Notes for Lie Groups & Representations
Instructor: Andrei Okounkov

Henry Liu

December 12, 2016

Abstract

These are my live-texed notes for the Fall 2016 offering of MATH GR6343 Lie Groups & Representations. There are known omissions from when I zone out in class, and additional material from when I'm trying to better understand the material. Let me know when you find errors or typos. I'm sure there are plenty.

Chapter 1

Lie Groups

1.1 Definition and Examples

Definition 1.1.1. A **Lie group** over a field k (generally $\mathbb{R}$ or $\mathbb{C}$) is a group G that is also a differentiable manifold over k such that the multiplication map $G \times G \to G$ is differentiable.

Remark. We will see later that $x \mapsto x^{-1}$ on a Lie group G is also differentiable.

Remark. There are complex Lie groups and real Lie groups. Every complex Lie group is a real Lie group, since being a complex manifold is stricter than being a real manifold.

Example 1.1.2. Some examples of Lie groups:

1. k^n as a vector space with additive group structure;
2. $\mathbb{T} \coloneqq \{z \in \mathbb{C}^* : |z| = 1\}$;
3. k^*, the multiplicative group of the field k;
4. $\mathrm{GL}(V)$, the group of matrices with non-zero determinant;
5. any finite group, or countable group with discrete topology;
6. $\mathrm{SL}_n(k)$, the group of matrices with $\det = 1$;
7. $\mathrm{GL}_n^+(k)$, the group of matrices with $\det > 0$;
8. $O_n(k)$, the group of matrices with $AA^T = A^T A$;
9. $\mathrm{SO}_n(k) \coloneqq O_n(k) \cap \mathrm{SL}_n(k)$;
10. $\mathrm{Sp}_n(k) \coloneqq \{S : S^T \Omega S = \Omega\}$ where $\Omega \coloneqq \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$;
11. $U_n(k)$, the group of matrices with $UU^* = U^* U$.

Note that $U_n(k)$ is **not** a complex Lie group, since its defining equation contains complex conjugation, which is not holomorphic.

Definition 1.1.3. A subgroup of a Lie group is a **Lie subgroup** if it is a submanifold.

Example 1.1.4. Consider the torus $\mathbb{T}^2 \coloneqq S^1 \times S^1 = \mathbb{R}^2/\mathbb{Z}^2$, and pick a line $\mathbb{R}$ in $\mathbb{R}^2$ of irrational slope. Clearly $\mathbb{R}$ is a Lie group and is a subgroup of $\mathbb{T}^2$, but it is definitely not a Lie subgroup. What went wrong: $\mathbb{R}$ needs to be a submanifold, not just a manifold in its own right.

Example 1.1.5. Examples of Lie subgroups:

1. any discrete subgroup is a Lie subgroup;

2. diagonal matrices in $\mathrm{GL}(V)$;

We have to be careful about which field Lie subgroups are taken over. For example, $\mathrm{GL}(\mathbb{C}^n)$ is both a complex and real Lie group, but $U(n) \subset \mathrm{GL}(\mathbb{C}^n)$ is only a real Lie subgroup (since it is not a complex Lie group).

Proposition 1.1.6. *Let G_1, G_2 be Lie groups over k. Then $G_1 \times G_2$ is also a Lie group over k with the standard structure of a product of groups and a product of manifolds.*

Definition 1.1.7. A group homomorphism $m\colon G_1 \to G_2$ of Lie groups is a **Lie group homomorphism** if it is differentiable.

Example 1.1.8. Some examples of Lie group homomorphisms:

1. the identity map id, or more generally embeddings of Lie subgroups;

2. any linear map;

3. the determinant map det;

4. the conjugation map $a(g)\colon x \mapsto gxg^{-1}$;

5. the exponential map $\mathbb{R} \to S^1$ given by $x \mapsto e^{ix}$.

Note that the map which is multiplication by a fixed group element g is not a Lie group homomorphism, since it is not a group homomorphism.

Definition 1.1.9. A Lie group homomorphism from $p\colon G \to \mathrm{GL}(V)$ is a **linear representation of** G.

Example 1.1.10. Some examples of linear representations:

1. $\mathbb{R} \xrightarrow{\exp} S^1 \hookrightarrow \mathrm{GL}(\mathbb{R}^2)$ given by rotations;

2. given R, S linear representations of G, we can construct $R \oplus S$, $R \otimes S$, etc.

Remark. A representation of a Lie group is its action on a vector space, but we want to talk about actions in general.

1.2 Lie group actions

Let G be a Lie group (or algebraic group) and let X be a manifold in the same category.

Definition 1.2.1. A **Lie group action** of G on X is a differentiable group action $G \times X \to X$ given by $(g, x) \mapsto g \cdot x$. Here group action means it satisfies

$$e \cdot x = x, \quad g_1 \cdot (g_2 \cdot x) = (g_1 g_2) \cdot x.$$

Remark. Note that this may not be a Lie group homomorphism, since for an arbitrary differentiable manifold X we cannot say anything about whether $\mathrm{Diff}(X)$ is a Lie group.

