \in V/V_1 \Leftrightarrow Lv_2 \subset V_1 \Leftrightarrow L(V_1 + \mathbb{K}v_2) \subset V_1$$

We can therefore set $V_2 = V_1 + \mathbb{K}v_2$ and proceed in this way until we obtain a complete flag as required. $\square$

Note that if we pick a basis for V adapted to the flag (equivalently an isomorphism of V with $\mathbb{K}^n$ taking the flag to the standard flag in $\mathbb{K}^n$), we have that the elements of L expressed in terms of the basis consist of strictly upper triangular matrices. In particular L is nilpotent.

Remark 10.5. *The Lie group $\mathrm{GL}(n, \mathbb{R})$ acts transitively on complete flags in $\mathbb{R}^n$. The isotropy group of the standard flag in $\mathbb{R}^n$ is the subgroup B of upper triangular matrices. Hence the space of complete flags in $\mathbb{R}^n$ is the homogeneous space $\mathrm{GL}(n, \mathbb{R})/B$.*

Note that one can have a nilpotent subalgebra L of $\mathfrak{gl}(V)$ whose elements are not nilpotent endomorphisms of V. Indeed, any endomorphism $x \in \mathrm{End}(V)$ spans an abelian (hence nilpotent) Lie subalgebra of $\mathfrak{gl}(V)$.

Proof of Engel's theorem. Consider Consider the extension

$$0 \to Z(L) \to L \to \mathrm{ad}(L) \to 0$$

By assumption $\mathrm{ad}(x)$ is nilpotent for every $x \in L$. Hence, Proposition 10.2 tells us that $\mathrm{ad}(L) \subset \mathrm{End}(L)$ is a nilpotent Lie algebra. Since $L/Z(L)$ is nilpotent, Proposition 9.9(iv) tells us that L is nilpotent. $\square$

The second basic recognition Theorem is Lie's Theorem characterising solvable linear Lie algebras. This has the same flavor as Engel's Theorem but requires that the base field be algebraically closed and characteristic zero.

Theorem 10.6 (Lie's theorem). *Let $\mathbb{K}$ be an algebraically closed field of characteristic* 0 *and $L \subset \mathfrak{gl}(V)$ be a solvable Lie algebra. Then, there exists a complete flag in V*

$$0 = V_0 \subset V_1 \subset \cdots \subset V_n = V \qquad \text{with } LV_i \subset V_i,$$

so L is isomorphic to a subalgebra of upper triangular matrices.

As before, the previous statement is clearly equivalent to the following seemingly weaker statement which we will prove instead.

Proposition 10.7. *Let $\mathbb{K}$ be an algebraically closed field of characteristic* 0 *and $L \subset \mathfrak{gl}(V)$ be a solvable Lie algebra. Then all the elements of L have a common eigenvector.*

Note that the previous statement is a big generalization of the familiar statement from Linear Algebra to the effect that a set of commuting endomorphisms over an algebraically closed field have a common eigenvector. Before we prove Proposition 10.7 we point out two important Corollaries of Lie's Theorem.

Corollary 10.8. *Let $\mathbb{K}$ be an algebraically closed field of characteristic* 0*. A Lie algebra L over $\mathbb{K}$ is solvable if and only if $[L, L]$ is nilpotent.*

Proof. If $[L, L]$ is nilpotent, it is also solvable. Since $L/[L, L]$ is abelian, it follows from Proposition 9.9(ii) that L is solvable.

Conversely, suppose L is solvable. Then $\mathrm{ad}(L) \subset \mathrm{End}(L)$ is solvable by Proposition 9.9(i). By Lie's theorem we have that $\mathrm{ad}(L)$ is isomorphic to a Lie algebra of upper triangular matrices. But the $[\mathrm{ad}(L), \mathrm{ad}(L)] = \mathrm{ad}([L, L])$ is isomorphic to a Lie algebra of upper triangular matrices.

Hence $\mathrm{ad}[L, L]$ consists of nilpotent endomorphisms of L. Since $[L, L] \subset L$ is a Lie subalgebra, the subspace $[L, L]$ is invariant under the elements of $\mathrm{ad}[L, L]$ and hence $\mathrm{ad}[L, L] \subset \mathrm{End}([L, L])$ also consists of nilpotent endomorphisms. It follows from Engel's Theorem that $[L, L]$ is nilpotent. □

Corollary 10.9. *Let $\mathbb{K}$ be an algebraically closed field of characteristic zero and L a Lie algebra over $\mathbb{K}$. If L is solvable all irreducible finite dimensional representations of L are one dimensional.*

Proof. Let $\rho \colon L \to \mathfrak{gl}(V)$ be a finite dimensional irreducible representation. Since $\rho(L)$ is solvable, Lie's Theorem guarantees the existence of a common eigenvector $v \neq 0$ for all the endomorphisms $\rho(x)$. Hence $\mathbb{K}v$ is a one dimensional subrepresentation of V. As V is irreducible it must be all of V, so V is one-dimensional. □

Remark 10.10. *Conversely if all irreducible representations are one dimensional L is solvable as you will see in the homework.*

Proof of Prop 10.7. The proof is by induction on the dimension of L. The case when dimension is one is clear using the fact that $\mathbb{K}$ is algebraically closed. Assume that $\dim L > 1$ and that the statement holds for all Lie algebras of dimension less than that of L. Since L is solvable we know that $[L, L] \supsetneq L$. The canonical extension

$$0 \to [L, L] \to L \xrightarrow{\pi} L/[L, L] \to 0$$

therefore implies that L contains a codimension one ideal K. Indeed an ideal of the abelian Lie algebra $L/[L,L]$ is the same as a subspace and the inverse image under π of a codimension one subspace of $L/[L,L]$ will therefore be a codimension one ideal of L.

By induction hypothesis, there is a common eigenvector v for all $x \in K$. If we write

$$xv = \lambda(x)v$$

then one immediately checks that $\lambda\colon K \to \mathbb{K}$ is a linear functional. Let

$$W_\lambda = \{w \in V : xw = \lambda(x)w \text{ for all } x \in K\}$$

Let $y \in L$ be such that $L = K + \mathbb{K}y$. It is enough to check that W_λ is invariant under y, because if this is true then any eigenvector of $y_{|W_\lambda}$ will be a common eigenvector for all of L. Now, given $x \in K$

$$x(yw) = [x,y]w + yxw = \lambda([x,y])w + \lambda(x)yw$$

where we have used that K is an ideal. Hence W_λ will be invariant under y if and only if $\lambda([x,y]) = 0$. Given $w \in W_\lambda \setminus \{0\}$, consider the subspace

$$S = \langle w, yw, y^2w, \ldots \rangle \subset V$$

and let n be the smallest integer such that $y^n w$ is a linear combination of $\langle w, yw, \ldots, y^{n-1}w \rangle$, so that $\{w, yw, \ldots, y^{n-1}w\}$ is a basis for S. Let's check that S is invariant under the action of K: given $x \in K$, we have

$$\begin{aligned} xw &= \lambda(x)w \\ x(yw) &= \lambda([x,y])w + \lambda(x)yw \\ x(y^2w) &= [x,y]yw + yxy(w) = [x,y]yw + y(\lambda[x,y]w + \lambda(x)yw) \end{aligned}$$

Since $[x,y] \in K$, the first term in the sum on the left is in $\langle w, yw \rangle$ so we see that $x(y^2w) \in \langle w, yw, y^2w \rangle$. Inductively the expansion above shows that the subspaces $\langle w, yw, \ldots, y^i w \rangle$ are invariant under K and, moreover, that, with respect to the given basis, an element $x \in K$ acts by an upper triangular matrix with $\lambda(x)$ on the diagonal.

Now as $[x,y] \in K$, we see that the trace of its action on W is $\operatorname{tr}([x,y]) = n\lambda([x,y])$. On the other hand, the trace of a commutator of two endomorphisms is 0 so we conclude that

$$n\lambda([x,y]) = 0$$

As the characteristic of $\mathbb{K}$ is zero, it follows that $\lambda([x,y]) = 0$, which completes the proof. □

Remark 10.11. *The previous proof shows that in Proposition 10.7 it is enough that the characteristic of $\mathbb{K}$ is large enough with respect to the dimension of V.*

11. Cartan's criteria for solvability and semisimplicity

The trace form played an important role in the proof of Lie's Theorem and it will be important from now on in our study of Lie algebras. We will therefore review some of its properties.

Let $\mathbb{K}$ be an arbitrary field. The trace form

$$\mathrm{tr}\colon \mathfrak{gl}(n,\mathbb{K}) \times \mathfrak{gl}(n,\mathbb{K}) \to \mathbb{K}$$

is defined by the expression

$$\mathrm{tr}(XY) = \sum_{i,j=1}^{n} X_{ij}Y_{ji}$$

If $\{e_{ij}\colon 1 \le i,j \le n\}$ is the standard basis for $\mathfrak{gl}(n,\mathbb{K})$ (e_{ij} has 1 as the ij-th entry and all other entries 0), we have

$$\mathrm{tr}(e_{ij}e_{kl}) = \delta_{il}\delta_{jk}$$

where $\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$ is the Kronecker symbol.

With respect to this basis, the trace form is represented by the block diagonal symmetric matrix which has 1 along the diagonal in the entry corresponding to the basis elements e_{ii} and $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ along the pairs of entries corresponding to basis elements $\{e_{ij}, e_{ji}\}$ with $i \neq j$.

Over $\mathbb{R}$, the signature of this symmetric form is therefore $(\frac{n(n+1)}{2}, \frac{n(n-1)}{2})$.

Proposition 11.1. *Let $\mathbb{K}$ be a field and $A, X, Y \in \mathfrak{gl}(n,\mathbb{K})$, $g \in \mathrm{GL}(n,\mathbb{K})$. Then*

(i) tr *is a non-degenerate symmetric bilinear form.*
(ii) (Conjugation invariance) $\mathrm{tr}(gXg^{-1}gYg^{-1}) = \mathrm{tr}(XY)$
(iii) (Associativity with respect to the commutator bracket) $\mathrm{tr}([A,X]Y) = \mathrm{tr}(A[X,Y]) = 0$

Proof.]

(i) The expression for the trace makes it clear that it a symmetric bilinear form. Since the associated symmetric matrix computed above is clearly non-singular the trace form is non-degenerate.
(ii) This follows from symmetry as $\mathrm{tr}(gXYg^{-1}) = \mathrm{tr}(g^{-1}gXY)$.
(iii) Over $\mathbb{R}$ or $\mathbb{C}$ this follows from *(ii)* via differentiation. In general, it follows from symmetry since this implies the cyclic property of the trace $\mathrm{tr}(XAY) = \mathrm{tr}(AYX)$ from which the required equality immediately follows.

□

We can now define a canonical symmetric bilinear form on a Lie algebra which will play a fundamental role in the rest of the course.

Definition 11.2. *Let L be a Lie algebra over the field $\mathbb{K}$. The* Killing form *of L is the symmetric bilinear form*

$$\kappa\colon L \times L \to \mathbb{K}$$

defined by $\kappa(x,y) = \mathrm{tr}(\mathrm{ad}(x)\,\mathrm{ad}(y))$.

Example 11.3. *Consider the Lie algebra* $\mathfrak{sl}(2,\mathbb{K})$ *with basis given by*

$$h = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad x = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

With respect to the ordered basis (h, x, y) *we have that*

$$\operatorname{ad}(h) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \quad \operatorname{ad}(x) = \begin{bmatrix} 0 & 0 & 1 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \operatorname{ad}(y) = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$$

therefore the matrix representing the Killing form κ *with respect to the basis* (h, x, y) *is*

$$\begin{bmatrix} 8 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \end{bmatrix}$$

The following Proposition is a manifestation of how canonical the Killing form is. It states that the Killing form is invariant under all automorphisms of the Lie algebra as well as under all "infinitesimal automorphisms".

Proposition 11.4. *Let* L *be a Lie algebra with Killing form* κ*. Then, for all* $x, y \in L$ *we have*

(i) If $\alpha \in \operatorname{Aut}(L)$ *then* $\kappa(\alpha(x), \alpha(y)) = \kappa(x, y)$*.*
(ii) If $D \in \operatorname{Der}(L)$ *then* $\kappa(Dx, y) + \kappa(x, Dy) = 0$*.*

Proof. (i) Since $\operatorname{ad}(\alpha(x))(y) = [\alpha(x), y] = \alpha([x, \alpha^{-1}(y)]) = \alpha \circ \operatorname{ad}(x) \circ \alpha^{-1}$, the statement follows from the conjugation invariance of the trace.

(ii) First note that

$$[D, \operatorname{ad}(x)] = \operatorname{ad}(Dx) \tag{7}$$

as $D(\operatorname{ad}(x)(y)) - \operatorname{ad}(x)(Dy) = D([x, y]) - [x, Dy] = [Dx, y] + [x, Dy] - [x, Dy] = [x, Dy]$. Hence

$$\begin{aligned} \kappa(Dx, y) + \kappa(x, Dy) &= \operatorname{tr}(\operatorname{ad}(Dx)\operatorname{ad}(y)) + \operatorname{tr}(\operatorname{ad}(x)\operatorname{ad}(Dy)) \\ &= \operatorname{tr}(D\operatorname{ad}(x)\operatorname{ad}(y)) - \operatorname{tr}(\operatorname{ad}(x)D\operatorname{ad}(y)) + \operatorname{tr}(\operatorname{ad}(x)D\operatorname{ad}(y)) - \operatorname{tr}(\operatorname{ad}(x)\operatorname{ad}(y)D) \\ &= \operatorname{tr}(D\operatorname{ad}(x)\operatorname{ad}(y)) - \operatorname{tr}(\operatorname{ad}(x)\operatorname{ad}(y)D) = 0 \end{aligned}$$

□

Remark 11.5. *Note that* (7) *states that the space* $\{\operatorname{ad}(x) \colon x \in L\}$ *is an ideal of the Lie algebra* $\operatorname{Der}(L)$*. Moreover, statement (ii) above implies that the Killing form is* associative with respect to the Lie bracket *as*

$$\kappa(\operatorname{ad}(x)y, z) + \kappa(y, \operatorname{ad}(x)z) = 0 \Leftrightarrow \kappa([y, x], z) = \kappa(y, [x, z])$$

Definition 11.6. *The* radical *(or kernel) of a symmetric bilinear form* $b \colon V \times V \to \mathbb{K}$ *is the set*

$$\operatorname{Rad} b = \{v \in V \colon b(v, \cdot) = 0\}$$

When the symmetric bilinear form is understood, given a subset $S \subset V$, we write

$$S^\perp = \{v \in V : b(v,s) = 0 \text{ for all } s \in S\}$$

and then $\operatorname{Rad} b = V^\perp$.

Proposition 11.7. *Let L be a Lie algebra with Killing form κ.*

(1) $\operatorname{Rad}\kappa$ *is an ideal of L. More generally, if $I \subset L$ is an ideal, then $I^\perp$ is an ideal*[7].

(2) If $I \subset L$ is an ideal, the Killing form of I is $\kappa_I = \kappa_{|I\times I}$.

Proof. (1) Suppose $x \in L, y \in I^\perp$. Then given $z \in I$ we have $\kappa(z,[x,y]) = \kappa([z,x],y) = 0$, hence $I^\perp$ is an ideal.

(2) Let $W \subset L$ be a subspace with $I \oplus W = L$. Then with respect to a basis adapted to this direct sum decomposition, $\operatorname{ad}(x)$ for $x \in I$ is a block upper triangular matrix with 0 as its second diagonal block, i.e. $\operatorname{ad}(x)$ is of the form

$$\operatorname{ad}(x) = \begin{bmatrix} * & * \\ 0 & 0 \end{bmatrix}$$

Therefore $\operatorname{tr}(\operatorname{ad}_I(x)\operatorname{ad}_I(x')) = \operatorname{tr}(\operatorname{ad}(x)\operatorname{ad}(x'))$ for all $x, x' \in I$, as required. □

We can now state Cartan's basic criteria for solvability and semisimplicity which once again affirm this basic dichotomy in the study of Lie algebras.

Theorem 11.8 (Cartan's criteria)**.** *Let L be a finite dimensional Lie algebra over a field $\mathbb{K}$ of characteristic zero, and κ be the Killing form of L. Then*

(1) L is solvable if and only if $\operatorname{Rad}\kappa \supset [L,L]$.

(2) L is semisimple if and only if κ is non-degenerate (i.e. $\operatorname{Rad}\kappa = \{0\}$*).*

Proof that (1)⇒(2). Assume first that L is semisimple and let $S = \operatorname{Rad}\kappa$. Then S is an ideal and by Proposition11.7(ii) we have that $\kappa_S = 0$. Cartan's criterion (1) implies that S is solvable. Since 0 is the only solvable ideal of L we conclude that $S = 0$, hence κ is nondegenerate.

Now assume that κ is nondegenerate. In order to show that L is semisimple, it is sufficient to show that L contains no nontrivial abelian ideals (as a nontrivial solvable ideal necessarily contains a nontrivial abelian ideal, namely the last nonzero term in its derived series). Suppose that $I \subset L$ is an abelian ideal and let $x \in I, y \in L$. Then the composite

$$L \xrightarrow{\operatorname{ad}(x)} L \xrightarrow{\operatorname{ad}(y)} L \xrightarrow{\operatorname{ad}(x)} L \xrightarrow{\operatorname{ad}(y)} L$$

is the 0 map. Indeed, the image of the first map is contained in I (as $x \in I$) and hence the composition of the first two maps also has image contained in I. Since I is abelian it follows that the composition of the first three maps is already 0. This means in particular that $\operatorname{ad}(y)\operatorname{ad}(x)$ is nilpotent and hence

$$\kappa(y,x) = \operatorname{tr}(\operatorname{ad}(y)\operatorname{ad}(x)) = 0 \text{ for all } y \in L$$

[7] But beware that $I \cap I^\perp$ may be different from $\{0\}$.

Since κ is nondegenerate we see that $I = \{0\}$, as required. □

The proof of Cartan's criterion for solvability depends on a Linear Algebra lemma. Before stating it we will make a remark concerning the Jordan-Chevalley decomposition of an endomorphism $x \in \mathrm{End}(V)$ (with V a vector space over an algebraically closed field $\mathbb{K}$)

$$x = x_s + x_n$$

Let $\{v_1, \ldots, v_n\}$ be a basis of eigenvectors for x_s with $x_s v_i = \lambda_i v_i$, and let e_{ij} be the standard basis for $\mathrm{End}(V)$ defined by $e_{ij}(v_k) = \delta_{jk} v_i$. Then

$$(\mathrm{ad}(x_s)e_{ij})(v_k) = x_s e_{ij} v_k - e_{ij} x_s v_k = x_s \delta_{jk} v_i - e_{ij} \lambda_k v_k = (\lambda_i - \lambda_k)\delta_{jk} v_i = (\lambda_i - \lambda_k) e_{ij}(v_k)$$

Therefore the e_{ij} are eigenvectors for $\mathrm{ad}(x_s)$ and hence $\mathrm{ad}(x_s)$ is a semisimple endomorphism of L. Now, since $\mathrm{ad}(x_n) = L_{x_n} - R_{x_n}$ is the difference of commuting, nilpotent endomorphisms of L, we have that $\mathrm{ad}(x_n)$ is nilpotent. Moreover, $[\mathrm{ad}(x_n), \mathrm{ad}(x_s)] = \mathrm{ad}([x_n, x_s]) = \mathrm{ad}(0) = 0$ so that $\mathrm{ad}(x_s)$ and $\mathrm{ad}(x_n)$ commute. It follows from uniqueness that

$$\mathrm{ad}(x) = \mathrm{ad}(x_s) + \mathrm{ad}(x_n)$$

is the Jordan-Chevalley decomposition of $\mathrm{ad}(x) \in \mathrm{End}(\mathrm{End}(V))$.

Lemma 11.9. *Let V be a vector space over an algebraically closed field of characteristic zero and $L \subset \mathrm{End}(V)$ be a Lie subalgebra such that $\mathrm{tr}(xy) = 0$ for all $x, y \in L$. Then $[L, L]$ is nilpotent.*

Proof. By Engel's Theorem, it is enough to show that $\mathrm{ad}(x)$ is nilpotent for every $x \in [L, L]$, and since $[L, L]$ is an ideal, it is enough to show that $\mathrm{ad}(x)$ are nilpotent as endomorphisms of L. Consider the Jordan-Chevalley decomposition $x = x_s + x_n$, and write $x_s = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Since the Jordan-Chevalley decomposition of $\mathrm{ad}(x)$ is

$$\mathrm{ad}(x) = \mathrm{ad}(x_s) + \mathrm{ad}(x_n)$$

it is enough to show that $x_s = 0$, or equivalently that all the eigenvalues of x_s are zero. We will do this by testing x_s against suitable endomorphisms y using the trace form. Given a linear functional $f \colon \mathbb{Q}\langle \lambda_1, \ldots, \lambda_n \rangle \to \mathbb{Q}$ we can define $y = \mathrm{diag}(f(\lambda_1), \ldots, f(\lambda_n))$ (so that $y =$"$f(x_s)$"). Then

$$\mathrm{ad}(y)(e_{ij}) = (f(\lambda_i) - f(\lambda_j))e_{ij} = f(\lambda_i - \lambda_j)e_{ij}$$

Fix a polynomial $p(x) \in \mathbb{Q}[x]$ such that $p(\lambda_i - \lambda_j) = f(\lambda_i - \lambda_j)$ for all i, j (this exists by Lagrange interpolation and note that in particular $p(0) = 0$). Then

$$\mathrm{ad}(y) = p(\mathrm{ad}(x_s))$$

Since $\mathrm{ad}(x_s) = r(\mathrm{ad}(x))$ for some $r(x)$ with zero constant term, we see that $\mathrm{ad}(y) = q(\mathrm{ad}(x))$ for some polynomial $q(x)$ with zero constant term. It follows that the action of $\mathrm{ad}(y)$ on $\mathrm{End}(V)$ preserves L even thought y itself need not belong to L. Now

$$\mathrm{tr}(xy) = \mathrm{tr}(x_s y) + \mathrm{tr}(x_n y)$$

Since x_n commutes with the projections onto the eigenspaces of x_s, it also commutes with y, therefore $\mathrm{tr}(x_n y) = 0$. We can compute the other term explicitly to get

$$\mathrm{tr}(xy) = \mathrm{tr}(x_s y) = \sum_{i=1}^{n} \lambda_i f(\lambda_i)$$

On the other hand, writing $x = \sum_j [a_j, b_j]$ we have

$$\mathrm{tr}(xy) = \sum_j \mathrm{tr}([a_j, b_j]y) = \sum_j \mathrm{tr}(a_j[b_j, y])$$

Since $[b_j, y] = -\operatorname{ad}(y)(b_j) \in L$, our assumption on the trace form yields $\mathrm{tr}(xy) = 0$. Therefore $\sum_{j=1}^{n} \lambda_j f(\lambda_j) = 0$ for any linear functional on the rational vector space spanned by the eigenvalues. This can only happen if all the λ_i are 0, which concludes the proof. □

Proof of Theorem 11.8 (1). Consider first the case when the base field $\mathbb{K}$ is algebraically closed. If L is solvable then, by Lie's Theorem, $\operatorname{ad}(L)$ may be regarded as a Lie algebra of upper triangular matrices. Hence $\operatorname{ad}[L, L] = [\operatorname{ad}(L), \operatorname{ad}(L)]$ will consist of strictly upper triangular matrices. Therefore, if $x \in [L, L]$ and $y \in L$, we will have

$$\kappa(x, y) = \mathrm{tr}(\operatorname{ad}(x) \operatorname{ad}(y)) = 0$$

which means that $[L, L] \subset \operatorname{Rad} \kappa$.

Conversely, suppose $[L, L] \subset \operatorname{Rad} \kappa$ and apply Lemma 11.9 to the Lie algebra $M = [\operatorname{ad}(L), \operatorname{ad}(L)] = \operatorname{ad}[L, L] \subset \operatorname{End}(L)$. The lemma applies because given $x = \operatorname{ad}(x'), y = \operatorname{ad}(y') \in M$ with $x', y' \in [L, L]$ we have

$$\mathrm{tr}(xy) = \kappa(x', y') = 0$$

It follows that $[M, M]$ is nilpotent. Since $M = \operatorname{ad}(L)^{(1)}$, we see that $\operatorname{ad}(L)$ is solvable and then the extension

$$0 \to Z(L) \to L \xrightarrow{\operatorname{ad}} \operatorname{ad}(L) \to 0$$

shows that L is also solvable as required.

Finally, note that if $\mathbb{K} \to \overline{\mathbb{K}}$ is a field extension, and L is a Lie algebra over $\mathbb{K}$, then

$$\kappa_{L \otimes \overline{\mathbb{K}}} = \kappa_L \otimes \overline{\mathbb{K}} \qquad \operatorname{Rad}(\kappa_{L \otimes \overline{\mathbb{K}}}) = \operatorname{Rad}(\kappa_L) \otimes_{\mathbb{K}} \overline{\mathbb{K}}$$

Therefore if L is a Lie algebra over an arbitrary field $\mathbb{K}$ of characteristic zero the validity of the statement of Cartan's criterion for L follows from the special case of an algebraically closed field proved above, upon consideration of the Lie algebra $L \otimes_{\mathbb{K}} \overline{\mathbb{K}}$ with $\overline{\mathbb{K}}$ the algebraic closure of $\mathbb{K}$. □

Corollary 11.10. *Let L be a semisimple Lie algebra over a field of characteristic zero. Then*

$$L = L_1 \times \cdots \times L_k$$

with L_i simple Lie algebras. In particular $L = [L, L]$.

Proof. It is enough to see that if $I \subset L$ is an ideal, then the Killing form breaks L up as a product of ideals

$$L = I \times I^{\perp},$$

for then, inductively we will get the required decomposition. Since κ is associative, $I^{\perp}$ is also an ideal: indeed, if $x \in I, y \in I^{\perp}$ and $z \in L$ then

$$0 = \kappa([x,z],y) = \kappa(x,[z,y])$$

Moreover, as κ is non-degenerate, $\dim I^{\perp} = \dim L - \dim I$. Now κ vanishes identically on the ideal $I \cap I^{\perp}$, therefore Cartan's criterion for solvability implies that $I \cap I^{\perp}$ is solvable. Since L is semisimple we must have $I \cap I^{\perp} = \{0\}$. Hence $L = I \oplus I^{\perp}$ as a vector space. Since for $x \in I, y \in I^{\perp}$ we have

$$\kappa([x,y],z) = \kappa(x,[y,z]) = 0$$

for every $z \in L$ we see that $[I, I^{\perp}] = 0$ so the Lie bracket structure on $I \oplus I^{\perp}$ is computed coordinatewise.

Finally, since for a simple Lie algebra L_i we have $[L_i, L_i] = L_i$, the same will be true for a cartesian product of such. $\square$

Remark 11.11. *Suppose L is semisimple over a field of characteristic 0 and $\rho : L \to V$ is arepresentation. Then*

$$\operatorname{tr}(\rho(x)) = 0 \quad \textit{for all } x \in L$$

as $\rho(L) = \rho([L,L]) \subset [\mathfrak{gl}(V), \mathfrak{gl}(V)] = \mathfrak{sl}(V)$. In particular a semisimple Lie algebra has no nontrivial one dimensional representations.

Exercise 11.12. *Show that if L is a semisimple Lie algebra over a field of characteristic zero, any ideal $I \subset L$ is a direct sum of orthogonal simple ideals. Conclude that if $f \colon L \to L'$ is Lie algebra homomorphism and L is semisimple then $f(L)$ is semisimple.*

12. Lie modules and cohomology

Let L be a Lie algebra. A representation $\rho\colon L \to \mathfrak{gl}(V)$ gives rise to a bilinear map $L \times V \to V$ defined by the expression

$$(x, v) \mapsto \rho(x)v$$

Writing xv for $\rho(x)v$, the condition that ρ is a map of Lie algebras becomes

$$[x,y]v = x(yv) - y(xv) \tag{8}$$

Conversely a bilinear map satisfying the above relation gives rise to a representation of L, as one immediately checks. Sometimes it is more convenient to think of representations in this way, which leads to the following definition.