Example 1.2.2. A **linear representation** is an action on a vector space by linear operators, i.e. $G \to \mathrm{GL}(V)$. For any group G, we have a few canonical actions:

1. the **left** (resp. right) **regular action** where $X = G$, and $G \times G \to G$ is just the multiplication $(g_1, g_2) \mapsto g_1 g_2$ (resp. $(g_1, g_2) \mapsto g_2 g_1^{-1}$);

2. the **adjoint action** $\mathrm{Ad}\colon G \times G \to G$ given by $(g,h) \mapsto ghg^{-1}$.

A homomorphism $\varphi\colon G \to H$ **induces** an action of G on H by $(g,h) \mapsto \varphi(g)h$.

Definition 1.2.3. For $x \in X$, the set $Gx \subset X$ is the **orbit**. The set of orbits is the quotient X/G. The **stabilizer** G_x is the set of elements $g \in G$ fixing x.

Proposition 1.2.4. *Let G act on X with $x \in X$. Then:*

1. *G_x is a Lie subgroup in G;*
2. *there is some open set U containing the identity $e \in G$ such that $U \cdot x$ is a submanifold.*

In this setting, $\dim U \cdot x + \dim G_x = \dim G$.

Proof. Define $\alpha_x\colon G \to X$ by $g \mapsto g \cdot x$. It has constant rank. Hence $G_x = \alpha_x^{-1}(x)$ is a regular submanifold by the constant rank theorem, and is also clearly a subgroup.

Similarly, by the constant rank theorem, for each $g \in G$ there is some neighborhood $U \ni g$ such that its image $\alpha_x(U)$ is a submanifold in X. For $g = e$, we get that $U \cdot x$ is a submanifold.

To see that $\dim U \cdot x + \dim G_x = \dim G$, note that rank-nullity holds for the differential $d\alpha_x$ at x. □

Remark. Some general questions we can ask about actions:

1. what are the orbits of the action?
2. what does the set of orbits X/G look like?

Lemma 1.2.5. *A Lie subgroup $H \subset G$ is closed.*

Proof. Suppose $H \subset G$ is a Lie subgroup. Then its closure $\bar{H}$ is a subgroup of G. In particular, $\bar{H}$ is H-invariant. By definition, H is a submanifold of G. Hence H is open in $\bar{H}$. Right-multiplication is continuous so $Hx = r_{x^{-1}}^{-1}(H)$ is open in $\bar{H}$ too. But $\bar{H}$ is the disjoint union of cosets, i.e. $\bar{H} \setminus H = \bigsqcup_{x \neq e} Hx$ is open, i.e. H is also closed in $\bar{H}$. Since $\bar{H}$ is the closure, $H = \bar{H}$ by definition. □

Remark. Note that naturally X/G is a topological space. The natural (set-theoretic) map $X \to X/G$ induces a topology on X/G via the quotient topology; however, this topology is usually non-Hausdorff.

Example 1.2.6. Here's an example of a non-Hausdorff topology on the quotient. Let $X = \mathbb{C}^2$ and let $G = (\mathbb{R}, +)$. There are two possible actions, and the first is non-Hausdorff:

$$\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}, \quad \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

The orbits of the first action look like hyperbolas, along with the four pieces of axes and the origin. The axes are not separable from the origin.

Definition 1.2.7. A function $X/G \to \mathbb{R}$ is **regular** if its lift to $X \to \mathbb{R}$ is a morphism in the category of X.

Example 1.2.8. Let $X = \mathbb{C}$ and $G = \{\pm 1\}$ acting via multiplication. Then a function on X/G is a function f such that $f(z) = f(-z)$. In other words it is a function $g(z^2)$. Hence $z^2\colon X/G \to \mathbb{C} = X$ is an isomorphism because the sets of regular functions on X/G and X are the same.

Example 1.2.9. Let $X = \mathbb{C}^2$ and $G = \{\pm 1\}$ acting via multiplication $(x,y) \mapsto \pm(x,y)$. Regular functions here are even functions in (x,y). Any such function factors through x^2, xy, y^2, i.e. there is a map from X/G to a cone. Here the image is a cone because there is the non-trivial relation $(x^2)(y^2) = (xy)^2$. (This is actually a diffeomorphism, not just a homeomorphism.)

Remark. Really, X/G is a topological space equipped with a sheaf of functions. The question is under what conditions is it a nicely behaved space.

Example 1.2.10. Consider the map

$$\mathbb{R} \ni t \mapsto \begin{pmatrix} e^{ita} & 0 \\ 0 & e^{itb} \end{pmatrix} \in U(1)^2 \subset \mathrm{GL}(2).$$

If $a/b \in \mathbb{Q}$, then the image of this map is closed. However if $a/b \notin \mathbb{Q}$, then the image is dense.

1.3 Proper actions

Definition 1.3.1. An action is **proper** if the following map is proper (as a map of topological spaces, i.e. the preimage of compact sets is compact):

$$A \colon G \times X \to X \times X, \quad (g, x) \mapsto (x, gx)$$

Example 1.3.2. A few examples of proper actions:

1. the left regular action gives $(g_1, g_2) \mapsto (g_2, g_1 g_2)$, which is an isomorphism, so clearly it is proper;
2. if $H \subset G$ be a Lie subgroup, the restriction to H of any proper action of G is still proper;
3. any action of a compact group is proper.

The "irrational flow" of $\mathbb{R}$ on $\mathbb{T}^2$ given in 1.1.4 is **not** a proper action of $\mathbb{R}$ on $\mathbb{T}^2$.

Lemma 1.3.3. *Fix $x \in X$. The evaluation map $\alpha_x \colon G \to X$ given by $g \mapsto gx$ is proper, and therefore also closed.*

Proof. Let $K \subset X$ be a compact set. Then $A^{-1}(\{x\} \times K) = B \times \{x\}$ for some B. But $B \times \{x\}$ is compact since A is proper, so $B = \alpha_x^{-1}(K)$ is also compact. Recalling that proper maps between locally compact Hausdorff spaces (every manifold is locally $\mathbb{R}^n$, which is locally compact by Heine–Borel) are closed, α_x is also closed. □

Proposition 1.3.4. *For a proper action, the stabilizer G_x is compact for all x. Hence the adjoint action is never proper unless G is compact.*

Proof. The evaluation map $\alpha_x \colon G \to X$ is proper, so $\alpha_x^{-1}(\{x\}) = G_x$ is compact. For the adjoint action $(g, h) \mapsto (h, ghg^{-1})$, note that $G_e = G$ must therefore be compact. □

Proposition 1.3.5. *Orbits of a proper action are closed embedded submanifolds, not just immersed submanifolds.*

Remark. This prevents pathologies like the "irrational flow" of $\mathbb{R}$ on $\mathbb{T}^2$.

Proof. Fix $x \in X$. It is clear that Gx is closed since the evaluation map $\alpha_x \colon G \to X$ given by $g \mapsto gx$ is closed (by lemma 1.3.3), so $\alpha_x(G) = Gx$ is closed.

To show Gx is an embedded submanifold, it suffices to show it locally. Take a compact ball B around x. Let $A \colon G \times X \to X \times X$ denote the map $(g, x) \mapsto (x, gx)$. Since the action is proper, A is proper, i.e. $A^{-1}((x, B)) = \{g \in G : gx \in B\}$ is compact.

We use compactness to get finiteness restrictions. By the constant rank theorem applied to the constant rank map $g \mapsto gx$, for each $g \in G$ there is an open neighborhood U such that Ux is an embedded submanifold of X. By compactness, $A^{-1}((x, B))$ has a finite cover by such open sets U, i.e. $B \cap Gx$ is a finite union of embedded submanifolds. We can shrink B until $B \cap Gx$ is contained within just one embedded submanifold. Hence Gx is an embedded submanifold. □

Proposition 1.3.6. *For a proper action G on X, the quotient X/G is Hausdorff.*

Remark. Suppose $R \subset X \times X$ is an equivalence relation. The general fact is that X/R is Hausdorff if and only if R is closed.

Proof. Using the remark, for us, the equivalence relation is precisely the map $G \times X \to X \times X$ given by $(g, x) \mapsto (x, gx)$. The image of this map is closed because G acts properly on X, and so we are done. $\square$

Proposition 1.3.7. *Assume the action of G on X is proper and free, i.e. $G_x = \{1\}$ for every $x \in X$. Then X/G is a smooth manifold. (Even more strongly, it is a Hausdorff ringed space.)*

Proof. Pick a point $\bar{x} \in X/G$, which corresponds to an orbit $G \cdot x$. The orbit is a smooth manifold. Let $a\colon G \times X \to X$ be the group action, so that $da\colon \mathfrak{g} \oplus T_x X \to T_x X$ is just addition of vectors $(\xi, v) \mapsto (\xi + v)$. Pick a small transverse slice S so that we have a map $G \times S \to X$. The claim is that S can be chosen small enough such that this map is an isomorphism with a neighborhood of the orbit G_x.

1. Locally near x this map is a diffeomorphism by the inverse function theorem.

2. It is a local diffeomorphism everywhere since G moves the diffeomorphism around in the orbits..

3. Hence we must show $G \times S \to X$ is bijective with its image (because local diffeomorphisms may not be bijective, e.g. covering maps). So suppose $g_1 s_1 = g_2 s_2$, i.e. $g s_1 = s_2$. Choose a sequence $\bar{S}_1 \supset \bar{S}_2 \supset \cdots$ compact, such that $\bigcap S_i = \{x\}$. There exists a neighborhood $U \ni e \in G$ such that for any $g \in U$, if $gS \cap S \neq \emptyset$, then $g = e$ (by looking the differential of such a map would be given by addition by 0, i.e. $g = e$). Now look at $G_n := \{g \in G \setminus U : g\bar{S}_n \cap \bar{S}_n \neq \emptyset\}$. This is compact by properness and $G_1 \supset G_2 \supset \cdots$, so that $\bigcap_n G_n \neq \emptyset$, i.e. there is some element g in the intersection such that $g \cdot x = x$.