Definition 12.1. *A* Lie module *over the Lie algebra L is a vector space V together with a bilinear map $L \times V \to V$ satisfying* (8) *for every $x, y \in L$ and $v \in V$.*

There are several ways of obtaining new L-modules from old. Here are some constructions with L-modules:

- Tensor product: Let V_1, V_2 be L-modules. Their tensor product $V_1 \otimes V_2$ becomes an L-module (check!) extending linearly the action on decomposable tensors given by
$$x(v_1 \otimes v_2) = (xv_1) \otimes v_2 + v_1 \otimes (xv_2)$$
A motivation for this formula comes from the fact that given representations V_1 and V_2 of a Lie group G, the tensor product naturally becomes a representation via the action $g \cdot (v_1 \otimes v_2) = (gv_1) \otimes (gv_2)$. If we differentiate this formula at $g = e$ we obtain the expression defining the action on a tensor product of Lie modules.
- Dual: The dual of the L-module V is the dual vector space V^* with the action of L defined by
$$(x \cdot \phi)(v) = -\phi(xv) \quad \text{for } x \in L, \phi \in V^*, v \in V$$
It is easy to check that the above formula gives V^* an L-module structure. Again the motivation for this formula comes from differentiating the standard action of a Lie group G on the dual of a G-representation V, which is given by
$$(g \cdot \phi)(v) = \phi(g^{-1}v)$$
(the inverse is needed for the resulting action on V^* to be a left action).
- Homomorphisms: If U, V are L-modules, the vector space $\mathrm{Hom}(U, V)$ of linear maps from U to V becomes an L-module via
$$(xf)(u) = x(f(u)) - f(xu) \quad \text{for } x \in L, f \in \mathrm{Hom}(U, V), u \in U \tag{9}$$
Note that, in the special case when $U = V$, the previous formula can alternatively be interpreted as saying that the action of x on f is given by the commutator of the endomorphism of V given by the action of x and f:
$$xf = [x\cdot, f] \in \mathrm{End}(V)$$
The identity (9) is in fact a consequence of the formulas for the action on the dual and tensor product as there is a canonical isomorphism $U^* \otimes V = \mathrm{Hom}(U, V)$ which sends a decomposable tensor $\phi \otimes v$ to the rank one homomorphism defined by
$$u \mapsto \phi(u)v$$
The action on $U^* \otimes V$ is given by $x(\phi \otimes v) = (x\phi) \otimes v + \phi \otimes (xv)$, which is taken via the canonical homomorphism to
$$u \mapsto -\phi(xu)v + \phi(u)xv$$

From the above we can get L-module structures on arbitrary tensors, including symmetric and anti-symmetric tensors. Note also that if $f\colon V_1 \to V_2$ is a map of L-modules (defined in the obvious way), then $\ker f$ and $\operatorname{Im} f$ are again L-modules. In this language an irreducible representation is called a simple module.

Definition 12.2. *Let L be a Lie algebra and V be an L-module. V is said to be* simple *if its only L-submodules are* 0 *and* V.

Our next aim is to prove a basic result concerning semisimple Lie algebras over fields of characteristic zero.

Theorem 12.3 (Weyl)**.** *Let L be a semisimple Lie algebra over a field of characteristic zero. Then all finite dimensional L-modules are direct sums of simple modules.*

This is in contrast with what happens with solvable Lie algebras (see Example 6.9) and in parallel to what happens for compact Lie groups (see Proposition 6.10). This theorem says that the representation theory of semisimple Lie algebras is "as simple as possible": it boils down to a classification of the irreducible representations.

Theorem 12.3 will follow immediately once we prove that every extension of L-modules for such an L

$$0 \to V \xrightarrow{i} W \xrightarrow{\pi} U \to 0 \tag{10}$$

(i.e. a short exact sequence of vector spaces where i and π are maps of L-modules) splits, meaning there exists a map of L-modules $s\colon U \to W$ with $\pi \circ s = \mathrm{id}_U$. In that case the map $i \oplus s\colon V \oplus U \to W$ will be an isomorphism and by induction on the dimension of a module one sees that any L-module breaks up as a sum of irreducible modules.

One basic point I would like to get across is that a question such as this is a question about *cohomology*. In order to understand this let us analyse the splitting of (10) in the most concrete way imaginable. We start by noticing that it always splits *linearly*, i.e. there always exists a map of vector spaces $s\colon U \to W$ such that $\pi \circ s = \mathrm{id}_U$. Such a linear splitting gives us an isomorphism of vector spaces

$$\begin{aligned} V \oplus U &\xrightarrow{\cong} W \\ (v, u) &\mapsto i(v) + s(u) \end{aligned}$$

Let us see how the L-module structure on W looks in terms of this isomorphism. Given $x \in L$, we have

$$x \cdot (i(v) + s(u)) = i(x \cdot v) + x \cdot s(u)$$

and since $\pi(x \cdot s(u)) = x \cdot \pi s(u) = x \cdot u$ we see that $x \cdot s(u) - s(x \cdot u) \in \operatorname{Im} i$. Writing $\phi\colon L \times U \to V$ for the unique bilinear map such that $i\phi(x, u) = x \cdot s(u) - s(x \cdot u)$ we have that

$$x \cdot (i(v) + s(u)) = i(x \cdot v + \phi(x, u)) + s(x \cdot u)$$

Thus the induced L-module structure on the direct sum via the isomorphism is

$$x \cdot (v, u) = (x \cdot v + \phi(x, u), x \cdot u) \tag{11}$$

The bilinear map $\phi\colon L \times U \to V$ measures the failure of s in providing an L-module splitting. Note moreover that ϕ completely determines the L-module structure on W and, indeed, it completely determines the isomorphism class of the extension in the following sense (exercise).

Definition 12.4. *Let L be a Lie algebra, U, V, W, W' be L-modules and*

$$0 \to V \xrightarrow{i} W \xrightarrow{\pi} U \to 0 \qquad 0 \to V \xrightarrow{i'} W' \xrightarrow{\pi'} U \to 0$$

be two extensions of U by V. The two extensions are isomorphic if there is an isomorphism $\psi\colon W \to W'$ making the following diagram commute.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & V & \xrightarrow{i} & W & \xrightarrow{\pi} & U & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \mathrm{id}_V} & & \downarrow{\scriptstyle \psi} & & \downarrow{\scriptstyle \mathrm{id}_U} & & \\ 0 & \longrightarrow & V & \xrightarrow{i'} & W' & \xrightarrow{\pi'} & U & \longrightarrow & 0 \end{array}$$

From now on we will omit i so as to simplify the notation. The map ϕ is not an arbitrary bilinear map, as we have

$$\begin{aligned} \phi([x,y],u) &= [x,y]\cdot s(u) - s([x,y]\cdot u) \\ &= x\cdot(y\cdot s(u)) - y\cdot(x\cdot s(u)) - s(x\cdot(y\cdot u)) + s(y\cdot(x\cdot u)) \\ &= x\cdot(y\cdot s(u)) - y\cdot(x\cdot s(u)) - (x\cdot s(y\cdot u) - \phi(x, y\cdot u)) + (y\cdot s(x\cdot u) - \phi(y, x\cdot u)) \\ &= x\cdot(y\cdot s(u) - s(y\cdot u)) + \phi(x, y\cdot u) - y(x\cdot s(u) - s(x\cdot u)) - \phi(y, x\cdot u) \\ &= x\cdot\phi(y,u) + \phi(x, y\cdot u) - y\cdot\phi(x,u) - \phi(y, x\cdot u), \end{aligned}$$

so regarding ϕ as a linear map

$$\begin{aligned} L &\to \mathrm{Hom}(U,V) \\ x &\mapsto \phi(x,\cdot) \end{aligned}$$

we have

$$\phi([x,y],u) - x\cdot\phi(y,u) - \phi(x,y\cdot u) + y\cdot\phi(x,u) + \phi(y,x\cdot u) = 0 \qquad \text{for all } u \in U$$

or, equivalently, with respect to the L-module structure on $\mathrm{Hom}(U,V)$

$$\phi([x,y]) = (x\cdot\phi)(y) - (y\cdot\phi)(x) \tag{12}$$

The definition of the function ϕ depended on the choice of a section $s\colon U \to W$. What effect does the choice of section have on ϕ? If $s'\colon U \to W$ is another section, we can write

$$s'(u) = s(u) + i\lambda(u)$$

for a unique linear map $\lambda\colon U \to V$, and then

$$\phi'(x,u) = x\cdot s'(u) - s'(xu) = \phi(x,u) + x\lambda(u) - \lambda(xu)$$

or, regarding ϕ, ϕ' as maps from L to $\mathrm{Hom}(U,V)$ and λ as an element in $\mathrm{Hom}(U,V)$,

$$\phi' = \phi + x\cdot\lambda$$

Definition 12.5. *Let L be a Lie algebra and V be an L-module. A* derivation *of L with values in V is a linear map $D\colon L \to V$ such that*

$$D([x,y]) = x\cdot D(y) - y\cdot D(x)$$

The vector space of derivations of L with values in V is denoted $\mathrm{Der}(L,V)$*. An* inner derivation *D_v determined by an element $v \in V$ is a derivation $D_v\colon L \to V$ given by the expression*

$$D_v(x) = x\cdot v$$

(check that this is indeed a derivation). The subspace of inner derivations is denoted $\mathrm{InnDer}(L, V)$*.*

Example 12.6. *(i) If $V = L$ with L acting on itself by $x \cdot v = [x, v]$ then* $\mathrm{Der}(L, L) = \mathrm{Der}(L)$ *is the set of derivations of the Lie algebra L. The inner derivations form the subset* $\mathrm{ad}(L) \subset \mathrm{Der}(L)$*.*

(ii) If V is abelian, i. e. if $x \cdot v = 0$ for all x, v, then a derivation $D \colon L \to V$ is precisely a linear map from L to V which vanishes on $[L, L]$ therefore $\mathrm{Der}(L, V) = \mathrm{Hom}(L/[L, L], V)$*.*

With this definition of derivations with values in a module we see that the functions $\phi \colon L \to \mathrm{Hom}(U, V)$ characterizing an extension are elements of $\mathrm{Der}(L, \mathrm{Hom}(U, V))$ and that two correspond to the same extension if and only if they differ by the inner derivation determined by an element $\lambda \in \mathrm{Hom}(U, V)$). In the homework you will check that any derivation ϕ gives rise to an extension of U by V via the formula (11). We have therefore proved the following result.

Proposition 12.7. *Let L be a Lie algebra and U, V be L-modules. Then isomorphism classes of extensions of U by V*

$$0 \to V \to W \to U \to 0$$

are in one-to-one-correspondence with

$$\mathrm{Der}(L, \mathrm{Hom}(U, V)) / \mathrm{InnDer}(L, \mathrm{Hom}(U, V))$$

We will now consider a similar, but slightly more complicated classification problem. Consider a Lie algebra extension

$$0 \to I \xrightarrow{i} L_1 \xrightarrow{\pi} L \to 0$$

i.e. a short exact sequence of vector spaces where both i and π are maps of Lie algebras. In particular i identifies I with the ideal $\ker \pi \subset L_1$. We refer to this as an extension of the Lie algebra L by the ideal I. The extension is said to be *abelian* if the ideal I is an abelian Lie algebra. In that case, I has a canonical L-module structure for, given $u \in I, x \in L$ and letting $\tilde{x}$ be any lift of x to L_1 (meaning an element of L_1 such that $\pi(\tilde{x}) = x$) we can define

$$x \cdot u = i^{-1}([\tilde{x}, i(u)]) \in U$$

This makes sense because I is an ideal and is well defined because if x' is another lift of x then $x' - \tilde{x} \in I$ and hence, as I is abelian, $[\tilde{x}, u] = [x', u]$ for all $u \in I$ (we are omitting and will continue to omit i in order to simplify the expressions).

Conversely, given an L-module V we can give $L_1 = L \oplus V$ a Lie algebra structure by setting

$$[(x, v), (x', v')] = ([x, x'], x \cdot v' - x' \cdot v)$$

(check) for which V is an abelian ideal sitting in an obvious extension of L by V for which the action of L on V is the initially given one.

Let us try to classify abelian extensions of Lie algebras as we did for extensions of modules. Given an abelian extension, pick a linear map $s : L \to L_1$ with $\pi s = \mathrm{id}_L$. This gives us an isomorphism of vector spaces $L \oplus I \to L_1$ given by

$$(x, y) \mapsto s(x) + y$$

Moreover (omitting i) we have

$$[s(x) + y, s(x') + y'] = [s(x), s(x')] + [y, s(x')] + [s(x), y'] + 0 = [s(x), s(x')] - x' \cdot y + x \cdot y'$$

As before there is a bilinear map $\phi\colon L \times L \to I$ measuring the failure of s in being a Lie algebra homomorphism. It is given by

$$(x, x') \mapsto [s(x), s(x')] - s([x, x'])$$

and, since

$$[s(x) + y, s(x') + y'] = s([x, x']) + \phi(x, x') - x' \cdot y + x \cdot y'$$

we again see that the Lie algebra extension is completely determined by ϕ. Note that ϕ is, by definition, skew symmetric. As before, this skew symmetric function is not arbitrary. We leave it as an exercise to check the following relation.

Lemma 12.8. *For all $x, y, z \in L$, the skew symmetric function $\phi\colon L \times L \to I$ satisfies the following equality*

$$-\phi([x, y], z) + \phi([x, z], y) - \phi([y, z], x) + x \cdot \phi(y, x) - y \cdot \phi(x, z) + z \cdot \phi(x, y) = 0 \tag{13}$$

Conversely, one checks (as you will in the homework) that a skew-symmetric function satisfying (13) gives rise to a Lie bracket on $L \oplus I$ sitting in an extension of L by I. The simplest extension of L by I is the one obtained by setting $\phi = 0$.

Definition 12.9. *Let L be a Lie algebra and I be an L-module. The Lie algebra structure on $L \oplus I$ defined by*

$$[(x, y), (x', y')] = ([x, x'], x \cdot y' - x' \cdot y)$$

is called semi-direct product *of L and I and denoted $L \ltimes I$. More generally, if L acts on a Lie algebra M by derivations, the semi-direct product of L by M with respect to the action is the Lie algebra $L \ltimes M$ which equals $L \oplus M$ as a vector space and has Lie bracket given by the expression*

$$[(x, y), (x', y')] = ([x, x'], x \cdot y' - x' \cdot y + [y, y'])$$

Let us now see how the choice of section s affects ϕ. Another section $s'\colon L \to I$ is of the form $s' = s + \lambda$, with $\lambda\colon L \to I$ an arbitrary linear map. Then

$$\begin{aligned}
\phi'(x, y) &= [s'(x), s'(y)] - s'([x, y]) \\
&= [s(x), s(y)] + [\lambda(x), s(y)] + [s(x), \lambda(y)] + [\lambda(x), \lambda(y)] - s([x, y]) - \lambda([x, y]) \\
&= \phi(x, y) \ - y \cdot \lambda(x) + x \cdot \lambda(y) - \lambda([x, y])
\end{aligned}$$

Note that, since I is abelian, the difference between $\phi'(x, y)$ and $\phi(x, y)$ is exactly given by the expression vanishing in (12)! Moreover there is clearly a pattern in conditions (12) and (13) which is a telltale of *cohomology.* It can be generalized as follows.

Definition 12.10. *Let $\mathbb{K}$ be a field, L a Lie algebra over $\mathbb{K}$ and V be an L-module. The* Chevalley-Eilenberg complex *of L with coefficients in V is the cochain complex*

$$C^k(L;V) = \mathrm{Hom}(\Lambda^k L, V), \qquad \text{(the space of k-multilnear maps $L \times \cdots \times L \to V$)}$$

for $k \geq 0$, with differential

$$\delta\colon C^k(L;V) \to C^{k+1}(L;V)$$

defined by

$$\begin{aligned}(\delta\varphi)(x_0,\ldots,x_k) &= \sum_{0\leq i<j\leq k} (-1)^{i+j}\varphi([x_i,x_j],x_0,\ldots,\hat{x}_i,\ldots,\hat{x}_j,\ldots,x_k) \\ &\quad + \sum_{i=0}^{k} (-1)^i x_i \cdot \varphi(x_0,\ldots,\hat{x}_i,\ldots,x_k)\end{aligned}$$

The cohomology of this complex is called the cohomology of L with coefficients in V*:*

$$H^k(L;V) = (\ker\delta : C^k \to C^{k+1})/(\mathrm{Im}\,\delta : C^{k-1} \to C^k)$$

It is left as an exercise to check that $\delta^2 = 0$ so that the above definition makes sense.

Remark 12.11. *Note that the deRham complex of a manifold M is the special case when $L = \mathcal{X}(M)$ is the Lie algebra of vector fields on the manifold M and $V = C^\infty(M)$ is the vector space of smooth functions on M with the usual action of $\mathcal{X}(M)$. That is*

$$H^*_{dR}(M) = H^*(\mathcal{X}(M); C^\infty(M))$$

(although for the above to be precise we must consider suitably "smooth" cochains, rather then arbitrary multilinear maps from L to V).

It is worth writing down explicitly the formula for the differential in low degrees.

$$C^0(L;V) = V \xrightarrow{\delta} C^1(L;V) = \mathrm{Hom}(L;V) \xrightarrow{\quad\delta\quad} C^2(L;V) = \mathrm{Hom}(\Lambda^2 L, V)$$

$$v \longmapsto (x \mapsto x\cdot v)$$

$$\varphi \longmapsto ((x,y) \mapsto -\varphi([x,y]) + x\cdot\varphi(y) - y\cdot\varphi(x))$$

In particular, the 1-cocycles are the derivations, and the 1-coboundaries are the inner derivations.

Proposition 12.12. *Let L be a Lie algebra and V be an L-module. Then:*

(*i*) $H^0(L;V) = V^L = \{v \in V : x\cdot v = 0$ *for all* $x \in L\}$ *is the space of invariants under the action of L;*

(*ii*) $H^1(L;V) = \mathrm{Der}(L;V)/\,\mathrm{InnDer}(L;V)$ *classifies extensions of U by W when $V = \mathrm{Hom}(U,W)$;*

(*iii*) $H^2(L;V)$ *classifies abelian extensions of L by V.*

(*iv*) *If* $\mathbb{K} \to \mathbb{F}$ *is a field extension and* L *is a Lie algebra over* $\mathbb{K}$ *then*

$$H^*(L \otimes_{\mathbb{K}} \mathbb{F}; V \otimes_{\mathbb{K}} \mathbb{F}) = H^*(L; V) \otimes_{\mathbb{K}} \mathbb{F};$$

(*v*) *If* $0 \to V_1 \to V_2 \to V_3 \to 0$ *is a short exact sequence of* L*-modules, then there is a long exact sequence on cohomology:*

$$\begin{tikzcd}
0 \arrow[r] & H^0(L;V_1) \arrow[r] & H^0(L;V_2) \arrow[r] & H^0(L;V_3) \arrow[dll, rounded corners] \\
& H^1(L;V_1) \arrow[r] & H^1(L;V_2) \arrow[r] & H^1(L;V_3) \arrow[dll, rounded corners] \\
& H^2(L;V_1) \arrow[r] & H^2(L;V_2) \arrow[r] & \ldots
\end{tikzcd}$$

Proof. (*i*) is immediate from the definition and we have already pointed out (*ii*). Statement (*iii*) is also immediate from the definition of cohomology and the discussion preceding the definition of the Chevalley-Eilenberg complex. Statement (*iv*) is also immediate from the definition of cohomology and the properties of the tensor product. To see (*v*), apply the snake lemma to the short exact sequence of cochain complexes

$$0 \to C^*(L; V_1) \to C^*(L; V_2) \to C^*(L; V_3) \to 0$$

□

Let G be a Lie group. Then, the subspaces

$$\Omega^*_\ell(G) \subset \Omega^*(G)$$

of left invariant forms on G constitute a subcomplex of the deRham complex (because the exterior derivative is linear and commutes with pullbacks - exercise!). Evaluation at $e \in G$ gives rise to an isomorphism of vector space

$$\Omega^k_\ell(G) \xrightarrow{\mathrm{ev}_e} C^k(\mathfrak{g}; \mathbb{R}) = \mathrm{Hom}(\Lambda^k \mathfrak{g}, \mathbb{R})$$

A form on G is completely determined by its evaluation on left invariant vector fields. Given $X_0, \ldots, X_k \in \mathcal{X}(G)$ left invariant we have

$$\begin{aligned}
d\omega(X_0, \ldots, X_k) &= \sum_{0 \le i < j \le k} (-1)^{i+j} \omega([X_i, X_j], X_0, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_k) \\
&\quad + \sum_{i=1}^{k} (-1)^i X_i \cdot \omega(X_0, \ldots, \hat{X}_i, \ldots, X_k) \\
&= \sum_{0 \le i < j \le k} (-1)^{i+j} \omega([X_i, X_j], X_0, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_k)
\end{aligned}$$

where the last equality holds because the evaluation of a left invariant form on left invariant vector fields is a constant function. We conclude the following.

Proposition 12.13. *The map* $\mathrm{ev}_e \colon \Omega^*_\ell(G) \to C^*(\mathfrak{g};\mathbb{R})$ *is an isomorphism of cochain complexes (where* $\mathbb{R}$ *denotes the trivial* $\mathfrak{g}$*-module).*

Now if G is a *compact* Lie group, there is an averaging map which turns any form on G into a left invariant form:

$$\pi \colon \Omega^*(G) \to \Omega^*_\ell(G)$$

is defined by

$$\pi(\omega) = \int_G L_g^* \omega dg$$

If $i \colon \Omega^k_\ell(G) \hookrightarrow \Omega^k(G)$ denotes the inclusion then clearly $\pi i = \mathrm{id}_{\Omega^*_\ell(G)}$. It is clear that π is a map of complexes, as is i obviously, so we conclude that, in this case,

$$H^*(\mathfrak{g};\mathbb{R}) \xrightarrow{i^*} H^*_{dR}(G)$$

is injective. However, using deRham's Theorem one can check (as you will do in the homework) that i^* is surjective and therefore we have.

Proposition 12.14. *If* G *is a compact Lie group, then* $H^*_{dR}(G) \cong H^*(\mathfrak{g};\mathbb{R})$*.*

Example 12.15. *Let* $L = \mathbb{R}^n$ *be an abelian Lie algebra. Then the differential* δ *in the Chevalley-Eilenberg complex is identically zero and therefore*

$$H^k(L;\mathbb{R}) = C^k(L;\mathbb{R}) = (\Lambda^k(\mathbb{R}^n))^*$$

so that $H^k(L;\mathbb{R})$ *is a real vector space of dimension* $\binom{n}{k}$*. This is indeed the deRham cohomology of an* n*-dimensional torus* $S^1 \times \cdots \times S^1$*, which is the unique compact Lie group with Lie algebra* L*.*

Note that the Lie groups with Lie algebra $\mathbb{R}^n$ *are those of the form* $\mathbb{R}^n/D$ *with* D *a discrete subgroup of* $\mathbb{R}^n$*. One can check (exercise) that these subgroups are all isomorphic to some* $\mathbb{Z}^k$ *spanning a* k*-dimensional subspace of* $\mathbb{R}^n$ *(for some* $k \leq n$*) and then* $\mathbb{R}^n/D \cong (S^1)^k \times \mathbb{R}^{n-k}$*.*

We will finish this section by giving another perspective on L-modules and cohomology which will be important for the computation of cohomology. Let V be a vector space over a field $\mathbb{K}$ and

$$T(V) = \mathbb{K} \oplus V \oplus (V \otimes V) \oplus V^{\otimes 3} \oplus \dots$$

denote the *tensor algebra on* V. The product is defined on decomposable tensors by the expression

$$(v_1 \otimes \cdots \otimes v_k) \cdot (w_1 \otimes \cdots \otimes w_\ell) = v_1 \otimes \cdots \otimes v_k \otimes w_1 \otimes \cdots \otimes w_\ell$$

and then extended linearly

Example 12.16. *If* $V \cong \mathbb{K}^2$*, then* $T(V)$ *is the polynomial algebra on "two non-commuting variables". Indeed, picking a basis* $\{x, y\}$ *for* V*, a basis for* $T(V)$ *is given by*

$$\{1, x, y, x \otimes x, x \otimes y, y \otimes x, y \otimes y, \dots\}$$

and so on can think of $T(V)$ *as a polynomial algebra in two non-commuting variables* x *and* y*.*

The algebra $T(V)$ is also called the *free associative algebra on* V for the following reason. If A is any associative algebra over $\mathbb{K}$ and $f : V \to A$ is a linear map, there is a unique algebra homomorphism $\bar{f}$ making the following diagram (of linear maps) commute:

$$\begin{tikzcd} V \arrow[rr, "f"] \arrow[dr, hook] & & A \\ & T(V) \arrow[ur, dashed, "\bar{f}"'] & \end{tikzcd}$$

Indeed there clearly is exactly one map $\bar{f}$ making the diagram commute; it must be given on decomposable tensors by the expression

$$\bar{f}(v_1 \otimes \cdots \otimes v_n) = f(v_1) \cdot \cdots \cdot f(v_n)$$

This property (called a universal property) completely determines $T(V)$ up to isomorphism. It says that "to give a map of $\mathbb{K}$-algebras out of $T(V)$ is the same as to give a linear map out of V".

Definition 12.17. *Let L be a Lie algebra over a field $\mathbb{K}$. The* universal enveloping algebra *of L is the algebra*

$$U(L) = T(L)/\langle x \otimes y - y \otimes x - [x, y] \colon x, y \in L\rangle$$

(where the quotient is by the two-sided ideal[8] *in $T(L)$ generated by the given elements).*

We have given a "hands on" definition of $U(L)$ but it is, like the tensor algebra, characterized by a universal property. Recall that any associative algebra A becomes a Lie algebra with the Lie bracket given by the commutator of elements of A.

Proposition 12.18. *Let L be a Lie algebra, A an associative algebra over the field $\mathbb{K}$, and $f \colon L \to A$ be a map of Lie algebras. Then there is a unique map $\bar{f} \colon U(L) \to A$ of associative algebras making the following diagram commute*

$$\begin{tikzcd} L \arrow[rr, "f"] \arrow[dr, "i"] & & A \\ & U(L) \arrow[ur, dashed, "\bar{f}"'] & \end{tikzcd}$$

where i is the canonical map $L \hookrightarrow T(L) \to U(L)$.

Proof. By the universal property of the tensor algebra, the map f can be extended to a unique map of algebras $\tilde{f} \colon T(L) \to A$. The existence and uniqueness of the required map $\bar{f}$ follows from the fact that maps from $U(L) \to A$ correspond to maps from $T(L) \to A$

[8] A subspace $I \subset A$ of an associative algebra is a two sided ideal i f $aI \subset I$ and $Ia \subset I$ for all $a \in A$. These are the kernels of homomorphisms of algebras, just as in the Lie case.

sending the ideal $\langle x \otimes y - y \otimes x - [x,y] \colon x, y \in L \rangle$ to $0 \in A$.

$$\begin{tikzcd} L \arrow[r, "f"] \arrow[d] & A \\ T(L) \arrow[ur, dashed, "\tilde{f}"] \arrow[d] & \\ U(L) \arrow[uur, dashed, bend right, "\bar{f}"'] & \end{tikzcd}$$

But the condition that the generators $x \otimes y - y \otimes x - [x,y]$ of the ideal go to 0 under $\tilde{f}$ is exactly equivalent to the condition that f is a map of Lie algebras. ☐

Example 12.19. *Consider the Lie algebra $\mathfrak{sl}(2;\mathbb{K}) = \langle h, x, y \rangle$ with the usual basis*

$$h = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad x = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

As $[x,y] = h,\ [h,x] = 2y$ and $[h,y] = -2x$, we obtain (for instance) the following basis for the universal enveloping algebra:

$$U(\mathfrak{sl}(2;\mathbb{K})) = \langle 1, h, x, y, h^2, xh, x^2, xy, y^2, hy, x^3, \dots \rangle$$

where we have not written yx because that element of $T(L)$ is identified in $U(L)$ with $xy - [x,y]$, and similarly for the other missing quadratic terms.

Remark 12.20. *When $L = \mathrm{Lie}(G)$ is the Lie algebra of a Lie group G, the universal enveloping algebra $U(L)$ can be interpreted as the algebra of left invariant linear differential operators on G.*

We now come to a central Theorem whose proof we will regrettably omit.

Theorem 12.21 (Poincaré-Birkhoff-Witt theorem)**.** *Let L be a finite dimensional Lie algebra over a field $\mathbb{K}$ and $\{v_1, \dots, v_n\}$ be a basis for L. Then*

$$\{[v_{i_1} \dots v_{i_\ell}] \colon i_1 \le i_2 \le \dots \le i_\ell\}$$

is a vector space basis for $U(L)$.