Hence for every S open in the quotient X/G, we have found a neighborhood of orbits. For every such neighborhood, we have a notion of regular functions: smooth functions which are G-invariant. This gives S a smooth structure. $\square$

Remark. In particular, G/H is a manifold for any Lie subgroup H.

Remark. What if the action is proper but not free? Then there is a point $x \in X$ with non-trivial stabilizer $G_x \neq \{1\}$. The orbit is still a smooth manifold, but now $Gx = G/G_x$. Now we can choose the slice S to be G_x-invariant: find a G_x-invariant Riemannian metric (see below) and then take S to be geodesics through $(T_x Gx)^\perp$, i.e. $S \cong (T_x Gx)^\perp$.

Proposition 1.3.8. *Every compact Lie group G has a G-invariant finite-measure regular measure dg.*

Remark. Note that the tangent bundle of any Lie group is trivial, since given a basis at $T_e G$ we can move it around via dL_g where L_g is left multiplication by g.

Proof idea. Since TG is trivial, G is orientable, and the left-invariant differential forms correspond to the tangent space $T_e G$. Hence there exists a unique left-invariant top form; explicitly, it is given by $\wedge_i (g_i^{-1} dg_i)$. (For manifolds this is a lot easier, because measures are represented by differential forms, and the Lebesgue measure is the only translation-invariant measure on $\mathbb{R}^n$.) $\square$

Remark. Left and right Haar measures both exist, and for compact Lie groups they coincide. Right translations act on the space of left-invariant Haar measures (which is $\mathbb{R}_+$), so for the left and right Haar measures to coincide, we require G has no homomorphism to the positive reals $\mathbb{R}_+$. Sufficient conditions include when G is compact, or simple, or has no 1-dimensional representations at all.

Corollary 1.3.9. *Let $\|\cdot\|_0$ be an arbitrary Riemannian metric. We can construct an invariant metric from it using*

$$\|v\|^2 := \int_{G_x} \|gv\|_0^2 \, dg.$$

Proposition 1.3.10. *Let G compact act on V an affine space, and suppose it preserves a convex set S in V. Then there exists a vector $v \in S$ fixed by G.*

Proof. Pick an arbitrary vector $v_0 \in S$, and set $v := \int_G \mu(dg)\, g \cdot v_0$. (View v as the barycenter of the orbit Gv_0.) □

Proposition 1.3.11. *Let G compact act on X a manifold. Then X has a G-invariant Riemannian metric.*

Remark. This is a generalization of the previous proposition.

Theorem 1.3.12. *Let G_x be the stabilizer of a point $x \in X$ a manifold. Let S be a G_x-invariant slice, isomorphic to $(T_x Gx)^\perp$ as a G_x-manifold. Then*

$$GS \cong G \times_{G_x} S := (G \times S)/G_x$$

as G-manifolds, i.e. manifolds with an action of G. (Here $A \times_H B := (A \times B)/H$, where $h(a,b) \mapsto (ah^{-1}, hb)$ is the standard fiber product.)

Proof. (Did we do this in class?) □

Corollary 1.3.13. *$X/G \cong S/G_x$ near Gx.*

Corollary 1.3.14. *X has a G-invariant Riemannian metric because $G \times S$ has a $G \times G_x$-invariant metric.*

Any finite-dimensional representation of a compact group is semi-simple, i.e. if we have a representation W, then $W = \bigoplus_i W_i$ where each W_i is simple. (This comes from how there is always a quadratic form that is G-invariant; given $W' \subset W$, we can always decompose $W = W' \oplus (W')^\perp$.)

Example 1.3.15. Let $\mathbb{R}$ act on $\mathbb{R}^2$ by $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$. It has two sub-representations that are trivial, but it is not the direct sum of two trivial representations.

Example 1.3.16 (Grassmannian)**.** Let $G = \mathrm{GL}(n, \mathbb{R})$ and H of upper triangular matrices with the first block being $k \times k$. Then $G/H = \mathrm{Gr}(n,k)$. Note that a matrix preserves the span of the first k basis vectors if and only if it is of the form given by H. Hence G acts on $\mathrm{Gr}(n,k)$ with H stabilizing $\mathrm{span}(e_1, \dots, e_k)$.

Alternatively, $\mathrm{Gr}(n, \mathbb{C}) = U(n)/(U(k) \times U(n-k))$, because $U(n)$ acts transitively on orthogonal bases for k-dimensional subspaces, and if an element fixes a k-dimensional subspace it also fixes the $(n-k)$-dimensional complement. This decomposition shows that $\mathrm{Gr}(n,k)$ is compact.

A chart near $L \in \mathrm{Gr}(n,k)$ is formed by linear maps $L \to V/L$; the graph of a map is a subspace. The Grassnammian $\mathrm{Gr}(n,k)$ is covered by $\binom{n}{k}$ charts of the form "$n \times k$ matrices with prescribed minor being non-zero" (there are $\binom{n}{k}$ such minors). This is a generalization of what we do for projective space, where $k = 1$ and we have just an n-tuple of numbers. Hence $\mathrm{Gr}(n,k) = M_{n,k}/\mathrm{GL}(k)$ as well, where $M_{n,k}$ is the set of all $n \times k$ matrices.

Remark. These ways of expressing $\mathrm{Gr}(n,k)$ hold over every field (except for $U(n)/(U(k) \times U(n-k))$). The question we should ask ourselves in general is if G is a linear (i.e. closed subspace of $\mathrm{GL}(n)$) algebraic group and $H \subset G$ is a subgroup, we want to make G/H an algebraic variety.

The way to do this for Grassmannians is to use the Plücker embedding: if we have $L \subset V$ where $\dim L = k$ and $\dim V = n$, then

$$\Lambda^k L \subset \Lambda^k V$$

where $\Lambda^k L$ is a line and $\Lambda^k V$ has a basis of $\binom{n}{k}$ elements. The coordinates of L we now define to be the coordinates of the line $\Lambda^k L$ inside $\Lambda^k V$, i.e. precisely the values of the minors in the $n \times k$ matrix representing L in $\mathrm{Gr}(n,k) = M_{n,k}/\mathrm{GL}(k)$. To recover the line L, let α represent $\Lambda^k L$, and take the kernel

$$V \to \Lambda^{k+1} V, \quad v \mapsto v \wedge \alpha.$$

The kernel is precisely L because $e_1 \wedge \beta = 0$ iff $\beta = e_1 \wedge \beta'$.

1.4 Some Lie group properties

	$\mathrm{GL}_n(\mathbb{R})$	$\mathrm{SL}_n(\mathbb{R})$	$O_n(\mathbb{R})$	$\mathrm{SO}_n(\mathbb{R})$	U_n	SU_n	$\mathrm{Sp}_{2n}(\mathbb{R})$
dim	n^2	n^2-1	$\frac{n(n-1)}{2}$	$\frac{n(n-1)}{2}$	n^2	n^2-1	$n(2n+1)$
π_0	$\mathbb{Z}_2$	1	$\mathbb{Z}_2$	1	1	1	1
π_1	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$	1	$\mathbb{Z}$.

We used the following facts (some of which are explained in the following subsections) in populating the table.

1. There is a surjective continuous map $\det\colon \mathrm{GL}(n,\mathbb{R}) \to \mathbb{R}^\times$, but $\mathbb{R}^\times$ is not connected. Hence $\mathrm{GL}(n,\mathbb{R})$ and even $O(n,\mathbb{R})$ is not connected. Given $M \in \mathrm{GL}^+(n,\mathbb{R})$, construct a path from M to I as follows: given a basis $v_1,\dots,v_n$, Gram–Schmidt provides an orthogonal basis
$$w_1 = v_1, \quad w_2 = v_2 - t\frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1, \quad \dots, \quad w_n = v_n - t\sum_{i<n}\frac{\langle v_n, w_i\rangle}{\langle w_i, w_i\rangle} w_n$$
where we added the parameter t to obtain a homotopy to $O(n,\mathbb{R})$; then use the homotopy $(\cos\theta)e_1 + (\sin\theta)w$ to move basis vectors to the standard basis while staying in $O(n,\mathbb{R})$. For the other groups, a similar argument works, except there is no obstruction arising from positive/negative determinant.

2. $U(n) = O(2n) \cap \mathrm{Sp}(2n,\mathbb{R})$ (complex vs real picture). This is useful because $\mathrm{Sp}(2n,\mathbb{R})$ retracts onto $U(n)$: given $A \in \mathrm{Sp}(2n,\mathbb{R})$, there is a **polar decomposition** $A = SU$ where $S := (A^TA)^{1/2}$ is symmetric and symplectic, and U is unitary, so by a preceding lemma, $A(t) = S^tU$ is the homotopy.

3. Using the long exact sequence of homotopy coming from the fibration $\mathrm{SU}(n-1) \to \mathrm{SU}(n) \to S^{2n-1}$, we get
$$\pi_1(\mathrm{SU}(n)) = \pi_1(\mathrm{SU}(n-1)) = \cdots = \pi_1(\mathrm{SU}(2)) = \pi_1(S^3) = 0.$$
Similarly, $\mathrm{SO}(n-1) \to \mathrm{SO}(n) \to S^{n-1}$ shows $\pi_i(\mathrm{SO}(n)) = \pi_1(\mathrm{SO}(3)) = \mathbb{Z}/2\mathbb{Z}$. Everything else retracts onto SO and SU.