Proof. See, for instance, [Hu, Chapter VI]. ☐

The previous Theorem states that $U(L)$ has the same "size" as a symmetric algebra on L (i.e. the polynomial algebra in $\dim L$ variables). In particular, *the canonical map $L \to U(L)$ is injective*, a fact which, although plausible, is far from obvious.

Remark 12.22. *If the Lie algebra L is abelian, then one easily checks that $U(L)$ is the symmetric algebra on L. One can think of $U(L)$ for a general L as a "deformation" of the symmetric algebra. The underlying vector spaces can be identified but the product in $U(L)$ is no longer the product in a polynomial algebra when L is not abelian.*

Now, the universal property of $U(L)$ implies that an L-module is precisely the same as a $U(L)$-module. This is our final way of regarding a representation of L. Indeed, an L-module is the same as a Lie algebra endomorphism $\rho\colon L \to \mathrm{End}(V)$ which corresponds via Proposition 12.18 to a map of associative algebras

$$\bar{\rho}\colon U(L) \to \mathrm{End}(V)$$

and this is the same as a $U(L)$ module structure on V (which is by definition a bilinear map $U(L) \times V \to V$ satisfying $1 \cdot v = v,\ a(bv) = (ab)v$ for all $a, b \in U(L)$ and $v \in V$). We can use this to rewrite the Chevalley-Eilenberg cochain complex

$$C^k(L;V) = \mathrm{Hom}_{\mathbb{K}}(\Lambda^k L, V) = \mathrm{Hom}_{U(L)}(U(L) \otimes_{\mathbb{K}} \Lambda^k L, V). \tag{14}$$

The gain in regarding the Chevalley-Eilenberg complex in this way is that we see that the differentials in the cochain complex are actually induced by maps of $U(L)$-modules between the free $U(L)$-modules $U(L) \otimes_{\mathbb{K}} \Lambda^* L$. Indeed $\delta\colon C^k(L;V) \to C^{k+1}(L;V)$ can be reinterpreted as

$$\delta = \mathrm{Hom}_{U(L)}(\partial, V) \qquad \text{with } \partial\colon U(L) \otimes_{\mathbb{K}} \Lambda^{k+1} L \to U(L) \otimes_{\mathbb{K}} \Lambda^k L$$

given (on generators) by the expression

$$\begin{aligned} \partial(x_0 \wedge \cdots \wedge x_k) &= \sum_{i<j} [x_i, x_j] \wedge \cdots \wedge \hat{x}_i \wedge \cdots \wedge \hat{x}_j \wedge \cdots \wedge x_k \\ &\quad + \sum_i ((-1)^i x_i) \cdot x_0 \wedge \cdots \wedge \hat{x}_i \wedge \cdots \wedge x_k \end{aligned}$$

Thus we see that the Chevalley-Eilenberg complex comes from a more fundamental object, *the Chevalley-Eilenberg chain complex of L* consisting of free $U(L)$-modules

$$(U(L) \otimes_{\mathbb{K}} \Lambda^* L, \partial).$$

Proposition 12.23. *The Chevalley-Eilenberg chain complex of a Lie algebra L is acyclic, meaning*

$$H_k(U(L) \otimes_{\mathbb{K}} \Lambda^k L, \partial) = \begin{cases} 0 & \text{if } k > 0 \\ \mathbb{K} & \text{if } k = 0 \end{cases}$$

Proof. This will be in the next homework. It makes use of the Poincaré-Birkhoff-Witt Theorem. □

Remark 12.24. *An acyclic complex as in the previous statement is called a* free resolution *of the trivial module $\mathbb{K}$. Relation* (14) *then identifies Lie algebra cohomology with the* derived functors *of the functor* $\mathrm{Hom}_{U(L)}(\mathbb{K}, \cdot) = (\cdot)^L$ *which computes the invariants of an L-module. These functors are usually denoted by* $\mathrm{Ext}_{U(L)}(\mathbb{K}, \cdot)$. *See for instance* [We] *for more on this.*

The new interpretation of Lie algebra cohomology afforded by (14) is very useful because it allows us to invoke not only functoriality in the L-module V, which was already apparent in the original definition, but crucially, also in the trivial module $\mathbb{K}$ which is being resolved by the Chevalley-Eilenberg chain complex.

13. The Whitehead Lemmas, Weyl's and Levi's Theorems, Lie's third Theorem

Suppose L is a semisimple Lie algebra over a field $\mathbb{K}$ of characteristic zero. Then the Killing form κ gives us an isomorphism (still denoted κ)

$$\kappa \colon L \to L^*$$

defined by

$$x \mapsto \kappa(x, \cdot)$$

The associativity property $\kappa([x,y],z) = \kappa(x,[y,z])$ implies that κ is in fact an isomorphism of L-modules:

$$\kappa([x,y],\cdot) = -\kappa(y,[x,\cdot]) = x \cdot (\kappa(y,\cdot))$$

Now consider the isomorphism

$$L \otimes L \xrightarrow{\mathrm{id} \otimes \kappa} L \otimes L^* = \mathrm{End}(L)$$

On the right we have a canonical element id_L, which in terms of a vector space basis $\{v_i\}$ for V and the corresponding dual basis v_i^* is written

$$\sum_{i=1}^{n} v_i \otimes v_i^*$$

This gives rise to a canonical element in $L \otimes L$ and hence in $U(L)$ under the canonical map

$$L \otimes L \hookrightarrow T(L) \to U(L)$$

Definition 13.1. *Let L be a semisimple Lie algebra over a field $\mathbb{K}$ of characteristic zero. The* Casimir element *of L is the element*

$$c(L) = \sum_{i=1}^{n} v_i v^i \in U(L)$$

where $\{v_1, \dots, v_n\}$ is any basis of L and $\{v^1, \dots, v^n\}$ is the dual basis with respect to the Killing form (i.e. the elements $v^i \in L$ are characterized by the relations $\kappa(v_i, v^j) = \delta_{ij}$).

The discussion preceding the definition shows that $c(L)$ is independent of the choice of basis for L.

Example 13.2. *Consider the Lie algebra $\mathfrak{sl}(2;\mathbb{K})$ with basis $\{h,x,y\}$ as in Example 12.19. With respect to this basis we have seen that κ is given by the symmetric matrix* $\begin{bmatrix} 8 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \end{bmatrix}$

Therefore the dual basis for $\mathfrak{sl}(2;\mathbb{K})$ with respect to κ is

$$h^* = \tfrac{1}{8}h, \qquad x^* = \tfrac{1}{4}y, \qquad y^* = \tfrac{1}{4}x$$

and the Casimir element in $U(\mathfrak{sl}(2;\mathbb{K}))$ is

$$hh^* + xx^* + yy^* = \frac{1}{8}h^2 + \frac{1}{4}xy + \frac{1}{4}yx = \frac{1}{8}h^2 + \frac{1}{4}h + \frac{1}{2}yx,$$

where we have used that $[x, y] = h$ in the second equality.

The significance of the Casimir element is the following.

Proposition 13.3. *The Casimir element $c(L) \in U(L)$ is a nontrivial element in the center of $U(L)$ and therefore gives rise to a canonical endomorphism for any L-module V.*

Proof. The action of L on the L-module $\mathrm{End}(L)$ is given by the commutator with $\mathrm{ad}(x)$. Therefore, for all $x \in L$, we have $x \cdot \mathrm{id}_L = 0$. Since $c(L)$ is the image of id_L under the L-module isomorphism $\mathrm{id} \otimes \kappa \colon L \otimes L \to \mathrm{End}(L)$ it follows that $x \cdot c(L) = 0$ for all $x \in L$.

Writing $c(L) = \sum v_i \otimes v^i$ this means that

$$\sum [x, v_i] \otimes v^i + v_i \otimes [x, v^i] = 0$$

and the image of the previous equality in $U(L)$ is

$$\sum x v_i v^i - v_i x v^i + v_i x v^i - v_i v^i x = 0 \Leftrightarrow x \sum v_i v^i = (\sum v_i v^i) x \Leftrightarrow x c(L) = c(L) x$$

Thus the Casimir element commutes with all the generators of $U(L)$ and is therefore in the center of $U(L)$.

The element $c(L)$ is nontrivial because the trace of its action on L (via the adjoint action) is nonzero: indeed, $c(L)$ acts on L as

$$(\sum_i v_i v^i) x = \sum_i [v_i, [v^i, x]] = \sum_i \mathrm{ad}(v_i)\,\mathrm{ad}(v^i)(x)$$

and the trace of $ad(v_i)\,\mathrm{ad}(v^i)$ is, by definition, $\kappa(v_i, v^i)$. Hence

$$\mathrm{tr}(c(L)) = \mathrm{tr}(\sum_i v_i v^i) = \sum_i \kappa(v_i, v^i) = \dim L \neq 0$$

where we are using again that $\mathbb{K}$ has characteristic zero.

Finally, any element z in the center of an associative algebra A determines, by left multiplication, an endomorphism of any module over A. ☐

Remark 13.4. *If G is a semisimple Lie group (meaning a Lie group whose Lie algebra is semisimple), then interpreting the universal enveloping algebra of $\mathfrak{g}$ as the algebra of left invariant linear differential operators, the Casimir element is a canonically defined, left invariant, second order linear differential operator on G, which commutes with all other left invariant linear differential operators. One can check that, up to a scalar multiple, it corresponds to the Laplacian determined by a left invariant metric on G.*

The construction giving rise to the Casimir element can be performed using any non-degenerate, associative, symmetric bilinear form $\beta \colon L \times L \to \mathbb{K}$. One obtains an element $c(\beta) = \sum_i v_i v^i \in U(L)$ where $\{v_1, \ldots, v_n\}$ is a basis for L and $\{v^1, \ldots, v^n\}$ is the dual basis with respect to the pairing β. This element is again in the center of $U(L)$ by the same argument as above. Here is how such bilinear pairings arise:

Lemma 13.5. *Let L be a semisimple Lie algebra over a field of characteristic* 0*. If $\rho\colon L \to \mathfrak{gl}(V)$ is a faithful representation, then*

$$\beta(x, y) = \mathrm{tr}(\rho(x)\rho(y))$$

defines a non-degenerate, associative, symmetric bilinear form on L.

Proof. Clearly β is bilinear and symmetric because the trace form is. We check associativity

$$\beta([x, y], z) = \mathrm{tr}([\rho(x), \rho(y)]\rho(z)) = \mathrm{tr}(\rho(x)[\rho(y), \rho(z)]) = \beta(x, [y, z])$$

where we have used associativity of the trace form. Non-degeneracy follows as in the proof of Cartan's criterion for solvability: If $I = \mathrm{Rad}\,\beta$, then associativity of β implies that I is an ideal in L. The trace form vanishes identically on $\rho(I) \subset \mathfrak{gl}(V)$ hence $\rho(I)$ is solvable. As ρ is faithful, it follows that I is a solvable ideal in L. As L is semisimple, I must be 0, i.e. β is nondegenerate. □

Given a faithful representation $\rho\colon L \to \mathfrak{gl}(V)$, we will write $c(\rho) \in U(L)$ for the *Casimir element of the representation ρ*. Recall this is given by

$$c(\rho) = \sum_i v_i v^i$$

with $\{v_1, \dots, v_n\}$ any basis for L and $\{v^1, \dots, v^n\}$ the dual basis determined by $\mathrm{tr}(\rho(v_i)\rho(v^j)) = \delta_{ij}$. As before, $c(\rho) \in Z(U(L))$. It is a nontrivial element of the center because the trace of its action on V in nontrivial (by definition of β, its trace is $\dim L$ - check!).

If $\rho\colon L \to \mathfrak{gl}(V)$ is not a faithful representation then $\ker \rho$ is an ideal and we can decompose L as a product of ideals

$$L = (\ker \rho) \times (\ker \rho)^\perp$$

(where the orthogonal is taken with respect to the Killing form) and this decomposition gives rise to a decomposition of the universal enveloping algebra of L as

$$U(L) = U(\ker \rho) \otimes_{\mathbb{K}} U((\ker \rho)^\perp)$$

Check this as an exercise with the universal property of the enveloping algebras. Now ρ restricts to a faithful representation of the semisimple Lie algebra $(\ker \rho)^\perp$ and we then define $c(\rho) \in U(L)$ to be the image of the Casimir element

$$c(\rho_{|(\ker \rho)^\perp}) \in U((\ker \rho)^\perp)$$

under the canonical inclusion

$$U((\ker \rho)^\perp) \hookrightarrow U(L)$$

The element $c(\rho)$ is central in $U(L)$ (because it is in $U((\ker \rho)^\perp)$) and the trace of its action on V is easily checked to be $\dim(\ker \rho)^\perp$, so $c(\rho)$ is non-trivial as soon as the representation ρ is not trivial.

Remark 13.6. *If we write the semisimple Lie algebra as the product of its simple factors $L = L_1 \times \cdots \times L_n$ with L_i then*

$$U(L) = U(L_1) \otimes_{\mathbb{K}} \cdots \otimes_{\mathbb{K}} U(L_n)$$

Writing $c_i \in U(L_i)$ for the Casimir element of L_i, it is not hard to see that for any representation $\rho : L \to \mathfrak{gl}(V)$ we have

$$c(\rho) = \sum_i \lambda_i c_i$$

for some scalars $\lambda_i \in \mathbb{K}$, which will be 0 *for the factors in the kerne of ρ and nonzero otherwise.*

We now come to a central result in representation theory (whose proof is completely trivial). This result holds for any kind of representations (Lie groups, Lie algebras, associative algebras) with the same proof but for simplicity we state it only in the case of Lie algebra representations.

Lemma 13.7 (Schur's lemma)**.** *Let L be a Lie algebra and $f\colon V \to W$ be an L-module homomorphism between* simple modules*. Then,*

(1) Either $f = 0$ or f is an isomorphism.
(2) If $V = W$ and the base field is algebraically closed, then f is multiplication by some scalar.

Proof. $f(V) \subset W$ is a submodule. Either it is 0 and then $f = 0$ or it is different from 0 and then $f(V) = W$ so f is surjective. In the latter case, since $\ker f \neq V$ is a submodule and it is not the whole of V, we must have $\ker f = 0$. We conclude that if $f \neq 0$ then f is an isomorphism.

Suppose now that $V = W$ and the base field is algebraically closed. An endomorphism $f\colon V \to V$ will have a non-trivial eigenspace which is easily checked to be a submodule of V. Since V is simple, the eigenspace must be all of V. □

Example 13.8. *Let L be a semisimple Lie algebra over an algebraically closed field of characteristic zero and let V be a simple L-module. Then*

$$v \mapsto c(\rho)v$$

is an endomorphism of the L-module V and therefore must correspond to multiplication by some scalar. Since $\mathrm{tr}(c(\rho)) = \dim(\ker\rho)^\perp$*, we conclude that the scalar must be* $\dim(\ker\rho)^\perp/\dim V$*.*

We can now exploit the Casimir element to obtain information about the cohomology of semisimple Lie algebras.

Theorem 13.9. *Let L be a semisimple Lie algebra over an algebraically closed field $\mathbb{K}$ of characteristic* 0 *and V be a simple L-module* other than $\mathbb{K}$ with the trivial action*. Then $H^i(L;V) = 0$ for all $i \geq 0$.*

Proof. Let $\rho\colon L \to \mathfrak{gl}(V)$ be the representation corresponding to the module structure on V. The assumption that V is not isomorphic to the trivial module $\mathbb{K}$ implies that $(\ker\rho)^\perp \neq 0$ (as a simple module on which L acts trivially clearly must have dimension one). Consider the map induced on cohomology by $c(\rho)$

$$c(\rho)\colon H^i(L;V) \to H^i(L;V)$$

By Example 13.8, this map must be multiplication by the nonzero scalar $\frac{\dim(\ker\rho)^\perp}{\dim V}$.

On the other hand, multiplication by $c(\rho)$ gives a self map of the cochain complex

$$\mathrm{Hom}_{U(L)}(U(L)\otimes_{\mathbb{K}}\Lambda^k L, V)\xrightarrow{c(\rho)}\mathrm{Hom}_{U(L)}(U(L)\otimes_{\mathbb{K}}\Lambda^k L, V)$$

which is induced by multiplication by $c(\rho)$ on the Chevalley-Eilenberg chain complex

$$U(L)\otimes_{\mathbb{K}}\Lambda^k L\xrightarrow{c(\rho)}U(L)\otimes_{\mathbb{K}}\Lambda^k L$$

Now these complexes have trivial homology except in degree 0, where the homology is $\mathbb{K}$ with trivial L-action. The map induced by acting with $c(\rho)$ on the trivial module $\mathbb{K}$ is the zero map, hence, in every degree, multiplication by $c(\rho)$ induces the 0 map on homology. A standard Theorem in Homological Algebra (exercise) says that the map induced on cohomology is the dual of the map on homology so the map on cohomology is also 0 in every degree. Above we proved that the map on cohomology is given by muItiplication by a nonzero scalar. This can only be if $H^i(L;V)=0$ for all $i\geq 0$. □

Corollary 13.10 (First Whitehead lemma)**.** *Let L be a semisimple Lie algebra over a field $\mathbb{K}$ of characteristic 0 and V any L-module. Then*

$$H^1(L;V)=0$$

Proof. Let $\overline{\mathbb{K}}$ denote the algebraic closure of $\mathbb{K}$. Since $H^1(L\otimes_{\mathbb{K}}\overline{\mathbb{K}}, V\otimes_{\mathbb{K}}\overline{\mathbb{K}})\cong H^1(L;V)\otimes_{\mathbb{K}}\overline{\mathbb{K}}$, we may assume that $\mathbb{K}$ is algebraically closed.

By Theorem 13.9 the result holds if V is a simple module other than the trivial module $\mathbb{K}$. In the latter case, we have $\mathrm{Der}(L,\mathbb{K})=\mathrm{Hom}(L/[L,L],\mathbb{K})=0$ as $L=[L,L]$ and therefore $H^1(L;\mathbb{K})=\mathrm{Der}(L;\mathbb{K})/\mathrm{InnDer}(L;\mathbb{K})=0$. We conclude that H^1 vanishes on all simple modules.

A short exact sequence of L-modules $0\to V\to W\to U\to 0$ gives rise to an exact sequence

$$H^1(L;V)\to H^1(L;W)\to H^1(L;U).$$

so induction on the dimension of the module W gives us $H^1(L;W)=0$ for all L-modules W. □

Remark 13.11. *Taking $V=L$ in the previous result we se that all derivations of a semisimple Lie algebra are inner derivations. If L is the Lie algebra of a connected Lie group G, this implies that the Lie group* $\mathrm{Aut}(L)$ *(which has* $\mathrm{Der}(L)$ *as its Lie algebra) has connected component of the identity $G/Z(G)$.*

We can now prove a basic result about semisimple Lie algebras.

Proof of Theorem 12.3. Extensions of modules are classified by H^1 which vanishes identically by the First Whitehead Lemma. Hence all extensions of modules split and then inductively on the dimension we see that every finite dimensional L-module is a direct sum of simple L-modules □

Proposition 13.12. *Let L be a semisimple Lie algebra over a field $\mathbb{K}$ of characteristic 0. Then,*

$$H^2(L;\mathbb{K}) = 0$$

(where $\mathbb{K}$ is the trivial module).

Proof. We need to show that any extension

$$0 \to \mathbb{K} \to L_1 \xrightarrow{\pi} L \to 0 \tag{15}$$

splits, i.e. that there is a map of Lie algebras $s\colon L \to L_1$ with $\pi \circ s = \mathrm{id}_L$. Define an L-module structure on L_1 by the expression

$$x \cdot y = [\tilde{x}, y] \quad \text{for } x \in L, y \in L_1 \text{ and } \tilde{x} \text{ any element in } L_1 \text{ lifting } x$$

This is well defined because if $\bar{x}$ is any other lift of x we have $\bar{x} = \tilde{x} + \lambda$ for some $\lambda \in \mathbb{K}$ and $[\lambda, y] = 0$ (we are assuming that the action of L on $\mathbb{K}$ determined by the extension is trivial so $\mathbb{K}$ is central). It is now easy to see that the formula above gives L_1 an L-module structure and, moreover, (15) is a short exact sequence of L-modules.

By the Weyl's Theorem, there is a map $s\colon L \to L_1$ of L-modules such that $\pi s = \mathrm{id}_L$, and this map gives the required splitting: given $x_1, x_2 \in L$ we have

$$s([x_1, x_2]) = s(x_1 \cdot x_2) = x_1 \cdot s(x_2) = [\tilde{x}_1, s(x_2)]$$

for any lift $\tilde{x}_1$ of x_1 and we may take $\tilde{x}_1$ to be $s(x_1)$. □

Corollary 13.13 (Second Whitehead lemma)**.** *Let L be a semisimple Lie algebra over a field of characteristic* 0*. Then $H^2(L;V) = 0$ for any L-module V.*

Proof. Exercise (this is just as in the First Whitehead Lemma). □

We can now prove a basic structure theorem about Lie algebras which will allow us to finally give a proof of Lie's third Theorem.

Theorem 13.14 (Levi's theorem)**.** *Let L be a Lie algebra over a field of characteristic* 0*. Then the canonical extension*

$$0 \to \operatorname{Rad} L \to L \xrightarrow{\pi} L/\operatorname{Rad}(L) \to 0$$

splits, i.e. there is a map of Lie algebras $s\colon L/\operatorname{Rad}(L) \to L$ such that $\pi s = \mathrm{id}_{L/\operatorname{Rad}(L)}$. Therefore any Lie algebra L over a field of characteristic zero is a semidirect product of a semisimple Lie algebra with a solvable Lie algebra:

$$L \cong (L/\operatorname{Rad}(L)) \ltimes \operatorname{Rad}(L)$$

Proof. Let $S = \operatorname{Rad}(L)$. If S is abelian the statement follows from the second Whitehead Lemma as L/S is semisimple and abelian extensions are classified by H^2. We now proceed by induction on the degree of solvability of S. Suppose the result holds when $S^{(k)} = 0$ and assume S is such that $S^{(k+1)} = 0$. Then

$$0 \to S/S^{(k)} \to L/S^{(k)} \xrightarrow{p} L/S \to 0$$

splits, so there is a Lie algebra $L_1 \subset L/S^{(k)}$ mapping isomorphically to L/S under p. Let $q\colon L \to L/S^{(k)}$ denote the canonical projection and set $L_2 = q^{-1}(L_1) \subset L$. Then L_2 contains $S^{(k)}$ and we have an extension

$$0 \to S^{(k)} \to L_2 \xrightarrow{q} L_2/S^{(k)} \cong L_1 \to 0$$

As L_1 is semisimple and $S^{(k)}$ is abelian this extension splits. Let $L_3 \subset L_2$ be a Lie algebra mapping isomorphically to $L_2/S^{(k)}$ under the canonical projection. Then it is easy to check that $L_3 \subset L_2 \subset L$ maps isomorphically to L/S under π, which completes the proof. □

Remark 13.15. *A Lie subalgebra $L' \subset L$ mapping isomorphically to $L/\operatorname{Rad}(L)$ is called a* Levi factor *of L. It is a Theorem of Mal'cev that Levi factors are unique up to automorphism of L.*

Theorem 13.16 (Lie's Third Theorem)**.** *Let L be a Lie algebra over $\mathbb{R}$. Then, there exists a Lie group G with Lie algebra isomorphic to L.*

Proof. Let us first see that if L_1 is a Lie algebra acting on L_2 by derivations and G_1, G_2 are simply connected Lie groups integrating L_1 and L_2 respectively then there is a Lie group integrating $L = L_1 \ltimes L_2$: since G_1 is simply connected, there is a Lie group homomorphism $G_1 \to \operatorname{Aut}(L_2)$ integrating the action map $L_1 \to \operatorname{Der}(L_2)$. By Lie's second Theorem we have $\operatorname{Aut}(L_2) \cong \operatorname{Aut}(G_2)$ and in a previous homework we observed that the corresponding action of G_1 on G_2 by Lie group automorphisms is smooth.

Let $G = G_1 \ltimes G_2$ be the semidirect product determined by this action. This means that $G = G_1 \times G_2$ as a smooth manifold and the product on G is given by the expression

$$(g_1, g_2) \cdot (h_1, h_2) = (g_1 h_1, (h_1^{-1} \cdot g_2) h_2)$$

where $h_1^{-1} \cdot g_2$ denotes the effect on g_2 of the automorphism of G_2 determined by h_1^{-1}. It is an easy exercise to check that this gives G a Lie group structure whose Lie algebra is isomorphic to L.

Now let L be any Lie algebra over $\mathbb{R}$. By the previous discussion and Levi's Theorem it suffices to prove that there is a Lie group integrating L when L is either semisimple or solvable. In the first case $L \cong \operatorname{ad}(L) \subset \operatorname{End}(L)$ is a linear Lie algebra so there is a subgroup of $\operatorname{GL}(L)$ integrating it.

Now suppose L is solvable. Then we can pick a codimension one ideal I in L containing $[L, L]$ (the inverse image under $L \to L/[L, L]$ of any codimension one subspace). Writing $L = I \oplus \mathbb{R}$ we have an extension

$$0 \to I \to L \xrightarrow{p} \mathbb{R} \to 0$$

which splits (one just needs to pick $x \in L$ with $p(x) = 1 \in \mathbb{R}$, the Lie algebra spanned by x will then provide the splitting). Thus $L \cong \mathbb{R} \ltimes I$ for some action of $\mathbb{R}$ on I by derivations. By induction it follows that there exists a Lie group integrating L (moreover the simply connected version of this Lie group is diffeomorphic to $\mathbb{R}^n$) □

Remark 13.17. *The proof above shows that a simply connected Lie group integrating a solvable Lie algebra L will be diffeomorphic to $\mathbb{R}^n$ and, more generally, that any simply* *connected Lie group is diffeomorphic to $\mathbb{R}^n \times H$, with H the simply connected Lie group integrating the semisimple quotient of the original Lie algebra.*

One can further show that the simply connected Lie group integrating a semisimple Lie algebra is diffeomorphic to the cartesian product of a compact Lie groups and an Euclidean space. See [Kn] *for example.*

Compact Lie algebras. We will finish this section by using our understanding of cohomology to characterize the Lie algebras of compact Lie groups.

Theorem 13.18. *Let G be a compact Lie group. Then, $\mathfrak{g} = \mathfrak{a} \times \mathfrak{s}$ with $\mathfrak{a}$ abelian and $\mathfrak{s}$ semisimple with negative definite Killing form.*

Proof. Assume G is connected and consider the canonical extension

$$0 \to Z(G) \to G \to \mathrm{Ad}(G) \to 0$$

with $\mathrm{Ad}(G) \subset \mathrm{Aut}(\mathfrak{g}) \subset GL(\mathfrak{g})$. Since $\mathrm{Ad}(G)$ is compact we can pick an $\mathrm{Ad}(G)$-invariant inner product on $\mathfrak{g}$. With respect to an orthonormal basis for this inner product, the matrices $\mathrm{Ad}(g)$ are orthogonal, so their eigenvalues are of the form $e^{i\theta}$ for some θ. Therefore the matrices $\mathrm{ad}(x)$ for $x \in \mathfrak{g}$ are diagonalizable with purely imaginary eigenvalues iy_α with $y_\alpha \in \mathbb{R}$. It follows that

$$\kappa(x, x) = \mathrm{tr}(\mathrm{ad}(x)\,\mathrm{ad}(x)) = \sum_\alpha (iy_\alpha)^2 = -\sum_\alpha y_\alpha^2 < 0$$

if $\mathrm{ad}(x) \neq 0$. We conclude that κ is negative definite on $\mathrm{ad}(\mathfrak{g})$. By Cartan's criterion for semisimplicity $\mathrm{ad}(\mathfrak{g})$ is semisimple. By the second Whitehead Lemma, the abelian extension

$$0 \to Z(\mathfrak{g}) \to \mathfrak{g} \to \mathrm{ad}(\mathfrak{g}) \to 0$$

splits and, since the action of $\mathrm{ad}(\mathfrak{g})$ on $Z(\mathfrak{g})$ is trivial, this means that

$$\mathfrak{g} = Z(\mathfrak{g}) \times \mathrm{ad}(\mathfrak{g})$$

as required. □

Theorem 13.19. *Suppose $\mathfrak{g} = \mathfrak{a} \times \mathfrak{s}$ is a Lie algebra over $\mathbb{R}$ with $\mathfrak{a}$ abelian and $\mathfrak{s}$ semisimple with negative definite Killing form. Then*

(1) There is a compact Lie group G with Lie algebra $\mathfrak{g}$.
(2) If $\mathfrak{a} = 0$, then the simply connected Lie group integrating $\mathfrak{g} = \mathfrak{s}$ is (and therefore all connected Lie groups integrating $\mathfrak{g}$ are) compact.