1.5 Symplectic matrices

Definition 1.5.1. A matrix M is **symplectic** if $M^TJM = J$, where $J = \begin{pmatrix} 0 & I \\ -I & 0\end{pmatrix}$. The collection of $2n \times 2n$ symplectic matrices is denoted $\mathrm{Sp}(n,k)$ (over a field k).

Definition 1.5.2. The **Pfaffian** of a skew-symmetric matrix ω is given by taking the associated 2-form $\omega = a_{ij}e^i \wedge e^j$, then computing $1/n!\omega^n = \mathrm{Pf}(\omega)e^1 \wedge \cdots \wedge e^{2n}$.

Lemma 1.5.3. $\mathrm{Pf}^2(A) = \det(A)$ *for any skew-symmetric matrix* A.

Lemma 1.5.4. *Symplectic matrices have determinant* 1.

Proof. Use the Pfaffian argument: $\mathrm{Pf}(\Omega) = \mathrm{Pf}(M^T\Omega M) = \det(M)\,\mathrm{Pf}(\Omega)$, and since $\mathrm{Pf}(\Omega) \neq 0$, we have $\det(M) = 1$. □

Proposition 1.5.5. *Let* $S \in \mathrm{Sp}(2n,\mathbb{R})$ *be positive definite. Then it can be diagonalized using a unitary change of basis, i.e. there exists* $U \in U(2n,\mathbb{R})$ *such that* $S = U^TDU$ *where* D *is diagonal.*

Remark. Here $U(2n,\mathbb{R})$ is the image of $U(n)$ inside $M(2n,\mathbb{R})$, under the identification $A+iB \mapsto \begin{pmatrix} A & B \\ -B & A\end{pmatrix}$. In particular, if $U \in U(2n,\mathbb{R})$, we have $U^TU = I$.

Corollary 1.5.6. *If* M *is a symmetric symplectic matrix, then* $M^\alpha \in \mathrm{Sp}(2n,\mathbb{R})$ *for* $\alpha > 0$.

Proof. Diagonalize $M = U^TDU$ and note that $M^\alpha = U^TD^\alpha U$, which is still in $\mathrm{Sp}(2n,\mathbb{R})$. We require symmetric so that taking the α power makes sense (i.e. diagonalizing and taking each eigenvalue to the α power). □

1.6 Fundamental groups of Lie groups

Proposition 1.6.1. *Let $\pi\colon \tilde{G} \to G$ be the universal cover of the Lie group G. Let $\tilde{e} \in \pi^{-1}(e)$. Then there exists a unique multiplicative structure on $\tilde{G}$ (with $\tilde{e}$ the identity), that makes π a homomorphism of Lie groups.*

Proof. Consider the commutative diagram

$$\begin{array}{ccc} \tilde{G} \times \tilde{G} & \longrightarrow & \tilde{G} \\ \Big\downarrow & & \Big\downarrow p \\ G \times G & \xrightarrow{\mu_G} & G. \end{array}$$

Let $\alpha\colon \tilde{G} \times \tilde{G} \to G$ be the diagonal map. Then $\operatorname{im}(\alpha_*)$ lies in $p_*(\pi_*(\tilde{G}))$, so we have a unique lift of α to $\tilde{\mu}$. Associativity follows from uniqueness. Facts:

1. the kernel of p is discrete and normal;
2. a discrete normal subgroup of a path connected Lie group is central. □

Corollary 1.6.2. $\pi_1(G)$ *is abelian.*

Proof. (I zoned out. Help?) □

Remark. It turns out that for Lie groups, $\pi_2(G) = 0$ and $\pi_3(G)$ is torsion-free.

Chapter 2

Lie Algebras

2.1 From Lie groups to Lie algebras

Recall that we have a smooth transitive action of G on itself via $L_g(h) \coloneqq gh$.

Definition 2.1.1. A vector field X on G is **left invariant** if $(L_g)_* X = X$, i.e. $(dL_g)_h(X_h) = X_{gh}$.

For a left invariant vector field, because the action of G is transitive, the vector field is fully determined by X_e, its value at the identity.

Proposition 2.1.2. *For X and Y vector fields on a smooth manifold M, the commutator $[X, Y]f = X(Yf) - Y(Xf)$ is a vector field on M.*

Proposition 2.1.3. *If $M = G$ is a Lie group, and X, Y are left-invariant, then so is $[X, Y]$.*

Proposition 2.1.4. *If $F\colon G \to H$ and X is a left invariant vector field on G, then there is a unique left invariant vector field on H such that*

$$dF_g(X_g) = Y_{F(g)}, \quad \forall g \in G.$$

Definition 2.1.5. The **Lie algebra** $\mathfrak{g}$ of a Lie group G is the set of left-invariant vector fields with the bracket $[\cdot, \cdot]$. A **representation** of a Lie algebra $\mathfrak{g}$ is a Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{gl}(V)$ for some vector space V.