Proof. (1) Since $S^1 \times \cdots \times S^1$ is a compact Lie group integrating an abelian Lie algebra $\mathfrak{a}$, it suffices to consider the case when $\mathfrak{a} = 0$. Let G be a connected Lie group integrating $\mathfrak{g}$ and consider the extension

$$0 \to Z(G) \to G \to \mathrm{Ad}(G) \to 0$$

As $\mathfrak{g}$ is semisimple, $Z(\mathfrak{g})$ is trivial and therefore $Z(G)$ is discrete. Since $\mathrm{Ad}(G)$ preserves the inner product determined by the Killing form, it is contained in an orthogonal group $O_\kappa(\mathfrak{g})$, which is compact. Thus we need only show that $\mathrm{Ad}(G)$ is a closed subgrop of $\mathrm{GL}(\mathfrak{g})$. Now, since all derivations of a semisimple Lie algebra are inner derivations (see Remark 13.11), we have

$$\mathrm{Lie}(\mathrm{Ad}(G)) = \mathrm{ad}(\mathfrak{g}) = \mathrm{Der}(\mathfrak{g}) = \mathrm{Lie}(\mathrm{Aut}(\mathfrak{g})),$$

Therefore, $\mathrm{Ad}(G)$ is the connected component of e in the closed subgroup $\mathrm{Aut}(\mathfrak{g})$ and is therefore closed.

(2) Given a connected Lie group G integrating $\mathfrak{g}$ we need to check that the compact Lie group $\mathrm{Ad}(G)$ has finite fundamental group, for then its universal cover, which is the simply connected Lie group integrating $\mathfrak{g}$, will be compact. As $\mathrm{Ad}(G)$ is a topological group, $\pi_1(\mathrm{Ad}(G))$ is abelian. Since $\mathrm{Ad}(G)$ is a compact manifold, its fundamental group must be finitely generated and so

$$\pi_1(\mathrm{Ad}(G)) \cong \mathbb{Z}^k \oplus \mathbb{Z}/n_1 \oplus \cdots \oplus \mathbb{Z}/n_k \tag{16}$$

Now $H^1_{dR}(\mathrm{Ad}(G)) \cong \mathrm{Hom}(\pi_1(\mathrm{Ad}(G)), \mathbb{R})$ (this follows from the deRham Theorem, or see the homework for an elementary proof). Since $\mathrm{Ad}(G)$ is compact, $H^1_{dR}(\mathrm{Ad}(G)) = H^1(\mathrm{ad}(\mathfrak{g}); \mathbb{R})$ and by the first Whitehead Lemma $H^1(\mathfrak{g}; \mathbb{R}) = 0$. We conclude that k in (16) equals zero and hence $\pi_1(\mathrm{Ad}(G))$ is finite.

□

14. Representations of $\mathfrak{sl}(2)$

Having discussed some of the general structure of Lie algebras we'll now concentrate on understanding the structure of semisimple algebras and their representation theory, starting with the simplest (and most important) example.

Let $\mathbb{K}$ be an algebraically closed field of characteristic 0 and consider the Lie algebra $\mathfrak{sl}(2, \mathbb{K})$ with basis

$$h = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad x = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad h = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

The Lie bracket on $\mathfrak{sl}(2, \mathbb{K})$ is determined by the relations

$$[h, x] = 2x, \qquad [h, y] = -2y, \qquad [x, y] = h.$$

Let V be a representation of $\mathfrak{sl}(2, \mathbb{K})$. As $\mathbb{K}$ is algebraically closed, the element $h \colon V \to V$ must have some eigenvector v with eigenvalue $\mu \in \mathbb{K}$. Then

$$\begin{aligned} h \cdot (x \cdot v) &= [h, x] \cdot v + x \cdot h \cdot v \\ &= 2x \cdot v + x \cdot \mu v \\ &= (\mu + 2)(x \cdot v) \end{aligned}$$

This xv will also be an eigenvector for h with eigenvalue $\mu + 2$, i.e. "x raises eigenvalues by 2". An analogous computation will show (exercise) that yv is an eigenvector of h with eigenvalue $\mu - 2$.

Starting with an eigenvector v, consider the span of the set $\{v, xv, \ldots, x^n v, \ldots\}$. Since V is finite dimensional there is a largest integer k such that $\{v, xv, \ldots, x^k v\}$ is linearly independent. As eigenvectors with distinct eigenvalues are linearly independent, it must be that $x^{k+1}v = 0$. The vectors $x^k \cdot v$ have a special role to play in the representation theory of $\mathfrak{sl}(2, \mathbb{K})$.

Definition 14.1. *Let V be an $\mathfrak{sl}(2; \mathbb{K})$-module. An element $v \in V$ is said to be a* highest weight vector *if v is an eigenvector for h and $xv = 0$. The* weight *of an eigenvector for h in an $\mathfrak{sl}(2, \mathbb{K})$-module is the corresponding eigenvalue. The eigenspaces for h will be called* weight spaces.

Proposition 14.2. *Let $v_0 \in V$ be a highest weight vector with weight μ ($x \cdot v_0 = 0$, $hv_0 = \mu v_0$). Define for $i \geq -1$*

$$v_i = \frac{1}{i!} y^i \cdot v_0$$

with the convention that $v_{-1} = 0$. Then

(i) $y \cdot v_i = (i+1)v_{i+1}$;
(ii) $h \cdot v_i = (\mu - 2i)v_i$;
(iii) $x \cdot v_i = (\mu - i + 1)v_{i-1}$;

Proof. Statement (i) is immediate from the definition of the v_i and we have already pointed out (ii) (y lowers the weight by 2). Statement (iii) is true for $i = 0$ because of our convention that $v_{-1} = 0$.

$$\begin{aligned} x \cdot v_i = x \cdot \frac{y \cdot v_i}{i+1} &= \frac{1}{i+1}([x, y]v_i + y \cdot x \cdot v_i) \\ &= \frac{1}{i+1}(h \cdot v_i + y(\mu - i + 1) \cdot v_{i-1}) \\ &= \frac{1}{i+1}((\mu - 2i)v_i + (\mu - i + 1)iv_i) \\ &= \frac{(i+1)\mu - 2i - (i-1)i}{i+1} v_i \\ &= (\mu - i)v_i \end{aligned}$$

□

An important consequence of the previous Proposition is that the *highest weight μ must be a non-negative integer*. Indeed, the argument that gave us the existence of a highest weight vector also gives us the existence of a "lowest weight vector", namely for some $\ell \geq 1$ we will have $v_\ell = 0$ and then condition (iii) will imply that $\mu = \ell - 1$.

For each non-negative integer ℓ we let

$$V(\ell) = \langle v_0, \ldots, v_\ell \rangle$$

denote the $\mathfrak{sl}(2)$-module defined by the equations in Proposition 14.2 (leaving you to check that these equations do indeed define an $\mathfrak{sl}(2)$-module structure). In terms of the given basis, the action of the basis elements of $\mathfrak{sl}(2,\mathbb{K})$ on $V(\ell)$ is given by the following formulas:

$$h = \begin{bmatrix} \ell & & & \\ & \ell-2 & & \\ & & \ddots & \\ & & & -\ell \end{bmatrix} \qquad x = \begin{bmatrix} 0 & \ell & & & \\ & 0 & \ell-1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{bmatrix} \qquad y = \begin{bmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 2 & 0 & & \\ & & \ddots & \ddots & \\ & & & \ell & 0 \end{bmatrix}$$

It is also easy to check that the representations $V(\ell)$ are irreducible: given any nonzero vector $v \in V(\ell)$ for a suitable k, we will have that $x^k v$ is a highest weight vector, and then acting on $x^k v$ with y repeatedly shows that $\mathfrak{sl}(2,\mathbb{K})v = V(\ell)$.

Conversely, if V is an irreducible representation of $\mathfrak{sl}(2,\mathbb{K})$ it will contain a highest weight vector (by the argument explained above) and then since V is irreducible it must equal the representation spanned by the highest weight vector. We have proved the following basic result.

Theorem 14.3. *Let $\mathbb{K}$ be an algebraically closed field of characteristic zero. The irreducible representations of $\mathfrak{sl}(2,\mathbb{K})$ are the representations $V(\ell)$ for each non-negative integer ℓ.*

An arbitrary finite dimensional representation of $\mathfrak{sl}(2,\mathbb{K})$ is isomorphic to a unique representation of the form $V(n_1) \oplus \ldots \oplus V(n_k)$, hence isomorphism classes of representations of $\mathfrak{sl}(2,\mathbb{K})$ are parametrized by tuples of non-negative integers.

It is easy to identify the isomorphism type of a representation of $\mathfrak{sl}(2,\mathbb{K})$. One need only compute the highest weight vectors and identify their weights. The irreducible representations can also be described more concretely.

Example 14.4. *(i) $V(0)$ is the trivial one-dimensional representation.*

(ii) $V(1)$ is clearly the defining representation of $\mathfrak{sl}(2,\mathbb{K})$, $\mathbb{K}^2$.

(iii) The eigenvalues of the action of h on $V(2)$ ar $2, 0, -2$, so $V(2)$ looks just like the adjoint representation of $\mathfrak{sl}(2,\mathbb{K})$. But then it must be the adjoint representation, by Theorem 14.3. We can check this by expressing the adjoint representation in terms of the basis $\langle x, -h, y\rangle$ for $\mathfrak{sl}(2,\mathbb{K})$. We have

$$\mathrm{ad}(h) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix} \qquad \mathrm{ad}(x) = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad \mathrm{ad}(y) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}.$$

Alternatively, $V(2) \cong \mathrm{Sym}^2 V(1)$ can be identified with the space of symmetric 2-tensors on $V(1)$: writing $\langle v_0, v_1\rangle$ for the basis of $V(1)$, a basis for $\mathrm{Sym}^2 V(1)$ is given by

$$\{v_0 \otimes v_0, v_0 \otimes v_1 + v_1 \otimes v_0, v_1 \otimes v_1\} \subset V(1) \otimes V(1)$$

or in more compact notation

$$\{v_0^2, v_0 v_1, v_1^2\}$$

The formula for the action on a tensor product gives us

$$\begin{aligned} h \cdot v_0^2 &= (hv_0)v_0 + v_0(hv_0) = v_0^2 + v_0^2 = 2v_0^2 \\ h \cdot v_0 v_1 &= (hv_0)v_1 + v_0 h v_1 = v_0 v_1 - v_0 v_1 = 0 \\ h \cdot v_1^2 &= 2 v_1 h v_1 = -2v_1^2 \end{aligned}$$

and the action of h alone, identifies $\mathrm{Sym}^2(V(1))$ *as* $V(2)$.

Exercise 14.5. *For an arbitrary non-negative integer ℓ check that $V(\ell)$ is isomorphic to* $\mathrm{Sym}^\ell V(1)$ *(the symmetric ℓ tensors on $V(1)$ which we can think of as homogeneous degree ℓ polynomials in two variables).*

15. The Cartan decomposition of a semisimple Lie algebra

In the next few sections we will see how the very simple structure of the representations of $\mathfrak{sl}(2)$ discussed in the previous section completely determines the structure of an arbitrary semisimple Lie algebra L over an algebraically closed field $\mathbb{K}$ of characteristic zero. For the next few sections L will be a fixed such Lie algebra.

Recall from the discussion preceding Lemma 11.9 that if $x \in \mathfrak{gl}(V)$ then $\mathrm{ad}\colon \mathfrak{gl}(V) \to \mathrm{End}(\mathfrak{gl}(V))$ preserves de Jordan-Chevalley decomposition. When L is a semisimple Lie algebra, the adjoint representation is faithful and the previous result suggests a way of decomposing an element x abstractly into a semisimple and a nilpotent part.

Definition 15.1. *Let L be a semisimple Lie algebra over an algebraically closed field of characteristic zero. An element $x \in L$ is said to be* semisimple *or* nilpotent *if* $\mathrm{ad}(x) \in \mathrm{End}(L)$ *is.*

Proposition 15.2. *Let L be a semisimple Lie algebra over an algebraically closed field of characteristic zero. Given $x \in L$, there exist (necessarily unique) elements $x_s, x_n \in L$ such that*

$$\mathrm{ad}(x_s) = \mathrm{ad}(x)_s \qquad \text{and} \qquad \mathrm{ad}(x_n) = \mathrm{ad}(x)_n$$

Moreover,

- *If $L \subset \mathrm{End}(V)$, x_s and x_n are the semisimple and nilpotent part of the endomorphism $x \in L$.*
- *If $\rho\colon L \to \mathfrak{gl}(V)$ is any representation of L and $x \in L$ then $\rho(x_s) = \rho(x)_s$ and $\rho(x_n) = \rho(x)_n$.*

Proof. We will skip this proof, which uses Weyl's Theorem. See [Hu, Sections 5.4 and 6.4]. □

The elements $x_s, x_n \in L$ are called the semisimple and nilpotent parts of $x \in L$ and the decomposition $x = x_s + x_n$ is called the *abstract Jordan-Chevalley decomposition* of the element x. An element $x \in L$ is *semisimple* if $x = x_s$ and *nilpotent* if $x = x_n$. The previous Proposition tells us moreover that the Jordan-Chevalley decomposition is preserved by arbitrary representations, and that when L is a semisimple Lie subalgebra of $\mathfrak{gl}(V)$, the abstract Jordan-Chevalley decomposition coincides with the usual Jordan-Chevalley decomposition of endomorphisms.

Definition 15.3. *A Lie subalgebra $\mathfrak{t} \subset L$ is said to be a* toral subalgebra *if it consists of semisimple elements. A* Cartan subalgebra *of L is a maximal toral subalgebra.*

Proposition 15.4. *There exist non-zero toral subalgebras. Moreover, a toral subalgebra of L is abelian.*

Proof. There must be some nonzero element in L which is semisimple. Otherwise, $\mathrm{ad}(x)$ would be nilpotent for every $x \in L$ and then, by Engel's Theorem, L would be nilpotent. If $x \in L \setminus \{0\}$ is semisimple then $\mathbb{K}x \subset L$ is a nontrivial toral subalgebra.

Let $\mathfrak{t} \subset L$ be a toral subalgebra. Given $x \in \mathfrak{t}\setminus\{0\}$, $\mathrm{ad}_{\mathfrak{t}}(x)$ is diagonalizable because $\mathrm{ad}(x)(\mathfrak{t}) \subset \mathfrak{t}$ and $\mathrm{ad}(x)$ is diagonalizable by assumption. If $\mathfrak{t}$ is not abelian, $\mathrm{ad}_{\mathfrak{t}}(x)$ has a nonzero eigenvalue, i. e., there is $y \in \mathfrak{t} \setminus \{0\}$ and $a \in \mathbb{K} \setminus \{0\}$ such that $[x, y] = ay$.

As $y \in \mathfrak{t}$ we can pick a basis $\{v_1, \ldots, v_n\}$ of $\mathfrak{t}$ formed by eigenvectors of y. Suppose $[y, v_i] = \lambda_i v_i$ and write $x = c_1v_1 + \cdots + c_nv_n$. Then

$$ay = [x, y] = -c_1\lambda_1 v_1 - \ldots - c_n\lambda_n v_n$$

and hence

$$0 = [y, [x, y]] = -c_1\lambda_1^2 v_1 - \ldots - c_n\lambda_n^2 v_n$$

But then $c_i\lambda_i^2 = 0 \Leftrightarrow c_i\lambda_i = 0$ for every i and hence $y = 0$, which is a contradiction. $\square$

Example 15.5. *It is easy to check that $L = \mathfrak{sl}(n; \mathbb{K})$ is semisimple using the Killing form. Later we will see that is is actually simple. The subspace*

$$\mathfrak{h} = \left\langle \begin{bmatrix} 1 & & & & \\ & -1 & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & 0 \end{bmatrix}, \begin{bmatrix} 0 & & & & \\ & 1 & & & \\ & & -1 & & \\ & & & \ddots & \\ & & & & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & 1 & \\ & & & & -1 \end{bmatrix} \right\rangle$$

is a Cartan subalgebra. Indeed, $\mathfrak{h}$ consists of semisimple elements which commute with each other so it is a toral subalgebra. It is maximal because one easily checks that only diagonal matrices commute with all elements of $\mathfrak{h}$.

If $\mathfrak{h}$ is a Cartan subalgebra, $\mathrm{ad}(\mathfrak{h}) \subset \mathrm{End}(L)$ is a set of commuting diagonalizable matrices so we can decompose L as a sum of common eigenspaces for all elements of $\mathfrak{h}$ simultaneously. If $v \in L$ is an eigenvector for all $h \in \mathfrak{h}$ and we write $\alpha(h)$ for the eigenvalue corresponding to h

$$[h, v] = hv = \alpha(h)v,$$

the function $\alpha\colon \mathfrak{h} \to \mathbb{K}$ is linear:

$$\alpha(h_1 + h_2)v = (h_1 + h_2) \cdot v = h_1v + h_2v = \alpha(h_1)v + \alpha(h_2)v = (\alpha(h_1) + \alpha(h_2))v$$

hence $\alpha(h_1 + h_2) = \alpha(h_1) + \alpha(h_2)$. Similarly we see that α preserves scalar multiplication.

Therefore the common eigenvalues are elements $\alpha \in \mathfrak{h}^*$ in the dual of $\mathfrak{h}$. Given $\alpha \in \mathfrak{h}^*$ we set

$$L_\alpha = \{x \in L\colon [h, x] = \alpha(h)x \text{ for all } h \in \mathfrak{h}\}$$

Of course most of these spaces will be zero. The number of distinct α's such that L_α is nonzero can be at most $\dim L$. A special role is played by L_0 which is the centralizer of $\mathfrak{h}$.

Definition 15.6. *The set*

$$\Phi = \{\alpha \in \mathfrak{h}^* \setminus \{0\} \colon L_\alpha \neq 0\}$$

is called the set of roots *of the Cartan subalgebra* $\mathfrak{h}$. *The spaces* L_α *with* $\alpha \in \Phi$ *are called the* root spaces. *The expression*

$$L = L_0 \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$$

is called the Cartan decomposition *of* L.

Example 15.7. *Let* $L = \mathfrak{sl}(2, \mathbb{K})$ *and consider the usual basis* $\{x, h, y\}$ *where* $h = \operatorname{diag}(1, -1)$. *Then* $\mathfrak{h} = \mathbb{K}h$ *and writing* H *for the dual of* h *we have*

$$\Phi = \{-2H, 2H\} \qquad \text{with } L_{2H} = \mathbb{K}x \quad \text{ and } L_{-2H} = \mathbb{K}y$$

(where x *and* y *are the standard basis elements of* $\mathfrak{sl}(2)$*).*

Example 15.8. *Let* $L = \mathfrak{sl}(3; \mathbb{K})$ *and* $\mathfrak{h} = \langle h_1, h_2 \rangle$ *where* $h_1 = \operatorname{diag}(1, -1, 0)$ *and* $h_2 = \operatorname{diag}(0, 1, -1)$. *Let* e_{ij} *denote the matrix which has every entry* 0 *except for* 1 *at the* ij*-th entry, so that* $h_1 = e_{11} - e_{22}$ *and* $h_2 = e_{22} - e_{33}$. *The elements* h_1, h_2 *and* e_{ij} *with* $1 \leq i \neq j \leq 3$ *form a basis for* L *as a vector space.*

Since

$$[e_{ij}, e_{kl}] = \delta_{jk} e_{il} - \delta_{li} e_{kj}$$

we obtain

$$\begin{array}{llllll} [h_1, e_{12}] = 2e_{12} & [h_1, e_{13}] = e_{13} & [h_1, e_{31}] = -e_{31} & [h_1, e_{21}] = -2e_{21} & [h_1, e_{23}] = -e_{23} & [h_1, e_{32}] = e_{32} \\ [h_2, e_{23}] = 2e_{23} & [h_2, e_{13}] = e_{13} & [h_2, e_{12}] = -e_{12} & [h_2, e_{32}] = -2e_{32} & [h_2, e_{31}] = -e_{31} & [h_2, e_{21}] = e_{21} \end{array}$$

Let $\{H_1, H_2\} \subset \mathfrak{h}^*$ *be the dual basis to* $\{h_1, h_2\}$ *(so* $H_i(h_j) = \delta_{ij}$*). The above computations show that the* e_{ij} *with* $i \neq j$ *are common eigenvectors for* h_1, h_2. *For instance*

$$[h_1, e_{12}] = 2e_{12}, \qquad [h_2, e_{12}] = -e_{12}$$

implies that $[h, e_{12}] = (2H_1 - H_2)(h) e_{12}$. *Similar computations show that the set of roots is*

$$\Phi = \{-H_1 + 2H_2, H_1 - 2H_2, 2H_1 - H_2, -2H_1 + H_2, H_1 + H_2, -H_1 - H_2\}$$

which forms an hexagon in the $\mathfrak{h}^*$ *plane. The Cartan decomposition is*

$$\mathfrak{sl}(3, \mathbb{K}) = \mathfrak{h} \oplus (\mathbb{K}e_{23} \oplus \mathbb{K}e_{32} \oplus \mathbb{K}e_{12} \oplus \mathbb{K}e_{21} \oplus \mathbb{K}e_{13} \oplus \mathbb{K}e_{31})$$

Let us now prove some elementary properties of the Cartan decomposition. Note that the first property already strongly restricts the expression of the Lie bracket on L.

Proposition 15.9. *Let* $\mathfrak{h}$ *be a Cartan subalgebra of* L.

(i) Given $\alpha, \beta \in \mathfrak{h}^*$ *we have* $[L_\alpha, L_\beta] \subset L_{\alpha+\beta}$.
(ii) If $\alpha \in \mathfrak{h}^*$ *is nonzero the elements of* L_α *are nilpotent.*
(iii) If $\alpha + \beta \neq 0$, *then* L_α *and* L_β *are orthogonal with respect to the Killing form.*
(iv) The restriction of the Killing form to L_0 *is non-degenerate.*

Proof. (i) Given $x \in L_\alpha$, $y \in L_\beta$ and $h \in \mathfrak{h}$ we have

$$[h,[x,y]] = [[h,x],y] + [x,[h,y]] = [\alpha(h)x,y] + [x,\beta(h)y] = (\alpha(h)+\beta(h))[x,y]$$

so $[x,y] \in L_{\alpha+\beta}$.

(ii) This follows from (i) as $\mathrm{ad}(x)^k(L_\beta) \subset L_{\beta+k\alpha}$ and, as Φ is finite, $L_{\beta+k\alpha} = 0$ for sufficiently large k.

(iii) Let $x \in L_\alpha$ and $y \in L_\beta$ with α, β such that $\alpha+\beta \neq 0$. Pick $h \in \mathfrak{h}$ such that $(\alpha+\beta)(h) \neq 0$. Then,

$$\alpha(h)\kappa(x,y) = \kappa([h,x],y) = -\kappa(x,[h,y]) = -\beta(h)\kappa(x,y)$$

It follows that $(\alpha(h)+\beta(h))\kappa(x,y) = 0$ and hence $\kappa(x,y) = 0$.

(iv) Given $\alpha \in \Phi$, (iii) tells us that $L_0 \perp L_\alpha$ with respect to the Killing form. If $x \in L_0$ is such that $\kappa(x,y) = 0$ for all $y \in L_0$ then, since $L = L_0 \oplus \oplus_{\alpha\in\Phi} L_\alpha$, we will have $\kappa(x,z) = 0$ for all $z \in L$. As κ is non-degenerate it follows that $x = 0$. We conclude that $\kappa_{|L_0 \times L_0}$ is non-degenerate.

□

Theorem 15.10. *$L_0 = \mathfrak{h}$, i. e., $\mathfrak{h}$ is self-centralizing. Hence the restriction of the Killing form to the Cartan subalgebra $\mathfrak{h}$ is non-degenerate.*

Proof. The proof exploits the fact that the Killing form is non-degenerate on L_0 and the Jordan-Chevalley decomposition. See [Hu, Proposition 8.2] □

Thus the Killing form gives us an isomorphism

$$\kappa\colon \mathfrak{h} \to \mathfrak{h}^*$$

(we still denote it by κ), which sends

$$h \mapsto \kappa(h,\cdot)$$

Using this isomorphism we may regard the set of roots Φ as a subset of $\mathfrak{h}$. We will soon see that we can in fact regard Φ as a subset of Euclidean space and, as such, it is an exceedingly symmetric set, akin to the vertices of a platonic solid in the case of three dimensions. As such it is very rigid and this will lead to the classification of semisimple Lie algebras.

Example 15.11. *Consider again $L = \mathfrak{sl}(3,\mathbb{K})$ (see Example 15.8). You will see in the homework that on $\mathfrak{sl}(n;\mathbb{K})$ the Killing form is given by the expression*

$$\kappa(A,B) = 2n\,\mathrm{tr}(AB)$$

Taking this for granted, in terms of the basis $\{h_1 = \mathrm{diag}(1,-1,0), h_2 = \mathrm{diag}(0,1,-1)\}$ for $\mathfrak{h}$ we see that κ is given by the symmetric matrix

$$6\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

(which defines an inner product on $\mathbb{Q}^2$). The eigenvalues of this matrix are 3 and 1 with corresponding eigenvectors are $(1,-1)$ and $(1,1)$, so $\{h_1-h_2, h_1+h_2\}$ is an orthogonal

basis for $\mathfrak{h}$ with respect to κ. The norms are $\|h_1 - h_2\|^2 = 36$, $\|h_1 + h_2\|^2 = 12$, so an orthonormal basis for $\mathfrak{h}$ is

$$v_1 = \frac{h_1 - h_2}{6} \qquad v_2 = \frac{h_1 + h_2}{2\sqrt{3}}$$

The set $\kappa(\Phi) \subset \mathfrak{h}$ is (in terms of the basis $\{h_1, h_2\}$) given by

$$\{\pm(0, \tfrac{1}{6}), \pm(\tfrac{1}{6}, 0), \pm(\tfrac{1}{6}, \tfrac{1}{6})\}$$

so, in the orthonormal coordinates described above, the image of Φ is the set of vertices

$$\left\{\pm\left(-\frac{1}{2}, \frac{1}{2\sqrt{3}}\right), \pm\left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right), \pm\left(0, \frac{1}{\sqrt{3}}\right)\right\}$$

of a regular *hexagon. Note that the inner product of any pair of these vectors is a rational number.*

We will use the following notation for the isomorphism $\kappa^{-1}\colon \mathfrak{h}^* \to \mathfrak{h}$ determined by the Killing form: For $\alpha \in \mathfrak{h}^*$ we write

$$t_\alpha = \kappa^{-1}(\alpha) \tag{17}$$

so that

$$\alpha = \kappa(t_\alpha, \cdot)$$

We shall now prove some more properties of the Cartan decomposition which together strongly hint at the fact that the set of roots completely determines the Lie product structure on L.

Proposition 15.12. *(i) Φ spans $\mathfrak{h}^*$ and $\Phi = -\Phi$.*

(ii) Given $x \in L_\alpha \backslash \{0\}$, $y \in L_{-\alpha}$, then

$$[x, y] = \kappa(x, y) t_\alpha.$$

Moreover, $\kappa(x, y) \neq 0$ for some y and hence $[L_\alpha, L_{-\alpha}] = \mathbb{K} t_\alpha$.