Proposition 2.1.6. *Given a Lie group representation $\rho\colon G \to \mathrm{GL}(V)$, the differential $d\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is a Lie algebra representation.*

Example 2.1.7. Let $\varphi_g(h) = ghg^{-1}$. Then $\varphi_g(e) = e$, so we can differentiate at e to get $d\varphi_g\colon \mathfrak{g} \to \mathfrak{g}$ given by $X \mapsto gXg^{-1}$ called $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$. Differentiating once more we get $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$.

Example 2.1.8. Consider $\det\colon \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^\times$. We find that $d_e(\det)(X) = \mathrm{tr}(X)$.

Example 2.1.9. The tensor product of two representations of a Lie group G is $g \cdot (v \otimes w) = (g \cdot v) \otimes (g \cdot w)$. Differentiating,

$$(d/dt)(g(t)v \otimes g(t)w)|_{t=0} = Xv \otimes w + v \otimes Xw,$$

giving the tensor product of two Lie algebra representations.

Theorem 2.1.10 (Existence)**.** *Let G, H be Lie groups with G simply connected. Then for any Lie algebra homomorphism $\varphi\colon \mathfrak{g} \to \mathfrak{h}$, there exists a map $f\colon G \to H$ such that $df = \varphi$.*

Proof sketch. Take a path $g(t)$ in G from e to g and define a path $\xi(t)$ in $T_e(G)$ by $g'(t) = dL_{g(t)}\xi(t)$. Consider a solution $h(t)$ in H of the differential equation

$$h'(t) = dL_{h(t)}\varphi(\xi(t))h(t).$$

Define $f(g) \coloneqq h(1)$. We need to check this is well-defined.

Suppose g_0, g_1 are two paths in G with $g_i(0) = e$ and $g_i(1) = g$. Since G is simply connected, these paths are homotopic; call the square given by the homotopy g. Define maps $A, B \colon [0,1] \times [0,1] \to \mathfrak{g}$ by taking $A(t, s_0)$ to be the velocity path for $g(t, s_0)$, and $B(t_0, s)$ to be the velocity path for $g(t_0, s)$, i.e.

$$\partial g(t,s)/\partial t = A(t,s)g(t,s), \quad \partial g(t,s)/\partial s = B(t,s)g(t,s).$$

Hence $(\partial B/\partial t - \partial A/\partial s)g = ABg - BAg = [A,B]g$. Define a map $h \colon [0,1] \times [0,1] \to H$ to be a solution

$$\partial h(t,s)/\partial t = \varphi(A(t,s))h(t,s).$$

If we can show that $h(1,s)$ does not depend on s, we are done. Look at the equation

$$\partial h/\partial s = \tilde{B}(t,s)h(t,s), \quad \partial \tilde{B}/\partial t = \partial(\varphi(A))/\partial s = [\varphi(A), \tilde{B}].$$

This differential equation in t is satisfied by $\varphi(B)$ and $\tilde{B}(0,s) = 0$. By uniqueness of solutions, $\tilde{B}(1,s) = \varphi(B(1,s)) = 0$, i.e. $h(1,s)$ is independent of s. □

Theorem 2.1.11 (Uniqueness)**.** *If G is a connected Lie group, then any map $f \colon G \to H$ is determined by its differential $df \colon \mathfrak{g} \to \mathfrak{h}$.*

Proof. (I zoned out. Help?) □

2.2 The Lie functor

There is a functor from the category of (real or complex) connected 1-connected Lie groups to the category of Lie algebras (over real or complex), given by

$$G \mapsto \mathfrak{g} \coloneqq T_e G, \quad G_1 \xrightarrow{f} G_2 \mapsto \mathfrak{g}_1 \xrightarrow{df} \mathfrak{g}_2.$$

For every given df, there is a unique f determined by solving the relevant differential equation. The *hard part* is, given $\mathfrak{g}$, find a Lie group G whose Lie algebra is $\mathfrak{g}$.

For any G, there is an exact sequence

$$1 \to H \to \hat{G} \xrightarrow{\gamma \mapsto \gamma(1) = g} G \to 1$$

where $\hat{G}$ is the universal cover, and H is a normal discrete subgroup (isomorphic to $\pi_1(G)$, which is abelian). Any map of Lie groups $G_1 \xrightarrow{f} G_2$ induces a map $\hat{G}_1 \xrightarrow{\hat{f}} \hat{G}_2$ which preserves the kernels of $\hat{G}_1 \to G_1$ and $\hat{G}_2 \to G_2$.

If $H = G_x$ for a G-action on X, then the Lie algebra of H is $\ker(\mathfrak{g} \to T_x X)$ where this map is the differential of $g \mapsto gx$.