(iii) Given $\alpha \in \Phi$, $x_\alpha \in L_\alpha \backslash \{0\}$, there exists $y_\alpha \in L_{-\alpha}$ so that setting $h_\alpha = [x_\alpha, y_\alpha]$ we have that $S_\alpha = \mathbb{K}\{x_\alpha, h_\alpha, y_\alpha\} \subset L$ is a Lie subalgebra isomorphic to $\mathfrak{sl}(2)$ (more precisely, $[h_\alpha, x_\alpha] = 2x_\alpha$ and $[h_\alpha, y_\alpha] = -2y_\alpha$). Moreover,

$$h_\alpha = \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}.$$

(iv) $\dim L_\alpha = 1$ *and if $\alpha \in \Phi$, then $c\alpha \in \Phi$ if and only if $c = \pm 1$. In particular the y_α in (iii) is unique given x_α.*

(v) Given $\alpha, \beta \in \Phi$ with $\beta \neq -\alpha$ the subspace of L defined by

$$W_{\alpha,\beta} = \bigoplus_{j \in \mathbb{Z}} L_{\beta + j\alpha}$$

is an irreducible module over the copy S_α of $\mathfrak{sl}(2)$ of point (iii) so there are integers q and r such that

$$W_{\alpha,\beta} = \bigoplus_{j=-r}^{q} L_{\beta+j\alpha}$$

Moreover

$$\langle \beta, \alpha \rangle \overset{def}{=} \beta(h_\alpha) = 2\frac{\kappa(t_\beta, t_\alpha)}{\kappa(t_\alpha, t_\alpha)} = r - q \tag{18}$$

(vi) The bilinear form $\mathbb{Q}\Phi \times \mathbb{Q}\Phi \to \mathbb{K}$ defined by

$$(\alpha, \beta) \mapsto \kappa(t_\alpha, t_\beta)$$

takes values in $\mathbb{Q}$ and determines an inner product on the rational span of the roots $\mathbb{Q}\Phi \cong \mathbb{Q}^{\dim_{\mathbb{K}} \mathfrak{h}}$.

Given $\alpha, \beta \in \mathfrak{h}^*$, the integers $\langle \beta, \alpha \rangle$ defined in (18) are called *Cartan integers*. Note that the pairing $\langle \cdot, \cdot \rangle \colon \mathfrak{h}^* \times (\mathfrak{h}^* \setminus 0) \to \mathbb{K}$ defined by the expression

$$\langle \beta, \alpha \rangle = 2\frac{\kappa(t_\beta, t_\alpha)}{\kappa(t_\alpha, t_\alpha)}$$

is only linear in the first variable.

Proof. (i) If $h \in \mathfrak{h}$ is such that $\alpha(h) = 0$ for all $\alpha \in \Phi$, then $[h, x] = 0$ for all $x \in L_\alpha$ and hence $\mathrm{ad}(h) = 0$, so $h \in Z(L) = \{0\}$.

Let $\alpha \in \Phi$. If $-\alpha \notin \Phi$ then for all $\beta \in \Phi$ we have $\alpha + \beta \neq 0$, so by Proposition 15.9(iii) we have $L_\alpha \perp L_\beta$ for all $\beta \in \Phi \cup \{0\}$, i.e. L_α is contained in the radical of the Killing form, which is 0. This is a contradiction which shows that $-\alpha \in \Phi$.

(ii) By Proposition 15.9(i) we have $[x, y] \in L_0 = \mathfrak{h}$. Moreover

$$\begin{aligned}\kappa([x,y],h) = \kappa(h,[x,y]) = \kappa([h,x],y) = \alpha(h)\kappa(x,y) = \kappa(t_\alpha,h)\kappa(x,y) \\ = \kappa(\kappa(x,y)t_\alpha, h) \quad \text{for all } h \in \mathfrak{h}\end{aligned}$$

Since κ is nondegenerate on $\mathfrak{h}$ it follows that $[x, y] = \kappa(x, y)t_\alpha$. If $\kappa(x, y) = 0$ for all $y \in L_{-\alpha}$ then again we would have that x is perpendicular to L_β for all $\beta \in \Phi \cup \{0\}$, which would contradict the non-degeneracy of the Killing form.

(iii) By *(ii)* we may pick $y \in L_{-\alpha}$ such that $\kappa(x_\alpha, y) = 1$ and $S = \langle x_\alpha, y, t_\alpha \rangle \subset L$ is then a Lie subalgebra satisfying the relations

$$[x_\alpha, y] = \kappa(x_\alpha, y)t_\alpha = t_\alpha, \quad [t_\alpha, x_\alpha] = \alpha(t_\alpha)x_\alpha = \kappa(t_\alpha, t_\alpha)x_\alpha, \quad [t_\alpha, y] = -\kappa(t_\alpha, t_\alpha)y$$

If $\kappa(t_\alpha, t_\alpha) = 0$ then S is a solvable subalgebra and therefore $\mathrm{ad}_L([S, S])$ is nilpotent. It follows that $\mathrm{ad}(t_\alpha)$ is nilpotent. However, being in $\mathfrak{h}$, $\mathrm{ad}(t_\alpha)$ is semisimple and so it must be 0, which is a contradiction.

Setting $y' = \lambda y$ with $\lambda \in \mathbb{K}$ we have

$$[x, y'] = \lambda t_\alpha, \quad [\lambda t_\alpha, x_\alpha] = \lambda\kappa(t_\alpha, t_\alpha)x_\alpha, \qquad [\lambda t_\alpha, y'] = -\lambda\kappa(t_\alpha, t_\alpha)y'$$

so if we take

$$\lambda = \frac{2}{\kappa(t_\alpha, t_\alpha)}, \qquad h_\alpha = \lambda t_\alpha = \frac{2}{\kappa(t_\alpha, t_\alpha)} t_\alpha$$

we see that $y_\alpha = \lambda y'$ together with x_α and t_α satisfy the structure relations of $\mathfrak{sl}(2)$.

(iv) By *(ii)* and Proposition 15.9(i), the subspace

$$V_\alpha = \langle h_\alpha \rangle \oplus \oplus_{\substack{r\alpha\in\Phi \\ r\in\mathbb{K}}} L_{r\alpha} \subset L$$

ia a representation of S_α (the copy of $\mathfrak{sl}(2;\mathbb{K})$ constructed in *(iii)*). The classification of $\mathfrak{sl}(2)$ representations tells us that

$$r\alpha(h_\alpha) = r\kappa(t_\alpha, h_\alpha) = 2r\frac{\kappa(t_\alpha, t_\alpha)}{\kappa(t_\alpha, t_\alpha)}$$

is an integer if $L_{r\alpha} \neq 0$ and this can only happen if $r \in \mathbb{K}$ is a half integer. As the 0 eigenspace for the action of h_α on V_α is one dimensional, there can only be one irreducible summand $V(n)$ with n even in V_α. As L_α is the eigenspace of the action of h_α on V_α corresponding to the eigenvalue 2, it must be one dimensional. Moreover since V_α contains the adjoint representation of S_α (which is isomorphic to $V(2)$) the irreducible component with even highest weight is in fact $V(2)$.

The statement in the previous paragraph holds for any root, so whenever α is a root, 2α can not be a root (otherwise α would have a weight space $L_{2\alpha}$ with weight 4). But then it follows that if α is a root then $\frac{1}{2}\alpha$ can not be a root. This means that the weight space of the eigenvalue 1 in V_α is trivial and hence V_α contains no irreducible summands with odd highest weights, i.e.

$$V_\alpha = L_{-\alpha} \oplus \mathbb{K}h_\alpha \oplus L_\alpha.$$

(v) We saw in point *(iv)* that the spaces $L_{\beta+j\alpha}$ are one dimensional. As $L_{\beta+j\alpha}$ is the weight space of h_α with weight $\beta(h_\alpha) + 2j$ with $\beta(h_\alpha) \in \mathbb{Z}$, the classification of representations of $\mathfrak{sl}(2)$ tells us that $W_{\alpha,\beta}$ is irreducible. If the highest weight of this irreducible representation is n, then writing r and q for the integers in the statement we have

$$\begin{array}{r} \beta(h_\alpha) + 2q = n \\ \beta(h_\alpha) - 2r = -n \end{array} \Rightarrow 2\beta(h_\alpha) = 2r - 2q \Leftrightarrow \beta(h_\alpha) = r - q.$$

(vi) Pick a basis $\alpha_1, \dots, \alpha_n \in \Phi$ for $\mathfrak{h}^*$ (over $\mathbb{K}$), which is possible by *(i)*. Given $\beta \in \Phi$, we can write

$$\beta = c_1\alpha_1 + \cdots + c_n\alpha_n, \qquad \text{with } c_i \in \mathbb{K}$$

Writing (α, β) for $\kappa(t_\alpha, t_\beta)$ we have

$$\frac{(\beta, \alpha_j)}{(\alpha_j, \alpha_j)} = \sum_i c_i \frac{(\alpha_i, \alpha_j)}{(\alpha_j, \alpha_j)}.$$

Since κ is nondegenerate, the matrix $\left[\frac{(\alpha_i,\alpha_j)}{(\alpha_j,\alpha_j)}\right] \in M_{n\times n}(\frac{1}{2}\mathbb{Z})$ is invertible over $\mathbb{Q}$. As the elements $\frac{(\beta,\alpha_j)}{(\alpha_j,\alpha_j)}$ are in $\mathbb{Q}$, it follows that the coefficients c_i must also be rational numbers, so the rational vector space spanned by the roots has dimension equal to $\dim_{\mathbb{K}} \mathfrak{h}$.

Given $\lambda, \mu \in \mathfrak{h}^*$ we have by definition of the Killing form

$$(\lambda, \mu) = \operatorname{tr}(\operatorname{ad}(t_\lambda) \operatorname{ad}(t_\mu)) = \sum_{\alpha \in \Phi} \alpha(t_\lambda)\alpha(t_\mu) = \sum_{\alpha \in \Phi} (\alpha, \lambda)(\alpha, \mu) \tag{19}$$

In particular, for β in Φ

$$(\beta, \beta) = \sum_{\alpha \in \Phi} (\alpha, \beta)^2 \Rightarrow \frac{1}{(\beta, \beta)} = \sum_{\alpha \in \Phi} \frac{1}{4} \langle \alpha, \beta \rangle^2 \in \mathbb{Q}$$

It follows that

$$(\alpha, \beta) = 2\langle \alpha, \beta \rangle (\beta, \beta) \in \mathbb{Q}$$

so the Killing form takes rational values on $\mathbb{Q}\Phi$. The formula (19) shows that $(\beta, \beta) > 0$ for every $\beta \in \Phi$ (as at least one summand in the formula is non-zero by non-degeneracy of κ and (i)). This completes the proof.

□

It is important to note the geometric meaning of the Cartan integers, which is implicit in (v). The set of roots

$$\beta + j\alpha \qquad j = -r, -r+1, \ldots, q$$

is called the α *string through* β. The Cartan integer $\langle \beta, \alpha \rangle = r - q$ gives the position of the root β in this string: β sits $\frac{r-q}{2}$ "steps" to the right of the center of the string. This has the following important consequence: for each $\beta \in \Phi$ we have

$$s_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha \in \Phi,$$

as this is the element which is $\frac{r-q}{2}$ positions to the left of the centre of the string (there are no holes in the string because it is an $\mathfrak{sl}(2)$ representation). Equivalently, we can observe that

$$s_\alpha(\beta)(h_\alpha) = \beta(h_\alpha) - 2\beta(h_\alpha) = -\beta(h_\alpha)$$

must be a weight in the representation of S_α, because $\beta(h_\alpha)$ is.

Now s_α is the reflection on the hyperplane orthogonal to α in $\mathfrak{h}^*$:

$$\begin{aligned} s_\alpha(\alpha) &= \alpha - \langle \alpha, \alpha \rangle \alpha = \alpha - 2\alpha = -\alpha \\ s_\alpha(\beta) &= \beta \text{ if } \kappa(t_\alpha, t_\beta) = 0. \end{aligned}$$

so we see that the roots associated to a Cartan subalgebra form an extremely symmetric subset of a Euclidean space. These sets can be completely classified using Euclidean geometry. We will briefly outline this in the next section.

16. Root systems and Serre's relations.

We have seen in the previous section that the set of roots of a semisimple Lie algebra constitute a very symmetric set of Euclidean space. In this section we will briefly summarize the classification of such objects skipping most proofs.

Definition 16.1. *Let $(E, (\cdot,\cdot))$ be a Euclidean space. A subset $\Phi \subset E\backslash\{0\}$ is called a* root system *if it satisfies the following axioms*

(R1) Φ *is finite and spans* E

(R2) *If* $\alpha \in \Phi$ *and* $t\alpha \in \Phi$ *then* $t = \pm 1$

(R3) *If* $\alpha, \beta \in \Phi$ *then* $\langle \beta, \alpha\rangle \overset{def}{=} 2\frac{(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}$

(R4) *Given* $\alpha, \beta \in \Phi$*, the reflection of* β *on the plane* $P_\alpha = \alpha^\perp$ *preserves* Φ*, i.e.*

$$s_\alpha(\beta) = \beta - \langle \beta, \alpha\rangle\alpha \in \Phi \qquad \textit{for all } \alpha, \beta \in \Phi$$

The rank *of a root system is the dimension of the underlying Euclidean space.*

Our main example of interest is of course the root system associated to a semisimple Lie algebra over an algebraically closed field of characteristic zero. In the previous section we proved that the set of roots Φ in the Euclidean space $(\mathbb{R}\Phi, (\cdot,\cdot))$ (where the inner product is the one given in Proposition 15.12(vi)) is a root system.

Definition 16.2. *Given a root system* $\Phi \subset E$*, the* Weyl group $W(\Phi)$ *is the subgroup of the orthogonal group* $O(E)$ *generated by the reflections* s_α *with* $\alpha \in \Phi$*.*

One basic first observation about the Weyl group is that it is finite. Indeed, as Φ spans E the action of an element of W on E is completely determined by the permutation of the finite set Φ that it induces. More generally, the *automorphism group of the root system*

$$\mathrm{Aut}(\Phi) = \{f \in \mathrm{GL}(E)\colon f(\Phi) = \Phi,\ \langle f(\alpha), f(\beta)\rangle = \langle \alpha, \beta\rangle \text{ for all } \alpha, \beta \in \Phi\}$$

is finite for the same reason. Note that isomorphisms of root systems are not required to be isometries of the underlying Euclidean spaces, but only to preserve the pairing $\langle \cdot,\cdot\rangle$. For instance scalar multiplication by $\lambda \in \mathbb{R} \setminus \{0\}$ is an isomorphism between the root systems Φ and $\lambda\Phi$.

Example 16.3. *The root system associated to* $\mathfrak{sl}(3)$ *is called* A_2 *- see Example 15.11. It is easy to check (exercise) that the Weyl group can be identified with the group of permutations of the set of pairs of opposing vertices in the hexagon. Hence* $W \cong D_3 \cong \Sigma_3$*. The automorphism group* $\mathrm{Aut}(A_2)$ *is the full group of symmetries of the hexagon* D_6*, so* W *is a subgroup of index* 2*, missing the automorphism* $x \mapsto -x$*. See Example 16.6 below for a picture of this root system and also for another example, called* G_2*. In this second example the ratio between the lengths of the roots corresponding to the outer and inner hexagon is* $\sqrt{3}$*.*

Exercise 16.4. *Check that if* Φ *is a root system, then*

$$\Phi^\vee = \left\{ \frac{2\alpha}{(\alpha,\alpha)} \colon \alpha \in \Phi \right\}$$

is also a root system, called the dual root system. *Note that if there are roots of different lengths this is not (or at least not obviously) isomorphic to* Φ.

The axiom R3 has the following important consequences regarding the relative position of the roots: Since

$$\langle \beta, \alpha \rangle = \frac{2(\beta, \alpha)}{(\alpha, \alpha)} = 2\frac{\|\beta\|\|\alpha\|}{\|\alpha\|^2} \cos\theta = 2\frac{\|\beta\|}{\|\alpha\|} \cos\theta$$

(with θ the angle between α and β) we have

$$\langle \beta, \alpha \rangle \langle \alpha, \beta \rangle = 4\cos^2\theta \in \mathbb{Z}$$

so

$$4\cos^2\theta \in \{0, 1, 2, 3, 4\}$$

Suppose for instance that $4\cos^2\theta = 3$. Then, $\cos\theta = \pm\frac{\sqrt{3}}{2}$, i. e., $\theta = \frac{\pi}{6}$ or $\frac{5\pi}{6}$. and

$$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 3$$

Since $\langle \alpha, \langle \beta \rangle$ and $\langle \beta, \alpha \rangle$ must be integers it follows that

$$\{\langle \alpha, \beta \rangle, \langle \beta, \alpha \rangle\} = \{-1, -3\} \text{ or } \{1, 3\}$$

Assuming for definiteness that $\|\beta\| \geq \|\alpha\|$ we must have $\langle \beta, \alpha \rangle = \pm 3$ with the sign according to whether $\theta = \frac{\pi}{6}$ or $\theta = \frac{5\pi}{6}$. In any case we have

$$2\frac{\|\beta\|}{\|\alpha\|}\frac{\sqrt{3}}{2} = 3 \Rightarrow \|\beta\| = \sqrt{3}\|\alpha\|$$

Considering all the possibilities leads to the range presented in Table 1.

$\frac{\|\beta\|}{\|\alpha\|}$ assuming $\|\beta\| \geq \|\alpha\|$	$\langle \alpha, \beta \rangle$	$\langle \beta, \alpha \rangle$	θ
arbitrary	0	0	$\frac{\pi}{2}$
1	1	1	$\frac{\pi}{3}$
	-1	-1	$\frac{2\pi}{3}$
$\sqrt{2}$	1	2	$\frac{\pi}{4}$
	-1	-2	$\frac{3\pi}{4}$
$\sqrt{3}$	1	3	$\frac{\pi}{6}$
	-1	-3	$\frac{5\pi}{6}$

TABLE 1. Possible angles and relative lengths of pairs of roots

One consequence of Table 1 is that

$$(\alpha, \beta) > 0 \Rightarrow \beta - \alpha \in \Phi$$
$$(\alpha, \beta) < 0 \Rightarrow \beta + \alpha \in \Phi$$

Indeed the table tells us that either $\langle \alpha, \beta \rangle$ or $\langle \beta, \alpha \rangle = \pm 1$. Assuming for instance that $(\alpha, \beta) > 0$ and therefore $\langle \alpha, \beta \rangle$ or $\langle \beta, \alpha \rangle$ equal 1 (the other case is similar) we have

$$s_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha = \beta - \alpha \text{ if } \langle \beta, \alpha \rangle = 1$$
$$s_\beta(\alpha) = \alpha - \langle \alpha, \beta \rangle \beta = \alpha - \beta \text{ if } \langle \alpha, \beta \rangle = 1$$

Thus, in any case, we have $\beta - \alpha \in \Phi$.

In general, Table 1 implies that the relative positions of the roots in a root system is severely constrained. For instance with regard to the root system associated to a semisimple Lie algebra, the fact that the maximum value of $\langle \beta, \alpha \rangle$ is 3 implies that the maximum dimension of the irreducible $\mathfrak{sl}(2)$ representations appearing in the adjoint representation is 4, or equivalently, the maximum dimension of an α string through β is 4. Moreover this can only occur for roots whose length ratio is $\sqrt{3}$. The previous statements follow easily from the interpretation of the Cartan integers given at the end of the previous section.

Definition 16.5. *A* base *for the root system Φ is a subset $\Delta = (\alpha_1, \ldots, \alpha_\ell) \subset \Phi$ such that:*

(i) Δ is a vector space basis for E;
(ii) Every $\alpha \in \Phi$ can be written as $\sum_i n_i \alpha_i$ with $n_i \in \mathbb{Z}$ in such a way that the n_i are either all positive or all negative.

It is not immediately clear that bases exist but it is true that every root system has one. A base allows us to partition Φ into a subset of positive and negative roots

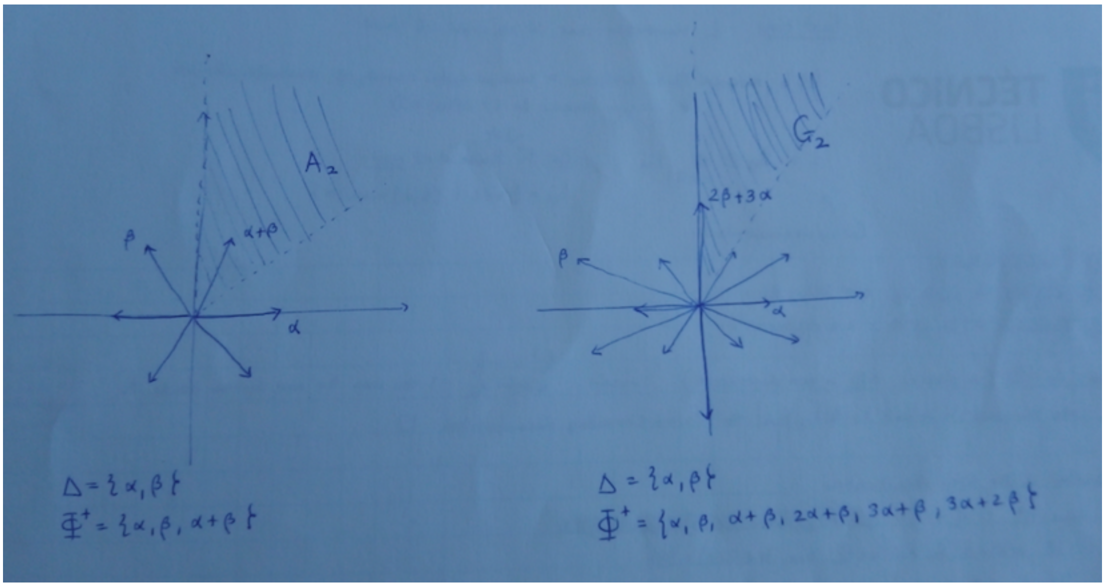
$$\Phi = \Phi^+ \coprod \Phi^-$$

(defined in the obvious way) and introduces a partial order $\prec$ on Φ:

$$\alpha \prec \beta \Leftrightarrow \beta - \alpha \in \Phi^+$$

The elements of a base Δ are called *simple roots* and if $\alpha = \sum n_i \alpha_i$, the integer $\sum_i n_i$ is called the *height* of the root (with respect to the given base).

Example 16.6. *Here are the examples A_2 and G_2 again, with a choice of base $\Delta = \{\alpha, \beta\}$ and the corresponding sets of positive roots indicated.*



The existence of bases follows from an analysis of the geometry of the action of the Weyl group on the root system.

Definition 16.7. *Let $P_\alpha = \alpha^\perp$ be the plane perpendicular to $\alpha \in \Phi$. The connected components of $E \setminus \bigcup_{\alpha\in\Phi} P_\alpha$ are called* Weyl chambers*. The planes P_α are said to be the walls of the Weyl chambers. An element $x \in E$ is* regular *if it does not belong to a wall and* singular *otherwise*

When the root system comes from a complex semisimple Lie algebra (and is regarded as a subset of $\mathfrak{h}$ via the isomorphism given by the Killing form) then the Euclidean space $\mathbb{R}\Phi$ is contained in the Cartan subalgebra $\mathfrak{h}$ and $x \in P_\alpha$ means that $\alpha(x) = 0$. This in turn means that x centralizes the copy S_α of $\mathfrak{sl}(2)$ corresponding to α. Thus the regular elements are those elements in $\mathbb{R}\Phi \subset \mathfrak{h}$ which have $\mathfrak{h}$ as centralizer, whilst singular elements have bigger centralizers.

Given a regular element $x \in E$, we define

$$\Delta^+(x) = \{\alpha \in \Phi \colon (\alpha, x) > 0, \alpha \text{ indecomposable.}\}$$

(indecomposable means that α can not be written as a nontrivial sum of other such roots).

Proposition 16.8. *The assignment*

$$x \to \Delta^+(x)$$

induces a one to one correspondence between Weyl chambers and bases of Φ.

Proof. See [Hu, Section 10.1]. ☐

The Weyl chambers corresponding to the bases indicated are shaded in the picture in Example 16.6. The walls of the Weyl chambers are the lines perpendicular to the simple roots indicated.

We will now state the basic properties of the action of the Weyl group on the root system.

Theorem 16.9. *(1) The Weyl group W acts simply transitively (i.e. freely and transitively) on the Weyl chambers and hence also on the bases.*
(2) Every root is part of some base.
(3) Given a base Δ, W is generated by the simple reflections *$\{s_{\alpha_i} : \alpha_i \in \Delta\}$, which are reflections on the walls of the Weyl chamber corresponding to Δ.*
(4) The closure of a Weyl chamber is a fundamental domain for the action of W on E, i. e., the orbit of each $x \in E$ under the action of W intersects the closure of a Weyl chamber at a single point.

Proof. See [Hu, Section 10.3] □

Definition 16.10. *Let $\Delta = (\alpha_1, \ldots, \alpha_\ell)$ be an ordered base for the root system Φ. The* Cartan matrix *of (E, Φ, Δ) is the $\ell \times \ell$ integer matrix*

$$[\langle \alpha_i, \alpha_j \rangle]$$

Example 16.11. *Referring back to Example 16.6 we see that the Cartan matrix for A_2 is*

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

while the Cartan matrix for G_2, where the angles between two adjacent roots is $\frac{\pi}{6}$ and the ratio between the lengths of long and short roots is $\sqrt{3}$, is

$$\begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}$$

The point of the Cartan matrix is that it encodes all the information contained in the root system. This is the content of the following Proposition.

Proposition 16.12. *The Cartan matrix determines the root system Φ up to isomorphism.*

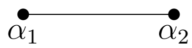
Proof. Given two root systems Φ and Φ' we can pick bases Δ and Δ' and define a linear isomorphism $f : E \to E'$ by extending linearly the assignment $\alpha_i \mapsto \alpha_i'$. By assumption this isomorphism will satisfy $\langle f(\alpha_i), f(\alpha_j) \rangle = \langle \alpha_i, \alpha_j \rangle$. The fact that such an isomorphism will commute with the simple reflections allows us to prove that it will in fact preserve all the brackets $\langle \alpha, \beta \rangle$, as the simple reflections generate W and any root is the image of a simple root by some element of W. See [Hu, Proposition 11.1] for the full details. □

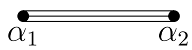
The information contained in the Cartan matrix can be pictorially encoded in a graph.

Definition 16.13. *The* Coxeter graph *of a root system Φ is the graph with one vertex for each simple root and $\langle \alpha_i, \alpha_j \rangle \langle \alpha_j, \alpha_i \rangle \in \{0, 1, 2, 3\}$ edges between the vertices corresponding to the roots α_i and α_j.*

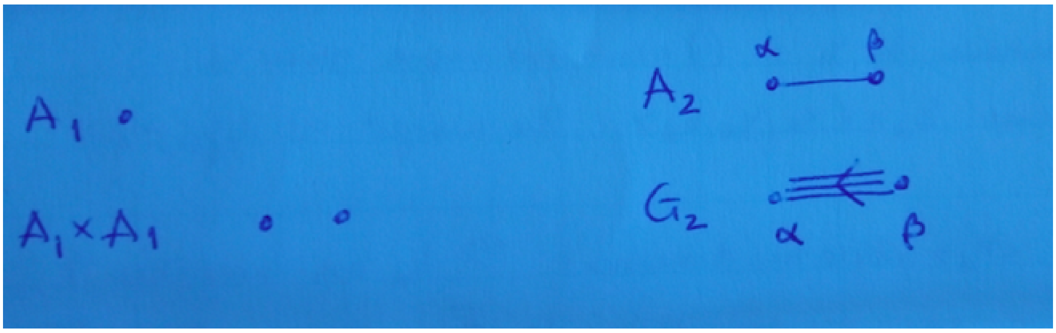
The Dynkin diagram *of Φ is the Coxeter graph together with an orientation of the edges connecting roots of different lengths pointing towards the shorter root.*

Here are the Coxeter graphs of A_2 and G_2:

A_2  $\langle \alpha_1, \alpha_2 \rangle \langle \alpha_2, \alpha_1 \rangle = (-1)^2 = 1$

G_2  $\langle \alpha_1, \alpha_2 \rangle \langle \alpha_2, \alpha_1 \rangle = (-1)(-3) = 3$

and the Dynkin diagrams of the root systems A_1, $A_1 \times A_1$, A_2 and G_2:



Clearly the Dynkin diagram contains exactly the same information as the Cartan matrix: the edges in the diagram give us the set of values $\{\langle \alpha_i, \alpha_j \rangle, \langle \alpha_i, \alpha_j \rangle\}$ as these numbers must be negative (by the discussion following Table 1 since $\alpha_i - \alpha_j$ can't be a root), while the orientation of the edges identifies which of the two numbers is bigger and hence where the numbers appear in the Cartan matrix.

Definition 16.14. *A root system is said to be* irreducible *if it cannot be partitioned into orthogonal subsets or, equivalently, if a base Δ cannot be partitioned into two orthogonal subsets.*

Clearly, any root system Φ can be written as a cartesian product of irreducible root systems. If $L = L_1 \times \cdots \times L_n$ is the decomposition of a semisimple Lie algebra into a cartesian product of simple factors then clearly

$$\Phi(L) = \Phi(L_1) \coprod \cdots \coprod \Phi(L_n)$$

Moreover the $\Phi(L_i)$ will be irreducible. Indeed the orthogonality of two sub root systems precisely translates into the fact that the Lie subalgebras spanned by the corresponding root spaces commute with each other (the sum of two roots one from each set will never be a root).

Proposition 16.15. *Let (E, Φ) be an irreducible root system with base Δ. Then:*

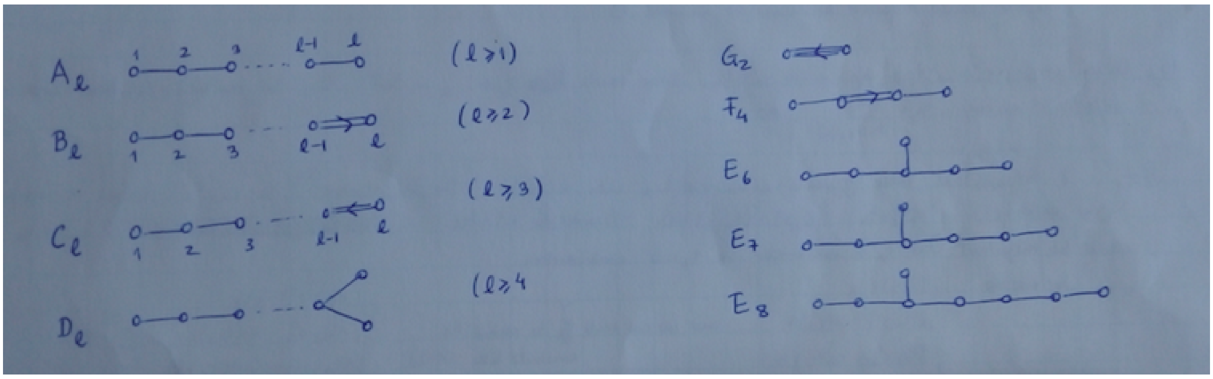
(1) There is a unique maximal root for the partial order $\prec$.
(2) The W-orbit of every $\alpha \in \Phi$ spans E.
(3) There are at most 2 roots lengths and each set of roots with the same length form an orbit of W.
(4) The maximal root is long.

Proof. See [Hu, Section 10.4] $\square$

In Example 16.6 above, the unique maximal roots are $\alpha + \beta$ and $2\beta + 3\alpha$ as indicated in the picture.

We can now state the classification of the irreducible root systems. This classification appears somewhat mysteriously in many areas of Mathematics including the classification of reflection groups, finite subgroups of Lie groups as well as singularities of algebraic varieties and smooth maps.

Theorem 16.16. *An irreducible root system Φ of rank ℓ is isomorphic to one of the following:*



The root systems $A_\ell, B_\ell, C_\ell, D_\ell$ can be shown to be the root systems of the classical Lie algebras, respectively:

- The special linear Lie algebra $\mathfrak{sl}(\ell+1) = \{x \in M_{\ell+1}(\mathbb{K})\colon \operatorname{tr} x = 0\}$
- The orthogonal Lie algebra $\mathfrak{so}(2\ell+1) = \{x \in M_{2\ell+1}(\mathbb{K})\colon x + x^T = 0\}$. This is isomorphic to the Lie algebra of automorphisms of an arbitrary non-degenerate symmetric bilinear form on $\mathbb{K}^{2\ell+1}$
- The symplectic Lie algebra $\mathfrak{sp}(2\ell) = \{x \in M_{2\ell}(\mathbb{K})\colon x^T J + Jx = 0\}$ where $J = \begin{bmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{bmatrix}$. This is the Lie algebra of automorphisms of an arbitrary non-degenerate skew-symmetric bilinear form on $\mathbb{K}^{2\ell}$.
- The orthogonal Lie algebra $\mathfrak{so}(2\ell)$.

You will check this in the homework in the case of the symplectic algebra.

The remaining five root systems are deemed *exceptional* and correspond to the *exceptional Lie algebras* also denoted G_2, F_4, E_6, E_7, E_8. The latter are all in some way related to the existence of the octonions. For instance G_2 is the complexification of the Lie algebra of derivations of the octonions.

Remark 16.17. *As an example of the exceptional symmetry of root spaces, the vertices of F_4 form a 24-cell in $\mathbb{R}^4$ (the exceptional regular solid in 4 dimensions). The E_8-lattice spanned by the roots in $\mathbb{R}^8$ also has many exceptional properties and was very recently proved to provide the highest density spherical packing of 8 dimensional space.*

As an example of how to produce the Cartan matrix from the Dynkin diagram, here is the Cartan matrix for B_ℓ:

$$\begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & \ddots & & & \\ 0 & \ddots & \ddots & & & \\ & & & & -1 & 0 \\ \vdots & & & -1 & 2 & -2 \\ 0 & \cdots & & 0 & -1 & 2 \end{bmatrix}$$

Our assignment of a root system to a semisimple Lie algebra L depended on the choice of a Cartan subalgebra. It may be shown that any two Cartan subalgebras of L differ by an automorphism of L (see [Hu, Sections 15 and 16]). We then have a function assigning a root system to any semisimple Lie algebra. Given the classification stated above one may show that this function is surjective by example, i.e. by identifying a specific Lie algebra corresponding to a root system in the list. This is however not entirely satisfactory.

The previous section together with the discussion of root systems in the present section strongly suggest that the root system associated to a semisimple Lie algebra determines the Lie algebra in question. This can in fact be done. One can see that an isomorphism between root systems arising from semisimple algebras can be promoted to an isomorphism between the Lie algebras (starting with a linear isomorphism between the Cartan subalgebras and extending) - see [Hu, Theorem 14.2]. This then implies that the correspondence between semisimple Lie algebras and root systems is a bijection, giving the classification of semisimple Lie algebras.

However, by far the most satisfactory formulation of the fact that the root system completely encodes a semisimple LIe algebra structure is the following famous Theorem of Serre.

Theorem 16.18 (Serre)**.** *Let Φ be a root system with base $(\alpha_1, \dots, \alpha_\ell)$. The Lie algebra freely generated by generators x_i, y_i, h_i, $i = 1, \dots, \ell$ with relations*

(S1) $[h_i, h_j] = 0$,
(S2) $[x_i, y_i] = h_i$, $[x_i, y_j] = 0$ *for* $i \neq j$,
(S3) $[h_i, x_j] = \langle \alpha_j, \alpha_i \rangle x_j$, $[h_i, y_j] = -\langle \alpha_j, \alpha_i \rangle y_j$,
(S^+_{ij}) $(\operatorname{ad} x_i)^{-\langle \alpha_j, \alpha_i \rangle + 1}(x_j) = 0$,
(S^-_{ij}) $(\operatorname{ad} x_i)^{-\langle \alpha_j, \alpha_i \rangle + 1}(y_j) = 0$

is a semisimple Lie algebra with root system Φ.

We refer to [Hu, Theorem 18.3] for the proof. Note that the above relations are certainly satisfied by a semisimple Lie algebras: picking a Cartan subalgebra and generators x_i of the root spaces L_{α_i} corresponding to a base $\{\alpha_1, \dots, \alpha_n\}$ we can take y_i and h_i to be the elements giving the corresponding copy of $\mathfrak{sl}(2)$. Relations (S1)-(S3) are then clear (note that $[x_i, y_j] = 0$ because $\alpha_i - \alpha_j$ is not a root). As for the exponents in ($S^\pm_{ij}$) they come from the facts that simple roots sit at the end of an α_i string through them and that the Cartan integers are the lengths of such strings.

A word about the meaning of the statement. A free Lie algebra on a vector space V is defined via a universal property like the one characterizing the free associative algebra $T(V)$. In fact, the Poincaré-Birkhoff-Witt Theorem allows us to identify the free Lie algebra on V with the Lie algebra generated by V inside $T(V)$. Imposing relations on a free Lie algebra amounts to quotienting the Lie algebra by the ideal generated by the relations.

Defining an object via generators and relations has the problem that we don't have a very concrete handle on it, but has the important advantage that we know very well how to map from it. In particular we get functoriality on the root system. Given a semisimple Lie algebra with root system Φ, it receives a map from Serre's universal Lie algebra and it is not hard to prove that the map must be an isomorphism.

17. The Weyl-Chevalley normal form. The compact form.

In this section we will explain in much more concrete terms how a semisimple Lie algebra is determined by its root system. This is called the Weyl-Chevalley normal form of a semisimple Lie algebra and it has many applications. We will use it to prove that every complex semisimple Lie algebra is the complexification of a compact semsimple Lie algebra.

Lemma 17.1. *Let Φ be a root system and $\alpha, \beta \in \Phi$ be linearly independent roots. Let*

$$\beta - r\alpha, \ldots, \beta + q\alpha$$

be the α-string through β. Then

(i) If $\alpha + \beta \in \Phi$, then

$$r + 1 = q\frac{(\alpha+\beta, \alpha+\beta)}{(\beta, \beta)}$$

(ii) Given $x_\alpha \in L_\alpha \setminus \{0\}$, let h_α and $x_{-\alpha}$ be such that $x_\alpha, h_\alpha, x_{-\alpha}$ are the standard generators of the copy of $\mathfrak{sl}(2)$ corresponding to α. Then

$$[x_{-\alpha}, [x_\alpha, x_\beta]] = q(r+1)x_\beta$$

Proof. (i) Such α and β span a rank 2 root system so it suffices to check this is the case for the three possibilities given by the classification (note also it is quite easy to classify rank 2 root systems directly - see [Sa, Section 2.7]).

(ii) If $\alpha + \beta \notin \Phi$ then both sides of the equality are zero, as $q = 0$. Otherwise, Since

$$\bigoplus_{j=-r}^{q} L_{\beta + j\alpha}$$

is the irreducible representation of $\mathfrak{sl}(2)$ with highest weight $q+r$, the equality follows from the formulas in Proposition 14.2.

□

Proposition 17.2. *It is possible to choose $x_\alpha \in L_\alpha \setminus \{0\}$ such that*

(i) For all $\alpha \in \Phi$, we have $[x_\alpha, x_{-\alpha}] = h_\alpha$;

(ii) For each $\alpha, \beta \in \Phi$ with $\alpha + \beta \in \Phi$, if we write $[x_\alpha, x_\beta] = c_{\alpha,\beta}x_{\alpha+\beta}$, then

$$c_{-\alpha,-\beta} = -c_{\alpha,\beta}.$$

For any such choice we have

$$c_{\alpha,\beta}^2 = q(r+1)\frac{(\alpha+\beta,\alpha+\beta)}{\beta,\beta} = (r+1)^2 \Rightarrow c_{\alpha,\beta} = \pm(r+1)$$

Proof. The idea is that L has a canonical automorphism called the *Cartan involution* coming from the root system automorphism $\alpha \mapsto -\alpha$. The existence of such an automorphism follows from Serre's Theorem, although for the classical Lie algebras it is easy to produce such an automorphism: it is just $A \mapsto -A^T$.

(i) Letting $\sigma\colon L \to L$ denote an involution sending each root α to $-\alpha$, we can pick $x'_\alpha \in L_\alpha \setminus \{0\}$ and set $x'_{-\alpha} = -\sigma(x'_\alpha) \in L_{-\alpha}$. Then

$$[x'_\alpha, x'_{-\alpha}] = \kappa(x'_\alpha, x'_{-\alpha})t_\alpha \Rightarrow [cx'_\alpha, cx'_{-\alpha}] = c^2\kappa(x'_\alpha, x'_{-\alpha})t_\alpha = c^2\kappa(x'_\alpha, x'_{-\alpha})\frac{(\alpha,\alpha)}{2}h_\alpha$$

Thus, up to sign, there is a *unique scaling* of x'_α which will satisfy (i).

(ii) By definition we have

$$[x_{-\alpha}, x_{-\beta}] = c_{-\alpha,-\beta}x_{-\alpha-\beta}$$

On the other hand, with σ the involution discussed above, we have

$$[x_{-\alpha}, x_{-\beta}] = [\sigma(x_\alpha), \sigma(x_\beta)] = \sigma([x_\alpha, x_\beta]) = \sigma(c_{\alpha,\beta}x_{\alpha+\beta}) = -c_{\alpha,\beta}x_{-\alpha-\beta}$$

and the desired result follows.

As for the value of the structure constants, we have

$$[[x_\alpha, x_\beta], [x_{-\alpha}, x_{-\beta}]] = -[c_{\alpha,\beta}x_{\alpha+\beta}, c_{\alpha,\beta}x_{-\alpha-\beta}] = -c_{\alpha,\beta}^2 h_{\alpha+\beta}$$

On the other hand, writing

$$\beta - r\alpha, \ldots, \beta + q\alpha \qquad \alpha - r'\beta, \ldots \alpha + q'\beta$$

for the α string through β and the β string through α respectively, we have

$$[[x_\alpha, x_\beta], [x_{-\alpha}, x_{-\beta}]] = [x_\alpha, \underbrace{[x_\beta, [x_{-\alpha}, x_{-\beta}]]}_{-q'(r'+1)x_{-\alpha}}] - [x_\beta, \underbrace{[x_\alpha, [x_{-\alpha}, x_{-\beta}]]}_{q(r+1)x_{-\beta}}] = q'(r'+1)h_\alpha - q(r+1)h_\beta$$

Putting everything together we have

$$-c_{\alpha,\beta}^2\left(\frac{2(t_\alpha+t_\beta)}{(\alpha+\beta,\alpha+\beta)}\right) = -\frac{2q'(r'+1)}{(\alpha,\alpha)}t_\alpha - \frac{2q(r+1)}{(\beta,\beta)}t_\beta$$

and as α and β are linearly independent this implies

$$\frac{c_{\alpha,\beta}^2}{(\alpha+\beta,\alpha+\beta)} = \frac{q'(r'+1)}{(\alpha,\alpha)} = \frac{q(r+1)}{(\beta,\beta)}$$

Hence

$$c_{\alpha,\beta}^2 = q(r+1)\frac{(\alpha+\beta,\alpha+\beta)}{(\beta,\beta)} = q(r+1)\frac{(r+1)}{q} = (r+1)^2$$

as required. □

Definition 17.3. *Let L be a semisimple Lie algebra. A basis for L satisfying the requirements of Proposition 17.2 is called a* Weyl-Chevalley basis *for L.*

Theorem 17.4. *Let L be a semisimple Lie algebra over an algebraically closed field of characteristic zero. Let $\Delta = \{\alpha_1, \ldots, \alpha_\ell\}$ be a base for the root system Φ determined by a Cartan subalgebra of L, and let $h_i = h_{\alpha_i}$. Then, with respect to a Weyl-Chevalley basis, the Lie bracket on L satisfies the following relations:*

$$\begin{aligned}
[h_i, h_j] &= 0 \\
[h_i, x_\alpha] &= \langle \alpha, \alpha_i \rangle x_\alpha \\
[x_\alpha, x_{-\alpha}] &= h_\alpha \textit{ is a } \mathbb{Z} - \textit{linear combination of the } h_i \\
[x_\alpha, x_\beta] &= \begin{cases} 0 & \textit{if } \alpha + \beta \notin \Phi \textit{ i.e., if } q = 0 \\ \pm(r+1) x_{\alpha+\beta} & \textit{if } \alpha + \beta \in \Phi \end{cases}
\end{aligned}$$

In particular, the structure constants of L may be chosen to be integers.

Proof. The only think to check is that h_α is a linear combination of the h_i. This follows from the fact that $\{h_\alpha \colon \alpha \in \Phi\}$ is also a root system (the dual root system - see Exercise 16.4) and the fact that the h_i form a base for this dual root system (exercise). □

The previous Theorem has many important consequences. For instance it allows for the classification of the real forms of a complex semisimple Lie algebra L (meaning real Lie algebras M such that $M \otimes \mathbb{C} \cong L$). We will content ourselves with proving that any complex semisimple Lie algebra is the complexification of a compact Lie algebra. See [Sa, Section 2.10] for more information along these lines.

Corollary 17.5. *Let L be a semisimple Lie algebra over $\mathbb{C}$. Then there exists a compact semisimple Lie algebra (i.e. semisimple with negative definite Killing form) U such that $L = U \otimes_{\mathbb{R}} \mathbb{C}$.*

Proof. We will just construct U explicitly from a Weyl-Chevalley basis for L. Let

$$\mathfrak{t} = \mathbb{R}\{ih_1, \ldots, ih_n\}$$

and set

$$u_\alpha = \frac{i}{2}(x_\alpha + x_{-\alpha}), \qquad v_\alpha = \frac{1}{2}(x_\alpha - x_{-\alpha}) \tag{20}$$

for each$\alpha \in \Phi^+$. One easily checks that the vector space U spanned by $\mathfrak{t}$ and the u_α, v_α for all $\alpha \in \Phi^+$ is a real Lie subalgebra of L. It is clear from the definition that $U \otimes_{\mathbb{R}} \mathbb{C} \cong L$.

Given $x = ih + \sum_{\alpha \in \Phi^+} t_\alpha u_\alpha + s_\alpha v_\alpha$ with $t_\alpha, s_\alpha \in \mathbb{R}$ we have

$$\kappa(x, x) = i^2 \|h\|^2 + \sum_{\alpha \in \Phi^+} t_\alpha^2 k(u_\alpha, u_\alpha) + 2 t_\alpha s_\alpha k(u_\alpha, v_\alpha) + s_\alpha^2 k(v_\alpha, v_\alpha)$$

Since $\kappa(x_\alpha, x_\alpha) = 0$, $\kappa(x_\alpha, x_{-\alpha}) = \frac{4}{(\alpha,\alpha)}$ and $\kappa(x_\alpha, x_{\pm\beta}) = 0$ if $\beta \neq \alpha$ we see that

$$\kappa(x, x) = \text{negative} + 0 + \text{negative},$$

and hence κ is negative definite, which shows that U is compact. □

Remark 17.6. *The way to remember the formulas* (20) *is to consider the basic example when* $L = \mathfrak{sl}(n, \mathbb{C})$ *and* $U = \mathfrak{su}(n)$ *is the subspace of skew-hermitian matrices with trace* 0*. Then the standard basis for* $\mathfrak{sl}(n)$ *is a Weyl-Chevalley basis: setting* $h_i = e_{i,i} - e_{i+1,i+1}$ *for* $i = 1, \ldots, n-1$ *we have the Cartan subalgebra* $\mathfrak{h} = \mathbb{C}\{h_1, \ldots, h_n\}$ *and the Cartan decomposition*

$$\mathfrak{sl}(n) = \mathfrak{h} \oplus \bigoplus_{i \neq j} \mathbb{C} e_{ij}$$

where the generators $x_\alpha = e_{ij}$ *for* $i < j$*, and* $x_{-\alpha} = e_{ij}$ *for* $i > j$*. Then*

$$u_\alpha = \frac{i}{2}(e_{ij} + e_{ji}), \qquad v_\alpha = \frac{1}{2}(e_{ij} - e_{ji}) \quad \textit{for } i < j$$

form the standard basis for the skew Hermitian matrices.

18. Representations of semisimple Lie algebras

In this short section we will briefly describe the classification of the irreducible representations of a semisimple Lie algebra L over an algebraically closed field of characteristic zero.

Let V be a representation of L and $\mathfrak{h}$ a Cartan subalgebra and assume fixed a base $\Delta = \{\alpha_1, \ldots, \alpha_\ell\}$ for the root system Φ determined by $\mathfrak{h}$. Recall that $\mathfrak{h}$ acts on V via semisimple elements, so we can decompose V as a direct sum of weight spaces

$$V = \bigoplus_{\lambda \in \mathfrak{h}^*} V_\lambda, \qquad \text{with } V_\lambda = \{v \in V \colon h \cdot v = \lambda(h) v\}$$

Given $v \in V_\lambda$ and $x_\alpha \in L_\alpha$ we have for every $h \in \mathfrak{h}$

$$\begin{aligned} h(x_\alpha v) &= [h, x_\alpha] v + x_\alpha h v = \alpha(h) x_\alpha v + x_\alpha(\lambda(h) v) \\ &= (\alpha(h) + \lambda(h)) x_\alpha v = (\alpha + \lambda)(h) x_\alpha v \end{aligned}$$

Hence

$$L_\alpha V_\lambda \subset V_{\lambda+\alpha}$$

For each $\alpha \in \Phi$, $L_\alpha \oplus L_{-\alpha}$ generates a copy of $\mathfrak{sl}(2)$. As $\lambda(h_\alpha)$ is a weight of V as a representation of this copy of $\mathfrak{sl}(2)$, it is an integer:

$$\lambda(h_\alpha) = \lambda\left(\frac{2t_\alpha}{(t_\alpha, t_\alpha)}\right) = 2\frac{(\lambda, \alpha)}{(\alpha, \alpha)} = \langle \lambda, \alpha \rangle \in \mathbb{Z}$$

Definition 18.1. *The* weight lattice *is the set*

$$\Lambda = \{\lambda \in \mathfrak{h}^* \colon \langle \lambda, \alpha \rangle \in \mathbb{Z} \textit{ for all } \alpha \in \Phi\} \subset \mathfrak{h}^*$$

This is indeed a lattice in $\mathfrak{h}^*$, namely a free abelian group of rank $\dim \mathfrak{h}$ spanning $\mathfrak{h}^*$, as the condition that defines Λ can alternatively be stated as $\lambda(h_{\alpha_i}) \in \mathbb{Z}$ for each element α_i in a base for Φ. Now the h_{α_i} form a base for the dual root system $\Phi^\vee$ (see homework), and hence a basis for $\mathfrak{h}$. It follows that

$$\Lambda = \mathbb{Z}\langle \lambda_1, \ldots, \lambda_\ell \rangle$$

with $\{\lambda_i\}$ the dual basis to $\{h_{\alpha_i}\}$.

Note that the root lattice $\mathbb{Z} \cdot \Phi$ is contained in Λ. If we write $\alpha_i = \sum n_{ij}\lambda_j$ with $n_{ij} \in \mathbb{Z}$ then

$$\langle \alpha_i, \alpha_k \rangle = \sum n_{ij} \langle \lambda_j, \alpha_k \rangle = \sum n_{ij}\delta_{jk} = n_{ik}$$

so the Cartan matrix expresses the base Δ for Φ in terms of the basis $\{\lambda_1, \ldots, \lambda_\ell\}$ of Λ. It follows that the order of the finite abelian group $\Lambda/(\mathbb{Z} \cdot \Phi)$ is the determinant of the Cartan matrix.

Consider the following subalgebra of L (called a *Borel subalgebra*):

$$B = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Phi^+} L_\alpha$$

This is a solvable subalgebra of L. By Lie's theorem, given a finite dimensional representation V of L we can pick a basis for V such that B acts via upper triangular matrices. Then

$$[B, B] = \bigoplus_{\alpha \in \Phi^+} L_\alpha$$

will act via strictly upper triangular matrices. The first basis element will be an eigenvector $v \in V$ for all of B for which

$$[B, B] \cdot v = \left(\bigoplus_{\alpha \in \Phi^+} L_\alpha \right) \cdot v = 0$$

This is called a *highest weight vector* of V. The weight of such a vector is the element $\lambda \in \mathfrak{h}^*$ such that

$$h \cdot v = \lambda(h) \cdot v \qquad \text{for all } h \in \mathfrak{h}$$

By the classification of $\mathfrak{sl}(2)$-modules we must have $\lambda(h_\alpha) \geq 0$ for every $\alpha \in \Phi^+$, or equivalently

$$\lambda(h_{\alpha_i}) = \langle \lambda, \alpha_i \rangle \geq 0 \qquad \text{for } i = 1, \ldots, \ell$$

Definition 18.2. *The set*

$$\Lambda^+ = \{\lambda \in \Lambda \colon \langle \lambda, \alpha \rangle \geq 0 \textit{ for all } \alpha \in \Phi^+\}$$

is called the set of dominant weights*.*

Now if V is an irreducible representation, then it will be generated by such a highest weight vector v (as $L \cdot v$ will be a nontrivial subrepresentation). The following basic Theorem, which generalizes the classification of $\mathfrak{sl}(2)$-modules states that the weight of such a highest weight vector is an arbitrary dominant weight which completely determines the representation.

Theorem 18.3. *Let L be a semisimple Lie algebra over an algebraically closed field of characteristic zero. Fix a Cartan subalgebra $\mathfrak{h} \subset L$ and a base $\{\alpha_1, \ldots, \alpha_\ell\}$ for the root system Φ determined by $\mathfrak{h}$. Then there is a one-to-one correspondence*

$$\{\textit{isomorphism classes of finite dimensional irreducible representations of } L\} \xrightarrow{\cong} \Lambda^+$$

which sends a representation to the weight of its unique (up to scalar) highest weight vector.

Proof. We have already explained how an irreducible representation V must contain a highest weight vector v. It is not hard to show using the PBW theorem that all weights appearing in such a representation must be lower than the highest weight (with respect to the partial order determined by the choice of base). Indeed, the PBW Theorem allows us to decompose the university enveloping algebra $U(L)$ as a vector space as

$$U(L) = U(N^-) \otimes_{\mathbb{K}} U(B)$$

where $N^- = \oplus_{\alpha \in \Phi^-} L_\alpha$. We then have

$$V = U(L)v = U(N^-)v$$

and the monomials on the standard basis elements of N^- will produce from v vectors of lower weight. It follows that the highest weight vector in an irreducible representation is unique up to scalar.

As for the construction of a representation with a given highest weight $\lambda \in \Lambda^+$, we can start by constructing an infinite dimensional representation (called a *Verma module*)

$$V_\lambda = U(L) \otimes_{U(B)} \mathbb{K}v$$

where $U(B)$ acts on the one dimensional space $\mathbb{K}V$ via the weight λ. This will have highest weight λ. The idea is that V_λ is the "largest representation with highest weight λ". One can then show that this representation has a unique irreducible quotient with highest weight λ which is the required irreducible representation corresponding to λ. See [Hu, Sections 20,21] for the complete details. $\square$

Example 18.4. *Referring back to Example 15.11, i.e. $\mathfrak{sl}(3)$, we see that if we pick as a base*

$$\alpha_1 = (\tfrac{1}{2}, -\tfrac{1}{2\sqrt{3}}), \qquad \alpha_2 = (0, \tfrac{1}{\sqrt{3}})$$

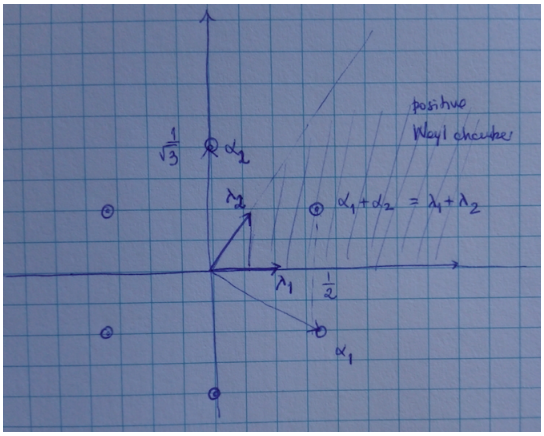
Then the dual root system has base

$$h_{\alpha_1} = (3, -\sqrt{3}), \qquad h_{\alpha_2} = (0, 2\sqrt{3})$$

and hence the dominant weights are

$$\lambda_1 = (\tfrac{1}{3}, 0), \qquad \lambda_2 = (\tfrac{1}{6}, \tfrac{1}{2\sqrt{3}})$$

For instance the highest weight of the adjoint representation (which is $\alpha_1 + \alpha_2$ in terms of the base) is equal to $\lambda_1 + \lambda_2$ in terms of the fundamental dominant weights (see the figure below).



19. Characters of compact Lie groups

Let G be a compact Lie group. We have seen that every representation of G is a direct sum of irreducible representations. In this section we will prove very basic properties of the representations of G in terms of some associated complex valued functions on G called *characters*. Even though the proofs are very easy (they basically come down to repeated applications of Schur's Lemma), the results are very striking and they are already very interesting and useful in the case when G is a *finite* group (in which case the integrals appearing below are finite sums).

For the sake of simplicity we will discuss only complex representations of G, although real and quaternionic representations are also of interest (they can be understood as complex representations with added structure). See [BtD, Section II.6] for more on this.

As we have seen in the case of Lie algebra representations, linear algebra gives us a way of constructing new representations from old. If V is a representation of G, the *dual representation* is the vector space V^* with the G-action given by

$$(g \cdot \varphi)(v) = \varphi(g^{-1} \cdot v) \qquad \text{for } g \in G, \varphi \in V^*, v \in V$$

Given G representations V and W, their *tensor product* is the vector space $V \otimes_{\mathbb{C}} W$ with the action

$$g \cdot (v \otimes w) = (g \cdot v) \otimes (g \cdot w)$$

The space $\mathrm{Hom}(V, W) = V^* \otimes W$ of linear maps from V to W also has a natural G-action (which can be derived from the two described above). It is given by

$$(g \cdot f)(v) = g \cdot f(g^{-1} v) \qquad \text{for } g \in G, f \in \mathrm{Hom}(V, W), v \in V$$

Note that the fixed point space of the G-action on $\mathrm{Hom}(V, W)$ is the set

$$\mathrm{Hom}^G(V, W) = \{f \in \mathrm{Hom}(V, W) \colon f(gv) = gf(v)\}$$

of *G-equivariant maps from V to W*, i.e. the set of maps of representations from V to W.

If V is a complex vector space, the *conjugate vector space* $\overline{V}$ is the complex vector space which has the same underlying set and vector sum as V but where the scalar multiplication is defined by

$$\lambda \cdot v = \overline{\lambda} v$$

(where on the right hand side, juxtaposition denotes scalar multiplication in V).

Recall that a (Hermitian) inner product on a complex vector space is a map

$$\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{C}$$

which is complex linear in the second variable and conjugate linear in the first variable. If $G \times V \to V$ is a representation then the exact same function will define a G-representation $\overline{V}$ (which need not be isomorphic to V; in concrete terms, if we pick a basis for V, the matrix $\overline{\rho}(g)$ representing the action of g on $\overline{V}$ will be the conjugate of the matrix $\rho(g)$ representing the action of g on V).

A Hermitian inner product determines a vector space isomorphism $\overline{V} \to V^*$ via

$$v \mapsto \langle v, \cdot \rangle$$

and clearly a G-invariant Hermitian inner product identifies the dual and conjugate representations.

Proposition 19.1. *Let W be a finite dimensional (complex) representation of (the compact Lie group) G. Let $\widehat{G}$ denote the set of isomorphism classes of irreducible representations of G and V_α be a representative of $\alpha \in \widehat{G}$. Then the evaluation map*

$$\Psi \colon \bigoplus_{\alpha \in \widehat{G}} \mathrm{Hom}^G(V_\alpha, W) \otimes_{\mathbb{C}} V_\alpha \to W$$

defined by $\Psi(f \otimes v) = f(v)$ is an isomorphism of G-representations ($\mathrm{Hom}^G(V_\alpha, W)$ is given the trivial G-action).

Proof. Since $\Psi(g \cdot (f \otimes v)) = \Psi(f \otimes (gv)) = f(gv) = gf(v)$ we see that Ψ is a map of representations. Schur's Lemma 13.7 implies that

$$\mathrm{Hom}^G(V_\alpha, V_\beta) = \begin{cases} \mathbb{C} & \text{if } V_\alpha \cong V_\beta \\ 0 & \text{otherwise} \end{cases}$$

therefore the statement holds when W is irreducible. Since Ψ clearly preserves direct sums and every representation W is a direct sum of irreducible representations, we see that Ψ is an isomorphism for any W. $\square$

We now come to a central definition in the study of representations.

Definition 19.2. *Let $\rho \colon G \to \mathrm{GL}(V)$ be a representation, the* character of ρ *is the function $\chi_\rho \colon G \to \mathbb{C}$ defined by*

$$\chi_\rho(g) = \mathrm{tr}(\rho(g))$$

A representative function for ρ *is a function $f \colon G \to \mathbb{C}$ of the form*

$$f(g) = \varphi(g \cdot v) \qquad \text{for some } v \in V, \varphi \in V^*$$

A representative function of G *is a representative function for some representation ρ of G. The set of representative functions is denoted $\mathcal{T} \subset C^\infty(G; \mathbb{C})$.*

Concretely, if we pick a basis $\{v_1, \dots, v_n\}$ for V, then $\rho(g)$ is represented by a matrix $[a_{ij}(g)]$ and a representative function for ρ will be some complex linear combination of the matrix entries $g \to a_{ij}(g)$. We note that the representative functions $\mathcal{T}$ form a subalgebra of the algebra $C^\infty(G;\mathbb{C})$ of complex valued smooth functions, which is moreover closed under complex conjugation.

Example 19.3. *Consider the Lie group $G = S^1$. A representation $\rho\colon G \to \mathrm{GL}(V)$ is determined by a Lie algebra map $d\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ and since $\mathfrak{g} \cong \mathbb{R}$, this amounts to picking an arbitrary element $A \in \mathrm{End}(V)$ such that $d\rho(1) = A$.*

Since $\exp\colon \mathbb{R} \to S^1$ sends 2π to 1, A is not completely arbitrary, it must satisfy $\exp(2\pi A) = \mathrm{Id}_V$. Conversely given such a matrix,

$$\rho(e^{i\theta}) = \exp(\theta A)$$

will define an S^1 representation.

If we pick an orthonormal basis with respect to some S^1-invariant inner product on V, the matrices $\rho(g)$ will be unitary and hence A will be skew-Hermitian. A skew-hermitian matrices is diagonalizable with purely imaginary eigenvalues, so the representation of $\mathfrak{g}$ (and hence that of G breaks up as a direct sum of one dimensional representation.

If iy is an eigenvalue of A, the condition that $\exp(2\pi A) = \mathrm{id}$ is that the real number y be an integer. We conclude that the irreducible representations are all 1-dimensional and are parametrized by the integers. The representation corresponding to the integer $n \in \mathbb{Z}$ is defined by

$$\rho(e^{i\theta})z = e^{in\theta}z \qquad \text{with } z \in \mathbb{C}$$

It follows that a representative function on S^1 takes the form

$$\sum_\alpha a_\alpha e^{in_\alpha\theta}, \qquad a_\alpha \in \mathbb{C}, n_\alpha \in \mathbb{Z}$$

i. e., $\mathcal{T} \subset C^\infty(S^1, \mathbb{C})$ is the subalgebra of trigonometric polynomials. *Note that in this case, since the irreducible representations are one dimensional, they may be identified with their characters.*

Recall that if X is a measure space, $L^2(X)$ denotes the space of square integrable complex valued function on X, and that we always give a compact Lie group a (bi)-invariant measure for which the volume of G is 1.

Theorem 19.4. *Let V, W be non isomorphic irreducible representations of the compact Lie group G. If $f, h\colon G \to \mathbb{C}$ are representative functions for V and W respectively, then $f \perp h$ in $L^2(G)$. In particular, the characters of V and W are orthogonal in $L^2(G)$.*

Proof. Pick invariant inner products on V and W. Then, representative functions of V and W take the form

$$f(g) = \langle v_1, g \cdot v_2 \rangle, \qquad h(g) = \langle w_1, g \cdot w_2 \rangle, \qquad \text{for some } v_1, v_2 \in V, w_1, w_2 \in W$$

thus we need to show that

$$\int_G \overline{\langle v_1, g \cdot v_2 \rangle} \langle w_1, g \cdot w_2 \rangle dg = 0. \tag{21}$$

For each fixed v_2, w_2, the integral above defines a bilinear map $V \times \overline{W} \to \mathbb{C}$

$$(v_1, w_1) \mapsto \int_G \overline{\langle v_1, g \cdot v_2 \rangle} \langle w_1, g \cdot w_2 \rangle dg$$

The invariance of the inner product and of the integral implies that this map is G-invariant (check). It therefore gives an invariant element in $(V \otimes \overline{W})^* = (\mathrm{Hom}(W, V))^*$, i.e. an element in $\mathrm{Hom}^G(W, V)^*$. However as V, W are not isomorphic and irreducible, Schur's lemma implies that $\mathrm{Hom}^G(W, V) = 0$ and hence the equality (21) holds for all v_1, v_2, w_1, w_2 as required. $\square$

Proposition 19.5 (Elementary properties of characters of compact Lie groups). *(i) If $V \cong W$, then $\chi_V = \chi_W$.*

(ii) $\chi(ghg^{-1}) = \chi(h)$ for all $g, h \in G$. Thus characters are class functions *meaning they are constant in each conjugacy class in G.*

(iii) $\chi_{V \oplus W} = \chi_V + \chi_W$.

(iv) $\chi_{V \otimes W} = \chi_V \chi_W$.

(v) $\chi_{V^*}(g) = \chi_{\overline{V}}(g) = \overline{\chi_V(g)} = \chi_V(g^{-1})$ for all $g \in G$.

(vi) $\chi_v(e) = \dim(V)$.

(vii) $\int_G \chi_V(g) dg = \dim(V^G)$.

(viii) $\langle \chi_V, \chi_W \rangle = \dim \mathrm{Hom}^G(V, W)$.

Proof. $(i) - (vi)$ are left as exercises. For the last equality in (v) note that we may assume that the representation is unitary and hence $\rho(g^{-1}) = \rho(g)^{-1} = \rho(g)^*$, where $*$ denotes the adjoin with respect to the invariant inner product (in terms of matrices it is the conjugate transpose). For (vii) let's check that the map $p \colon V \to V^G$ defined by

$$p(v) = \int_G gv dg$$

is a projection onto V^G: If $v \in V^G$ then $p(v) = \int_G v dg = v$, and left invariance of the integral implies that the image of p is contained in V^G. Now the trace of a projection is the dimension of its range.

As for $(viii)$, we have

$$\langle \chi_V, \chi_W \rangle = \int_G \overline{\chi_V(g)} \chi_W(g) dg = \int_G \chi_{V^* \otimes W}(g) dg = \int_G \chi_{\mathrm{Hom}(V,W)} dg$$

By (vii) this is equal to $\dim(\mathrm{Hom}(V, W))^G = \mathrm{Hom}^G(V, W)$. $\square$

We can now easily see that a representation is completely determined by its character (thus justifying the name).

Corollary 19.6. *Let G be a compact Lie group.*

(i) Two representations V, W are isomorphic if and only if $\chi_V = \chi_W$.

(ii) If $V = \bigoplus V_j^{n_j}$ with V_j irreducible, then

$$\|\chi_V\|^2 = \sum_j n_j^2.$$

Therefore V is irreducible if and only if $\|\chi_V\|_{L^2(G)} = 1$.

Proof. By Proposition 19.5 and Schur's Lemma 13.7 we have for V_i, V_j irreducible

$$\langle \chi_{V_i}, \chi_{V_j} \rangle = \begin{cases} 1 & \text{if } V_i \cong V_j \\ 0 & \text{if } V_i \not\cong V_j \end{cases}$$

Writing $V = \sum V_j^{n_j}$ we have (by Pythagoras' theorem)

$$\|\chi_V\| = \sum_j \|n_j \chi_{V_j}\|^2 = \sum_j n_j^2$$

As $n_j = \langle \chi_V, \chi_{V_j} \rangle$ is determined by the character of V we see that a representation is determined by its character. □

Here is an interesting and useful Corollary of the previous result.

Proposition 19.7. *Let G, H be compact Lie groups. The irreducible representations of the cartesian product $G \times H$ are of the form $V \otimes W$ with V an irreducible representation of G and W an irreducible representation of H.*

Proof. Let V, W be irreducible representations of G and H respectively. Since $(g, h) \cdot (v \otimes w) = (gv) \otimes (hw)$ we have $\chi_{V \otimes W}(g, h) = \chi_V(g)\chi_W(h)$ and therefore

$$\int_{G \times H} |\chi_{V \otimes W}(g, h)|^2 dg dh = \int_G \int_H |\chi_V(g)|^2 |\chi_W(h)|^2 dg dh = (\int_G |\chi_V(g)|^2 dg)(\int_H |\chi_W(h)|^2 dh) = 1$$

where we have used Fubini's Theorem and Corollary 19.6(ii) which then in turn implies that $V \otimes W$ is an irreducible representation of $G \times H$.

Now, let U be an arbitrary representation of $G \times H$. Since U is a representation of G we have an isomorphism of G-representations given by evaluation (see Proposition 19.1)

$$\bigoplus_{\alpha \in \widehat{G}} \mathrm{Hom}^G(V_\alpha, U) \otimes V_\alpha \xrightarrow{\cong} U$$

Since the action of H commutes with the G, $\mathrm{Hom}^G(V_i, U)$ are naturally representations of H and the previous map is actually an isomorphism of representations of $G \times H$ (check). Decomposing each of the summands as a representation of H using Proposition 19.1 again we obtain an isomorphism of H-representations

$$\bigoplus_{\alpha \in \widehat{G}, \beta \in \hat{H}} \mathrm{Hom}^H(W_\beta, \mathrm{Hom}^G(V_\alpha, U)) \otimes V_\alpha) \otimes W_\beta \xrightarrow{\cong} U$$

(where the action of $G \times H$ on the left factors in the tensor product is trivial). But again, the fact that G and H commute means that the map above is actually $(G \times H)$-equivariant. This decomposes any representation of $G \times H$ as a direct sum of $V_\alpha \otimes W_\beta$ proving that these are all the irreducible representations of $G \times H$.

Note that

$$\mathrm{Hom}^H(W_\beta, \mathrm{Hom}^G(V_\alpha, U)) = \mathrm{Hom}^{G \times H}(V_\alpha \otimes W_\beta, U)$$

so the expression obtained above for the decomposition of U is none other than the formula of Proposition 19.1 applied to the Lie group $G \times H$. $\square$

Example 19.8. *Suppose G is a compact abelian group. Then $G = (S^1)^\ell \times \mathbb{Z}/n_1 \times \mathbb{Z}/n_k$ for some non-negative integers ℓ, k and positive integers $n_i \geq 2$.*

A representation of a finite cyclic group $\mathbb{Z}/n$ is the same as a complex matrix A, which we may assume unitary, such that $A^n = \mathrm{id}$. The eigenvalues of such a matrix are (arbitrary) n-th roots of unity so we see that the irreducible representations of $\mathbb{Z}/n$ are one dimensional and classified by $j \in \{0, 1, \ldots, n\}$ with the irreducible representation corresponding to j given by

$$[k] \cdot z = e^{\frac{2\pi jki}{n}} z$$

By Proposition 19.7 and Example 19.3, irreducible representations of G are all one dimensional and are classified by tuples

$$(k_1, \ldots, k_\ell, j_1, \ldots, j_k) \qquad \text{with } k_i \in \mathbb{Z} \text{ and } j_i \in \{0, \ldots, n_i - 1\}$$

An element $\left(e^{i\theta_1}, \ldots, e^{i\theta_\ell}, e^{\frac{2\pi i m_1}{n_1}}, \ldots, e^{\frac{2\pi i m_k}{n_k}}\right)$ *will act by*

$$z \mapsto e^{i\left(k_1\theta_1 + \ldots + k_\ell\theta_\ell + \frac{2\pi m_1 j_1}{n_1} + \ldots + \frac{2\pi m_k j_k}{n_k}\right)} z.$$

20. The Peter-Weyl Theorem

We'll conclude our discussion of the basic representation theory of compact Lie groups with the following fundamental result.

Theorem 20.1 (Peter-Weyl)**.** *Let G be a compact Lie group. Then the algebra $\mathcal{T}$ of representative functions of G is dense in $C^0(G;\mathbb{C})$ with respect to the supremum norm, and hence also in $L^2(G)$ with its usual topology.*

It is not immediately apparent from the statement why this Theorem is so important. This is made more clear by the following *equivalent formulation.*

Theorem 20.2. *Every compact Lie group has a faithful representation.*

Proof. We will prove that the above statement is equivalent to Theorem 20.1. Suppose $\rho\colon G \to \mathrm{GL}(V)$ is a faithful representation of G. Then, given distinct elements $g, g' \in G$ we have $\rho(g) \neq \rho(g')$ and hence $a_{ij}(g) \neq a_{ij}(g')$ for some i, j in some matrix representation $[a_{ij}(g)]$ of $\rho(g)$. This shows that the algebra of representative functions $\mathcal{T} \subset C^0(G;\mathbb{C})$ separates points. Since it is also closed under conjugation, the Stone-Weierstrass Theorem[9] implies that $\mathcal{T}$ is dense in $C^0(G;\mathbb{C})$.

Conversely, suppose that $\mathcal{T}$ is dense in $C^0(G;\mathbb{C})$. Then given $g \in G \setminus \{e\}$ there exists a representative function which takes different values at e and g and hence there is $\rho\colon G \to \mathrm{GL}(V)$ such that $\rho(g) \neq \rho(e)$. Let $G_1 = \ker(\rho)$. Then $G_1 \subsetneq G$ is a closed subgroup. Let

[9] This Theorem is a generalisation of Weierstrass' Theorem to the effect that polynomials are uniformly dense in functions on an interval. The proof is very similar and can easily be done as an exercise. Alternatively, look it up in any book on Real or Functional Analysis.

$g_1 \in G_1 \setminus \{e\}$ and pick $\rho_1 \colon G \to \mathrm{GL}(V_1)$ with $\rho_1(g_1) \neq \rho_1(e)$. Let $G_2 = \ker(\rho \oplus \rho_1 \colon G \to \mathrm{GL}(V \oplus V_1))$. Then $G_2 \subsetneq G_1$. Continuing in this way we can obtain a strictly decreasing sequence of compact Lie groups as long as the representation we have constructed has a nontrivial kernel.

However, if $K \subsetneq G$ is a strict inclusion of compact Lie groups, then either K has fewer components or smaller dimension than G and hence any strictly decreasing sequence of compact Lie groups must terminate. It follows that the procedure of the previous paragraph will end with a faithful representation after a finite number of steps. □

Here are some other Corollaries of Theorem 20.1 for which we regrettably do not have time.

Corollary 20.3. *Let G be a compact Lie group.*

(i) The characters of G are uniformly dense in the space of class functions on G. In particular, when G is finite, the number of irreducible representations of G equals the number of conjugacy classes in G.
(ii) If V is any faithful representation of G then any irreducible representation of G is a summand in some tensor power $V^{\otimes k} \otimes (V^*)^{\otimes l}$.
(iii) If H is a closed subgroup of G, then there is a G-representation V and $v \in V$ such that the isotropy subgroup G_v equals H.

Proof. For *(ii)* we just need to observe that the span of representative functions of the tensor powers will form a subalgebra of $C^0(G;\mathbb{C})$ closed under conjugation and separating points. Hence such functions will be dense in $L^2(G)$, which can only be if all irreducible representations appear as direct summands. See [BtD, Section III.4] for more details and proofs of the other two statements. □

Note that point *(ii)* in the previous corollary allows for pretty concrete descriptions of irreducible representations of compact Lie groups.

In order to prove the Peter-Weyl Theorem we will study the representation of G on the complex valued functions on G. This can take many forms: we can consider for instance the action on the space $C^0(G;\mathbb{C})$ of continuous functions or on the space $L^2(G)$ of square integrable complex valued functions. This representation is called *the regular representation*. It is an infinite dimensional complex representation unless G is finite.

In fact, there are two possible ways of representing G on complex valued functions on G. We can define

$$(g \cdot f)(h) = f(g^{-1} \cdot h) \qquad \text{or} \qquad (g \cdot f)(h) = f(h \cdot g). \tag{22}$$

Luckily the two representations above are isomorphic: if $i \colon G \to G$ is the inverse map defined by $i(g) = g^{-1}$ one easily checks that

$$f \mapsto fi$$

is an isomorphism between the two possible definitions of the regular representation (in any of its forms). However, it is important to notice that the two actions (22) commute with each other so that the regular representation is actually a representation of $G \times G$.

Proposition 20.4 (Characterization of representative functions on G)**.** *Let G be a compact Lie group. A function $f \in C^0(G;\mathbb{C})$ is a representative function of G if and only if the vector space spanned by the action of G on f*

$$\mathbb{C}\{g \cdot f : g \in G\} \subset C^0(G;\mathbb{C})$$

is finite dimensional.

Proof. Given $f \in \mathcal{T}$ let V be a representation of G and $v \in V, \varphi \in V^*$ be such that $f(g) = \varphi(g \cdot v)$. If $\{\varphi_1, \dots, \varphi_n\}$ is a basis for V^*, then there are complex numbers $c_i \in \mathbb{C}$ such that $\varphi = c_1\varphi_1 + \ldots + c_n\varphi_n$. If $h \in G$ we have $(h \cdot \varphi_i)(g \cdot v) = \sum_{j=1}^n a_{ij}(h)\varphi_j(gv)$ where a_{ij} denote the matrix coefficients of the representation V^* with respect to the basis $\{\varphi_1, \dots, \varphi_n\}$. It follows that $\{h \cdot f : h \in G\}$ is contained in the linear span of the representative functions $\{g \mapsto \varphi_i(gv) : i = 1, \dots, n\}$ which is a finite dimensional vector space.

Conversely, suppose that $f \in C^0(G;\mathbb{C})$ is such that $\mathbb{C}\{g \cdot f : g \in G\}$ has basis $\{f_1, \dots, f_n\}$. Taking $g = e$, we see that $f = c_1 f_1 + \ldots + c_n f_n$ for some scalars $c_i \in \mathbb{C}$. The vector space spanned by the action of G on f is clearly closed under the G action and hence a continuous[10] representation of G. Any continuous homomorphism between Lie groups is smooth (this is an exercise with the exponential map - see for instance [Wa, Theorem 3.39]) so this is in fact a smooth representation. Writing

$$g \cdot f_i = \sum_{j=1}^n a_{ij}(g) f_j$$

we see that

$$f_i(g) = (g^{-1} \cdot f)(e) = \sum_{j=1}^n a_{ij}(g^{-1}) f_j(e)$$

The functions $g \mapsto a_{ij}(g^{-1})$ are again representative functions (use for instance the formula for the inverse of a matrix) therefore

$$f(g) = \sum_{i,j=1}^n (c_i f_j(e)) a_{ij}(g^{-1})$$

is also a representative function. □

Proof of Theorem 20.1. Let $\delta : G \to [0, \infty[$ be a smooth function supported on a small neighbourhood of e such that

$$\delta(g) = \delta(g^{-1}) \qquad \text{and} \qquad \int_G \delta(g) dg = 1.$$

Consider the operators

$$K : L^2(G) \to C^0(G;\mathbb{C})$$

[10]This is a slightly subtle point which we leave as an exercise. Note that the action is continuous on $C^0(G;\mathbb{C})$ and the subspace has the induced topology.

which takes the convolution with δ (we omit δ from the notation for K). Thus

$$K(f) = \delta * f \qquad \text{where } (\delta * f)(g) = \int_G \delta(h) f(gh)dh = \int_G \delta(kg^{-1})f(k)dk \tag{23}$$

Note that the second expression for $\delta * f$ in (23) shows (Leibniz's rule) that $K(f)$ is a smooth function on G.

(1) The operators K are continuous linear operators: We have

$$\|\delta * f\|_\infty = \max_{g\in G} |\delta * f(g)|$$

and

$$\begin{aligned} |(\delta * f)(g)| &= |\langle \delta(\cdot g^{-1}), f\rangle_{L^2}| \\ &\leq \|\delta(\cdot g^{-1})\|_{L^2}\|f\|_{L^2} \qquad \text{(by Cauchy-Schwarz)} \\ &= \|\delta\|_{L^2}\|f\|_{L^2} \end{aligned}$$

where in the last equality we have used the invariance of the integral under right translation. We conclude that $\|\delta * f\|_\infty \leq \|\delta\|_{L^2}\|f\|_{L^2}$.

(2) The union of the images of the operators K is dense in $C^0(G;\mathbb{C})$: Since G is compact, a continuous function f is uniformly continuous, i. e., for all $\varepsilon > 0$ there exists $a > 0$ such that $d(g,h) < a \Rightarrow |f(g) - f(h)| < \varepsilon$ (where we use a bi-invariant distance function d on G inducing the topology, for instance the distance determined by a bi-invariant Riemannian metric on G). Now

$$|(\delta * f)(g) - f(g)| = \left|\int_G \delta(h)(f(gh) - f(g))dh\right| \leq \int_{\operatorname{supp}\delta} \delta(h)|f(gh) - f(g)|dh$$

As long as $\operatorname{diam}(\operatorname{supp}(\delta)) < a$, we will have $d(gh, g) < a$ when $\delta(h) \neq 0$ and hence

$$\|K(f) - f\|_\infty = \max_{g\in G} |(\delta * f)(g) - f(g)| \leq \int_{\operatorname{supp}\delta} \delta(h)\varepsilon dg = \varepsilon$$

(3) The operators K are compact: we have to see that if $B \subset L^2(G)$ is bounded then $\overline{K(B)}$ is compact. Since K is continuous, $K(B)$ is certainly uniformly bounded. By Ascoli's theorem it suffices to show that the set $\{K(f)\colon f \in B\}$ is equicontinuous. We have

$$|Kf(g) - Kf(g')| = \left|\int_G (\delta(kg^{-1}) - \delta(kg'^{-1}))f(k)dk\right|.$$

Given $\varepsilon > 0$, since δ is uniformly continuous, there exists $a > 0$ such that $d(g, g') < a \Rightarrow |\delta(kg^{-1}) - \delta(kg'^{-1})| < \varepsilon$. Hence, for all $f \in B$ and all $g, g' \in G$ with $d(g, g') < a$

$$\begin{aligned} |K(f)(g) - K(f)(g')| &= \left|\int_G (\delta(kg^{-1}) - \delta(kg'^{-1}))f(k)dk\right| \\ &\leq \|\delta(\cdot g^{-1}) - \delta(\cdot g'^{-1})\|_{L^2}\|f\|_{L^2} \\ &\leq \varepsilon\|f\|_{L^2} \leq \varepsilon M \end{aligned}$$

with M an upper bound for the L^2 norm on B. We conclude that $K(B)$ is equicontinuous.

(4) $K\colon L^2(G) \to L^2(G)$ is a compact self-adjoint operator: it is compact because $K : L^2(G) \to C^0(G;\mathbb{C})$ is compact and the inclusion $C^0(G;\mathbb{C}) \hookrightarrow L^2(G)$ is continuous. Recall that being selfadjoint means $\langle Kf_1, f_2\rangle = \langle f_1, Kf_2\rangle$ for all $f_1, f_2 \in L^2(G)$. Let us check this identity:

$$\begin{aligned}
\langle Kf_1, f_2\rangle &= \int_G \overline{\left(\int_G \delta(kg^{-1})f_1(k)dk\right)} f_2(g)dg \\
&= \int_{G\times G} \delta(kg^{-1})\overline{f_1(k)}f_2(g)dkdg \\
&= \int_{G\times G} \delta(gk^{-1})\overline{f_1(k)}f_2(g)dkdg \\
&= \langle f_1, Kf_2\rangle,
\end{aligned}$$

where we have used that $\delta(x) = \delta(x^{-1})$. By the spectral theorem for compact self-adjoint operators we have that

$$K = \sum_{n=0}^{\infty} \lambda_n P_n$$

where $P_n\colon L^2(G) \to L^2(G)$ are orthogonal projections onto the eigenspaces of K

$$E_n = \{f \in L^2(G)\colon K(f) = \lambda_n f_n\}$$

Moreover the eigenvalues λ_n are real and $\lim_{n\to\infty} \lambda_n = 0$. Setting $\lambda_0 = 0$ (so that $E_0 = \ker K$), the spaces E_n are all finite dimensional with the possible exception of E_0. Moreover the eigenspaces E_n are all pairwise orthogonal.

(5) K is equivariant, i. e., $K(g \cdot f) = g \cdot K(f)$. This follows immediately the definition of convolution and the left invariance of the integral. It follows that the eigenspaces E_n of K are representations of G. For $n > 0$, we have that $E_n \subset \operatorname{Im} K \subset C^0(G;\mathbb{C})$ are finite dimensional representations and hence by Proposition 20.4 are contained in $\mathcal{T}$.

We can now finish the proof. We see that

$$K(f) = \sum_{n=0}^{\infty} \lambda_n P_n f = \sum_{n=1}^{\infty} \lambda_n P_n f = \lim_{n\to\infty} \sum_{n=1}^{N} \lambda_n P_n f$$

is an L^2 limit of functions in $\mathcal{T}$. But the elements $\sum_{n=1}^{N} \lambda_n P_n f$ belong to the compact set

$$\overline{K(B_{\|f\|_{L^2}}(0))} \subset C^0(G;\mathbb{C})$$

so $\sum_{n=1}^{N} \lambda_n P_n f$ has a convergent subsequence in C^0. Since its L^2 limit is Kf the C^0 limit must also be Kf. We conclude that $\mathcal{T}$ is uniformly dense in the image of K, which completes the proof. $\square$

The example when $G = S^1$ makes it clear that there is a relation between the Peter-Weyl Theorem and classical Fourier analysis. Indeed, in that case, the Fourier series gives a way of expressing an arbitrary square integrable function on L^2 as a limit of representative functions, a.k.a trigonometric polynomials. We will now explain how to generalise this to Fourier decomposition on any compact Lie group, following the discussion in Terry Tao's blog.

Recall that $\widehat{G}$ denotes the set of isomorphism classes of irreducible unitary representations of G. This set is called the *unitary dual* of G. We have included the adjective unitary because this is the standard term, but we note that for a compact Lie group, any representation admits an invariant inner product and is therefore equivalent to a unitary representation. For each $\xi \in \widehat{G}$, let V_ξ be a fixed representative of ξ and consider the map

$$\mathrm{Hom}(V_\xi, V_\xi) = (V_\xi)^* \otimes V_\xi \xrightarrow{i_\xi} L^2(G)$$

determined by

$$\varphi \otimes v \mapsto (g \mapsto \varphi(gv))$$

Both the domain and image of i_ξ have a $G \times G$ action and the map i_ξ is easily checked to be $G \times G$-equivariant:

$$i_\xi((g\varphi) \otimes (g'v))(h) = (g\varphi)(hg'v) = \varphi(g^{-1}hg'v) = (g, g') \cdot \varphi(\cdot v)$$

On $V_\xi^* \otimes V_\xi = \mathrm{Hom}(V_\xi, V_\xi)$ there is a natural $G \times G$ invariant inner product

$$\langle S, T\rangle = \mathrm{tr}(S^*T)$$

where S^* denotes the adjoint of an endomorphism S with respect to the given inner product. With respect to this inner product and the L^2 inner product on the range, the maps i_ξ are almost isometries:

Proposition 20.5. *Given $\xi \in \widehat{G}$, $T, S \in \mathrm{End}(V_\xi)$ we have*

$$\langle i_\xi(T), i_\xi(S)\rangle_{L^2(G)} = \frac{1}{\dim(V_\xi)}\langle f, g\rangle_{\mathrm{End}(V_\xi)}$$

Proof. It is enough to show the identity for T, S of the form $\varphi \otimes v$. Using the inner product on V_ξ we will write a functional φ as $\langle a, \cdot\rangle$ with $a \in V_\xi$ and write a^*b for the endomorphism $v \mapsto \langle a, v\rangle b$. Then one easily checks that $(a^*b)^* = b^*a$ and hence

$$\langle a^*b, c^*d\rangle_{\mathrm{End}(V_\xi)} = \mathrm{tr}(b^*a \circ c^*d) = \mathrm{tr}(\langle b, d\rangle c^*a) = \langle c, a\rangle\langle b, d\rangle$$

On the other hand we have

$$\begin{aligned}
\langle i_\xi(a^*b), i_\xi(c^*d)\rangle_{L^2(G)} &= \int_G \overline{\langle a, gb\rangle}\langle c, gd\rangle dg = \int_G \langle gb, a\rangle\langle c, gd\rangle dg \\
&= \int_G \mathrm{tr}((gb)^*(gd) \circ c^*a)\ dg = \mathrm{tr}\left(\left(\int_G (gb)^*(gd) dg\right) \circ c^*a\right) \\
&= \mathrm{tr}\left(\left(\int_G g \cdot (b^*d) dg\right) \circ c^*a\right)
\end{aligned}$$

Now recall that for $T\colon V \to V$, the expression $\int_G g \cdot T dg$ is a projection of T onto $\mathrm{Hom}^G(V,V)$. It is easy to see that it is in fact the orthogonal projection when we give $\mathrm{End}(V)$ the invariant inner product considered above (or equivalently the invariant inner product determined by an invariant inner product on V). In our case, $\mathrm{Hom}^G(V_\xi, V_\xi) \cong \mathbb{C}\,\mathrm{Id}_{V_\xi}$ by Schur's Lemma and clearly

$$T \mapsto \frac{\mathrm{tr}(T)}{\dim V_\xi}\,\mathrm{Id}_{V_\xi}$$

is the orthogonal projection onto the line spanned by the Id_{V_ξ}. We conclude that

$$\begin{aligned}
\langle i_\xi(a^*b), i_\xi(c^*d)\rangle_{L^2(G)} &= \left\langle \frac{\mathrm{tr}\, b^*d}{\dim V_\xi}\,\mathrm{Id}_{V_\xi}, c^*a \right\rangle_{\mathrm{End}(V_\xi)} = \frac{\mathrm{tr}\, b^*d}{\dim V_\xi}\,\mathrm{tr}\, c^*a \\
&= \frac{1}{\dim V_\xi}\langle b, d\rangle\langle c, a\rangle = \frac{1}{\dim V_\xi}\langle a^*b, c^*d\rangle_{\mathrm{End}(V_\xi)}
\end{aligned}$$

as required. □

Theorem 20.6. *Let G be a compact Lie group. Then the map of $G \times G$ representations*

$$\widehat{\oplus}_{\xi\in\widehat{G}}\left(\sqrt{\dim V_\xi}\ i_\xi\right) : \widehat{\bigoplus_{\xi\in\widehat{G}}} V_\xi^* \otimes V_\xi \to L^2(G)$$

is an isomorphism of Hilbert spaces, where on the left $\widehat{\oplus}$ denotes the L^2-completion of the direct sum.

Proof. The map $\oplus_\xi\sqrt{\dim V_\xi}\ i_\xi$ is a map of $G\times G$ representations. The summands in the domain of the map are by definition orthogonal and then it follows from Proposition 20.5 and Theorem 19.4 that $\oplus_\xi\sqrt{\dim V_\xi}\ i_\xi$ is an isometry onto its image. The Peter-Weyl Theorem says that this isometry, whose image is the algebra $\mathcal{T}$ of representative functions, has dense image. The desired statement follows immediately. □

Remark 20.7. *When G is finite, the previous theorem implies that*

$$\sum_{\xi\in\widehat{G}}(\dim V_\xi)^2 = |G|$$

For instance, the symmetric group Σ_3 of order 6 must have either six one dimensional irreducible representations or two one dimensional and one two dimensional irreducible representations (the latter holds of course).

The inverse of the isomorphism given by the Theorem 20.6 is usually denoted

$$f \mapsto (\hat{f}(\xi))_{\xi\in\widehat{G}} \qquad \text{with } \hat{f}(\xi) \in \mathrm{End}(V_\xi)$$

The elements $\hat{f}(\xi)$ are called *the Fourier coefficients of f along ξ*. In concrete terms the $\hat{f}(\xi)$ are square matrices of dimension $\dim V_\xi$.

The fact that the map in the Theorem 20.6 is an isomorphism of Hilbert spaces implies the *Plancherel identity*

$$\sum_{\xi\in\widehat{G}} \|\hat{f}(\xi)\|^2 = \|f\|^2_{L^2(G)}$$

The component $\hat{f}(\xi)$ is computed by the adjoint $\sqrt{\dim V_\xi}\, i_\xi^*$ of the map

$$\sqrt{\dim V_\xi}\, i_\xi \colon \operatorname{End}(V_\xi) \to L^2(G)$$

Exercise 20.8. *Check that for $f \in L^2(G)$ and $\xi \in \widehat{G}$ we have*

$$i_\xi^*(f) = \int_G f(g)\rho_\xi(g^{-1})dg \tag{24}$$

The expression in the previous exercise for the Fourier coefficients is called *the Fourier inversion formula.* Another standard property of Fourier series that is still valid is the interpretation of convolution in terms of the Fourier coefficients.

Exercise 20.9. *Given $f_1, f_2 \in L^2(G)$, the* convolution *of f_1 with f_2 is the function $f_1 * f_2 \in L^2(G)$ defined by the expression*

$$(f_1 * f_2)(x) = \int_G f_1(xg^{-1})f_2(g)dg$$

Prove that $(\widehat{f_1 * f_2})(\xi) = \hat{f}_1(\xi) \circ \hat{f}_2(\xi)$.

Example 20.10.
1. *When $G = S^1$, Theorem 20.6 gives the usual decomposition of a complex valued function on the circle as a Fourier series. The map in the Theorem computes the Fourier series given a sequence $\hat{f}(n) \in \ell^2(\mathbb{Z})$ and (24) is the usual formula for the Fourier coefficients of a function.*
2. *Let $G = S^3 \cong SU(2)$. Since $SU(2)$ is simply connected, the representations of $SU(2)$ are the same as those of $\mathfrak{su}(2)$. The complex representations of this Lie algebra are the same as those of its complexification*

$$\mathfrak{su}(2) \otimes \mathbb{C} \cong \mathfrak{sl}(2;\mathbb{C})$$

Thus $\widehat{SU}(2) = \mathbb{N}_0$, with the non-negative integer n corresponding to the representation of $SU(2)$ on the degree n homogeneous polynomials on $\mathbb{C}^2$. Given $f \colon S^3 \to \mathbb{C}$, its n-th Fourier coefficient is

$$\hat{f}(n) = \sqrt{n+1}\int_{S^3} f(g)\rho_n(g^{-1})dg \in \operatorname{End}(\mathbb{C}^{n+1})$$

For instance, when $n = 1$, ρ_1 is the standard representation of S^3 on $\mathbb{H} \cong \mathbb{C}^2$ by left multiplication and so

$$\hat{f}(1) = \sqrt{2}\int_{S^3} f(q)\bar{q}\; dq \in \operatorname{End}(\mathbb{C}^2)$$

21. The structure of compact Lie groups

Definition 21.1. *Let G be a compact Lie group. A* maximal torus *of G is a Lie subgroup of G isomorphic to $T = S^1 \times \cdots \times S^1$ which is not contained in any strictly larger subgroup of the same form.*

Alternatively, a maximal torus can be defined as *maximal connected abelian subgroup of* G. Indeed, since the closure of a connected abelian group is still connected and abelian, such a group must be closed, hence compact, and connected hence a torus (which must be maximal).

Maximal tori always exist because given $X \in \mathfrak{g}$

$$\overline{\{\exp(tX) \,:\, t \in \mathbb{R}\}} \quad \text{is a torus}$$

and any torus is contained in a maximal torus (if T, T' are tori and $T \subsetneq T'$, then $\dim T < \dim T'$ therefore any strictly increasing chain of tori in a compact Lie group is finite).

One useful thing about tori is that they have *topological generators*: if $\alpha_1, \ldots, \alpha_n$ are rationally independent irrational numbers then the subgroup of $T = (S^1)^n$ generated by the element $(e^{2\pi i\alpha_1}, \ldots, e^{2\pi i\alpha_n}) \in T$ will have closure equal to T. This is certainly familiar when $n = 1$ and, in general, follows by induction on n (exercise). This concept often allows us to deal with the whole torus focusing on a single element in the torus. For instance if $t \in T$ is a topological generator for a maximal torus T and t belongs to some other torus T' then we must have $T = T'$ (because the closure of the subgroup generated by t will be contained in the torus T').

Definition 21.2. *Let G be a compact Lie group and $T \subset G$ a maximal torus. The* Weyl group *of G is*

$$W = N_G(T)/T$$

where $N_G(T) = \{g \in G \colon gTg^{-1} = T\}$ is the normalizer of T.

It is not immediately clear that the above definition makes sense as different maximal tori could give rise to different Weyl groups. The main Theorem on maximal tori which we will state shortly will imply that all maximal tori in G are conjugate and therefore the isomorphism class of the Weyl group of a compact Lie group G is well defined.

Note that there is a natural map

$$W \to \mathrm{Aut}(T)$$

sending gT to the automorphism $c_g \colon T \to T$ given by conjugation by g. One can check that this map is injective (see [BtD, Corollary IV.2.4]) so that in this way we can regard the Weyl group as a group of automorphisms of T. It follows that W *is a finite group* as W is clearly a compact group and one easily checks that $\mathrm{Aut}(T)$ is isomorphic to $\mathrm{GL}(n, \mathbb{Z})$ (exercise) and hence discrete.

Example 21.3. *(i) Let $G = U(n)$. Then*

$$T = \left\{ \begin{bmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n} \end{bmatrix} : \theta_1, \ldots, \theta_n \in \mathbb{R} \right\}$$

(the subset of diagonal matrices) is a maximal torus. It is even a maximal abelian subgroup as it is easy to check that a matrix which commutes with all matrices in T must be diagonal. The normalizer of T is the group generated by T together with the permutation matrices (check). Thus the Weyl group W is isomorphic to the symmetric group on n elements. It acts on the torus T by permuting the diagonal entries. In this case the Weyl group can be regarded as a subgroup of G (the subgroup of permutation matrices) but that is not always the case.

(ii) Let $G = SU(n)$. Then

$$T = \left\{ \begin{bmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n} \end{bmatrix} : e^{i\theta_1 + \ldots + i\theta_n} = 1 \right\}$$

is a maximal torus (check that the diagonal matrices form a maximal abelian subgroup). The Weyl group is still isomorphic to Σ_n and acts on T by permuting the diagonal entries. Unlike in the previous example, it is now not possible to regard W as being naturally embedded in G.

(iii) Let $G = SO(2n)$. Then

$$T = \left\{ \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 & & & \\ \sin\theta_1 & \cos\theta_1 & & & \\ & & \ddots & & \\ & & & \cos\theta_n & -\sin\theta_n \\ & & & \sin\theta_n & \cos\theta_n \end{bmatrix} \right\}$$

is again a maximal abelian subgroup. One can show that the Weyl group W is isomorphic to $\Sigma_n \ltimes (\mathbb{Z}/2)^n$, where Σ_n acts on the torus by permuting the block diagonal matrices while $(\mathbb{Z}/2)^n$ acts by switching the signs of the θ_i (see [BtD, Section IV.3]*).*

The following result is the main Theorem concerning maximal tori and is absolutely fundamental to understanding the structure of compact Lie groups

Theorem 21.4. *Let G be a compact Lie group.*

(i) Every $g \in G$ belongs to some maximal torus.
(ii) All maximal tori in G are conjugate, i.e. given maximal tori T, T', there exists $g \in G$ such that $gTg^{-1} = T$.

Note that in the case of classical compact Lie groups like the orthogonal or unitary groups, the statement is familiar from basic linear algebra. It comes down to the fact that orthogonal or unitary matrices are diagonalisable.

Here is an important immediate consequence of Theorem 21.4

Corollary 21.5. *If G is a compact Lie group, the exponential map* $\exp\colon \mathfrak{g} \to G$ *is surjective.*

Proof. The exponential map is certainly surjective on tori and every element of G is contained in some torus. $\square$

Corollary 21.5 is in fact equivalent to Theorem 21.4(i) because if we know that exp is surjective[11], then given $g \in G$ we may find $X \in \mathfrak{g}$ with $\exp(X) = g$ and then

$$\overline{\{\exp(tX)\colon t \in \mathbb{R}\}}$$

is a torus containing g and this torus will be contained in some maximal torus.

We also note that the homework includes a proof of Theorem 21.4(ii) based on Lie algebras.

Here is another fundamental consequence for the representation theory of compact Lie groups.

Corollary 21.6. *Let G be a compact Lie group and T be a maximal torus in G. Then every conjugacy class of G intersects T. In particular, any representation ρ of G is determined by its restriction $\rho_{|T}$ to the maximal torus.*

Proof. Given $g \in G$, we have that $g \in T'$ for some maximal torus T'. Let $h \in G$ be such that $hT'h^{-1} = T$, then $hgh^{-1} \in T$ so the conjugacy class of g intersects T.

The second statement follows because a representation ρ is determined by its character $\chi_\rho(g) = \operatorname{tr}(\rho(g))$ which is constant on conjugacy classes. Thus χ_ρ is determined by its restriction to T and this restriction is exactly the character of $\rho_{|T}$. $\square$

Theorem 21.4 implies more specifically that the canonical map

$$T/W \to \{\text{conjugacy classes in } G\}$$

is surjective (and hence a quotient map of topological spaces). In particular the character of a representation of G will be a character of T which is invariant under the action of the Weyl group. However, one can show more precisely that the previous map is a bijection (and hence a homeomorphism - see [BtD, Lemma IV.2.5]).

Example 21.7. *Recall from Example 20.10(ii) that the irreducible complex representations of $SU(2)$ are parametrized by the non-negative integers, and we can take the irreducible representation $V(n)$ corresponding to n to be the vector space of degree n homogeneous polynomials on $\mathbb{C}^2$.*

A maximal torus of $SU(2)$ is

$$T = \left\{ \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

[11] Recall by the way that this follows from the Hopf-Rinow Theorem once one identifies the exponential map with the geodesic exponential for a bi-invariant metric on G.

The standard basis $z_1^k z_2^{n-k}$ *for the homogeneous polynomials consists of eigenvectors for the action of* T *(the element* $z_1^k z_2^{n-k}$ *is acted on with weight* $k-(n-k)=2k-n$*) and we obtain the following expression for the restriction of the character* χ_n *of* $V(n)$ *to* T

$$\chi_0(e^{i\theta}) = 1, \quad \chi_1(e^{i\theta}) = e^{i\theta} + e^{-i\theta}, \quad \chi_2(e^{i\theta}) = e^{2i\theta} + 1 + e^{-2i\theta}$$

and, more generally,

$$\chi_n(e^{i\theta}) = e^{in\theta} + e^{i(n-2)\theta} + \cdots + e^{-in\theta}$$

The action of the Weyl group on the torus switches the sign of the angle and the expressions above are indeed invariant under this action.

The (complex) representations of $SO(3)$ *are the representations of* $SU(2)$ *that send* $\pm\,\mathrm{Id}$ *to the identity matrix, so these are exactly the* $V(2n)$ *with* $n \geq 0$*.*

Let us now discuss the relation between compact Lie groups and complex semisimple Lie algebras. Recall that if $\mathfrak{g}$ is the Lie algebra of a compact Lie group G, we have proved that

$$\mathfrak{g} = \mathfrak{z} \times \mathfrak{k}$$

with $\mathfrak{z}$ abelian and $\mathfrak{k}$ a semisimple Lie algebra (over $\mathbb{R}$) with negative-definite Killing form. (moreover $\mathfrak{z}$ is the center of $\mathfrak{g}$). Then

$$\mathfrak{g} \otimes \mathbb{C} = (\mathfrak{z} \otimes \mathbb{C}) \times (\mathfrak{k} \otimes \mathbb{C})$$

with $\mathfrak{k} \otimes \mathbb{C}$ a complex semisimple Lie algebra. For instance, when $G = U(n)$ we have the decomposition $\mathfrak{g} = \mathbb{R} \times \mathfrak{su}(n)$ with $\mathbb{R}$ the Lie algebra of the circle corresponding to the multiples of the identity, and $\mathfrak{g} \otimes \mathbb{C} \cong \mathbb{C} \times \mathfrak{sl}(2, \mathbb{C})$.

It is easy to check (see homework) that writing $\mathfrak{t}$ for the Lie algebra of a maximal torus T, we have

$$\mathfrak{t} \otimes \mathbb{C} = (\mathfrak{z} \otimes \mathbb{C}) \times \mathfrak{h}$$

with $\mathfrak{h}$ a Cartan subalgebra of complex semisimple Lie algebra $\mathfrak{k} \times \mathbb{C}$.

The adjoint action of T on $\mathfrak{g} \otimes \mathbb{C}$ will be trivial on $\mathfrak{z} \otimes \mathbb{C}$ but on its complement $\mathfrak{k} \otimes \mathbb{C}$ it will be equivalent to the adjoint action of $\mathfrak{h}$ on $\mathfrak{k} \otimes \mathbb{C}$. The weights of this T-action will define a root system in $\mathfrak{t}^*$ (the root system will not in general span $\mathfrak{t}^*$ but only $\mathfrak{k}^*$, the annihilator of $\mathfrak{z}$ inside $\mathfrak{t}^*$) which can be identified with the root system of the corresponding complex semi-simple Lie algebra. One can then identify the Weyl group of the compact Lie group with a reflection group acting on the root system and hence identify the compact Lie group Weyl group with the Weyl group of the associated complex semisimple Lie algebra.

The representation theory of the compact connected Lie groups can be understood in terms of the representation theory of compact semisimple Lie algebras (and the extremely simple representation theory of tori): we begin by observing that there is a "largest" compact Lie group integrating a compact Lie algebra

$$\mathfrak{g} = \mathfrak{z} \times \mathfrak{k},$$

namely

$$T^n \times K$$

where T^n is a torus with dimension $n = \dim_{\mathbb{R}} \mathfrak{z}$ and K is the (unique) simply connected Lie group integrating $\mathfrak{k}$. Indeed, the simply connected Lie group integrating $\mathfrak{g}$ is $\mathbb{R}^n \times K$. Any other Lie group integrating $\mathfrak{g}$ is a quotient $(\mathbb{R}^n \times K)/A$ with A a discrete central subgroup. Writing $\pi\colon \mathbb{R}^n \times K \to \mathbb{R}^n$ for the projection, we have that $\pi(A)$ is a free abelian group isomorphic to $\mathbb{Z}^k$ for some k (and if the quotient is to be compact k must equal n). As $\pi(A)$ is free, the short exact sequence

$$0 \to \ker \pi_{|A} \to A \xrightarrow{\pi} \pi(A) \to 0$$

splits, so $A = \mathbb{Z}^k \times D$ with $D = \ker \pi_{|A} = A \cap K$ a finite central subgroup of K.

In conclusion, the Lie groups integrating $\mathfrak{g}$ take the form $\mathbb{R}^{n-k} \times T^k \times (K/D)$, and the compact Lie groups integrating $\mathfrak{g}$ take the form $T^n \times (K/D)$. In particular, $T^n \times K$ is the "largest" compact subgroup integrating $\mathfrak{g}$.

The irreducible representations of the product $T^n \times K/D$ are tensor products of irreducible representations of the factors (by Proposition 19.7) and the irreducible representations of T^n are one dimensional and classified by their weights in $\mathbb{Z}^n$ (see Example 19.8), so it suffices to describe the irreducible representations of K/D.

The irreducible representations of the compact Lie group K are in $1-1$ correspondence with those of the complex semisimple Lie algebra $\mathfrak{k} \otimes \mathbb{C}$, and are classified by their highest weight. A representation of K/D is exactly a representation of K which sends D to the identity linear transformation, from which one immediately sees that *the irreducible representations of K/D are the irreducible representations of K where D acts trivially.* One can show that a finite central subgroup D is contained in every maximal torus (in fact the center of a compact Lie group is the intersection of all its maximal tori - see [BtD, Theorem 4.2.3]) so one can easily check whether an irreducible representation of K descends to K/D using the restriction of its character to the maximal torus.

Example 21.8. *Consider the group $U(2)$. Its Lie algebra $\mathfrak{u}(2)$ is isomorphic to $\mathbb{R} \times \mathfrak{su}(2)$ and so the largest compact Lie group with Lie algebra $\mathfrak{u}(2)$ is $S^1 \times SU(2)$. The covering $S^1 \times SU(2) \to U(2)$ is a double covering given by the expression*

$$(e^{i\theta}, A) \mapsto e^{i\theta} A$$

so we have

$$U(2) = (S^1 \times SU(2))/D \qquad \text{with } D = \left\{ \pm \left(1, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \right\} \cong \mathbb{Z}/2$$

The representations of $S^1 \times SU(2)$ are parametrized by $\mathbb{Z} \times \mathbb{N}_0$. We have the standard maximal torus

$$T^2 = \left\{ \pm \left(e^{i\theta}, \begin{bmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{bmatrix} \right) \right\}$$

and we easily compute that for $(n,k) \in \mathbb{Z} \times \mathbb{N}_0$ we have

$$\chi_{n,k}(e^{i\theta}, e^{i\alpha}) = \sum_{j=0}^{k} e^{i(n\theta + (k-2j)\alpha)}$$

(see Example 21.7). Such a representation will send the generator $(e^{i\pi}, e^{i\pi})$ *of* D *to the identity if and only if* $n + (k - 2j)$ *are even for all* j*, i.e. if* n *and* k *are both even or both odd. Thus the irreducible representations of* $U(2)$ *are parametrized by*

$$\{(n, k) \in \mathbb{Z} \times \mathbb{N}_0 \colon n + k \textit{ is even.}\}$$

To summarize: the theory of compact Lie groups is closely related to the theory of complex semisimple Lie algebras. Except for central torus factors which are easily dealt with, the difference is completely explained by a finite abelian group D, the fundamental group of the compact semisimple factor of the compact Lie group. The group D can be described combinatorially in terms of the way that the dual root lattice sits inside the lattice of one parameter groups of the maximal torus. See [BtD, Section V.7].

Sketch proof of Theorem 21.4. For a fixed maximal torus $T \subset G$ consider the conjugation map

$$G/T \times T \xrightarrow{q} G$$

defined by

$$(gT, t) \mapsto gtg^{-1}$$

It is enough to show that this map is surjective. Indeed, if that is the case then an arbitrary element $h \in G$ can be written as gtg^{-1} and is therefore contained in the maximal torus gTg^{-1}. Moreover if T' is any maximal torus in T then letting $h \in T'$ denote a topological generator, we see that $h \in gTg^{-1}$ for some g, hence $T' \subset gTg^{-1}$ for some g and, T' being a maximal torus we must have $T' = gTg^{-1}$.

There are several ways of proving the map q is surjective, the nicest of which envolve some algebraic topology. One elementary way is to compute the degree of q. If the degree is nonzero then (by the standard definition of degree as the signed count of preimages of a regular value), the map must be surjective.

Let $h \in T$ be a topological generator of T, then

$$q^{-1}(h) = \{(gT, t) \colon gtg^{-1} = h\}$$

Given a point in the fiber, $t = g^{-1}hg$ must also be a topological generator of a maximal torus and, since it belongs to T, must be a topological generator of T. It follows that $gTg^{-1} = T$, i.e. gT is an element of the Weyl group of T. Conversely any element $gT \in W$ corresponds to the element $(gT, g^{-1}hg) \in q^{-1}(h)$. Thus the fiber over a topological generator of T can be identified with the Weyl group, which is a finite group.

The tangent spaces of the domain and range of q at the points in $q^{-1}(h)$ and h respectively, can naturally be identified with the vector space $\mathfrak{g}/\mathfrak{t} \oplus \mathfrak{t}$ and with regard to fixed orientations one can check that the sign of the determinant of dq is the same at all points in $q^{-1}(h)$ (one can obtain an expression for dq in terms of the adjoint action of h on the Lie algebra - see [BtD, Lemma IV.1.7] for the full details).

It follows that the degree of q is $|W|$, hence q is surjective, which concludes the proof. □

References

[Br] G. Bredon, *Introduction to compact transformation groups*, Pure and Applied Mathematics vol. 46, Academic Press, 1972.

[BtD] T. Bröcker and T. tomDieck, *Representations of Compact Lie Groups*, Springer Graduate Texts in Math. 98, 1985.

[DK] H. Duistermaat and D. Kolk, *Lie groups*, Springer Universitext, 1999.

[H] B. Hall, *Lie groups and Lie algebras. An elementary introduction*, 2nd edition, Springer Graduate Texts in Math. 222, 2015.

[HK] K. Hoffman and R. Kunze, *Linear Algebra*, 2nd edition, Prentice-Hall, 1971.

[Hu] J. Humphreys, *Introduction to Lie algebras and Representation Theory*, Springer Graduate Texts in Math. 9, 1972.

[Kn] A. Knapp, *Lie groups. Beyond an Introduction*,2nd edition, Progress in Mathematics 140, Birkhäuser, 2002.

[Ko] S. Kobayashi, *Transformation groups in Differential Geometry*, Springer Classics in Mathematics, 1995.

[Ol] P. J. Olver, *Applications of Lie Groups to Differential Equations*, 2nd edition, Springer Graduate Texts in Math. 107, 1993.

[tD] T. tom Dieck, *Transformation groups*, deGruyter Studies in Mathematics, 1987.

[Sa] H. Samelson, *Notes on Lie Algebras*, Springer Universitext, 1990.

[Wa] F. Warner, *Foundations of Differential Geometry and Lie Groups*, Springer Graduate Texts in Math. 94,1983.

[We] C. Weibel, *An introduction to homological algebra*, Cambridge Studies in Advanced Mathematics 38, Cambridge University Press, 1994.