Definition 2.2.1. A **Poisson algebra** is a commutative algebra and a Lie algebra, but with bracket $\{\cdot,\cdot\}$, satisfying the Leibniz rule

$$\{a, bc\} = \{a,b\}c + \{a,c\}b.$$

In other words, $a \mapsto \{a, \cdot\}$ is a map $A \to \mathrm{Der}(A, \{\cdot,\cdot\})$. (This is the Hamiltonian vector flow.) Analogously, $\mathrm{ad} \colon \xi \mapsto [\xi, \cdot]$ is also a map $\mathfrak{g} \mapsto \mathrm{Der}(\mathfrak{g}, [\cdot,\cdot])$.

Remark. If one has a family of associative products $*_\hbar$ such that

$$(a *_\hbar b)|_{\hbar=0} = ab,$$

then define

$$\{a, b\} = \lim_{\hbar \to 0} \frac{a *_\hbar b - b *_\hbar a}{\hbar}.$$

Since the numerator is the commutator, it satisfies the Jacobi identity, and therefore so does $\{a, b\}$. Hence we should view Poisson algebras as first-order approximations to non-commutative algebras, at the point where they are commutative.

Example 2.2.2. Take $\mathbb{R}^{2n}$ with coordinates $p_1, \ldots, p_n, q_1, \ldots, q_n$. We make it a Poisson algebra by declaring $\{p_i, q_j\} = \delta_{ij}$. What non-commutative algebra is this the first-order approximation of? Take $P_i = \hbar \partial_{q_i}$, which satisfies $[P_i, q_j] = \hbar \delta_{ij}$. In fact, $\mathrm{Sp}(2n)$ has a very concrete description: it consists of polynomials in p_i, q_j of degree 2, under the Poisson bracket $\{\cdot, \cdot\}$.

Recall that $\pi_1(\mathrm{SO}(n)) = \mathbb{Z}/2$ for $n \geq 3$, and $\mathbb{Z}$ for $n = 2$. Hence we can construct the universal cover of $\mathrm{SO}(n)$ as follows. Take a quadratic form Q on a vector space V, and define the Clifford algebra by $v \cdot v = Q(v)$.

Example 2.2.3. If we take $V = \mathbb{R}$ and $Q(x) = -x^2$, then the Clifford algebra is $\mathbb{C}$. If instead we take $Q(x) = x^2$, we get $\mathbb{R} \oplus \mathbb{R}$.

Example 2.2.4. Take $e_i e_j + e_j e_i = \delta_{ij}$, and note that $[e_1 e_2, e_j]$ is linear in e and preserves $e_j^2 = Q(e_j)$. Hence the dimension of the Clifford algebra Cl associated to this quadratic form Q is $2^{\dim V}$. The space of quadratic vectors in Cl is the Lie algebra of $\mathrm{SO}(n)$. The corresponding Lie group, called the **Spin group** $\mathrm{Spin}(Q)$, is the set of invertible elements $x \in \mathrm{Cl}$ that preserve V under $v \mapsto xvx^{-1}$. Clearly this map is in $\mathrm{SO}(V, Q)$ since it preserves the quadratic form Q, and is a two-fold cover with kernel ± 1.

2.3 Lie algebra to Lie group

How do we get from the Lie algebra to the Lie group? Let $\mathfrak{g}$ be a Lie algebra. Step 1 is to apply Ado's theorem.

Theorem 2.3.1 (Ado)**.** *Any finite-dimensional Lie algebra has a faithful linear representation* $\mathfrak{g} \to \mathfrak{gl}(V)$*.*

Proof sketch. One representation we have is $\mathfrak{g} \xrightarrow{\mathrm{ad}} \mathfrak{gl}(\mathfrak{g})$. The kernel is given by the center, so we must deal with it. We have a faithful representation of $\mathfrak{g}/Z(\mathfrak{g})$, so by inducting on the dimension of the center, we can move this faithful representation up to $\mathfrak{g}$. ☐

Then look for $G \subset \mathrm{GL}(V)$ (which need not be a Lie subgroup). The Lie algebra $\mathfrak{g}$ sits in the tangent space $T_e \mathrm{GL}(V)$. Using the local triviality of the tangent bundle $T \mathrm{GL}(V)$, we can make the foliation by G in $\mathrm{GL}(V)$ have tangent space $(dl_h)\mathfrak{g}$ at the point h of $\mathrm{GL}(V)$. These tangent spaces form an involutive distribution, and are therefore integrable by Frobenius.

Theorem 2.3.2 (Frobenius)**.** *A field of k-planes is integrable if and only if the subspace of vector fields tangent to any field of k-planes is a Lie algebra.*

Proof. Choose a local frame ∂_{x_i} for the distribution and check that the commutator of two basis vectors is zero. So we can change coordinates such that ∂_{e_i} is the local frame. ☐

Hence we can lift the Lie algebra $\mathfrak{g}$ to a manifold G by integrating the distribution. That G is a subgroup follows from exponentiating the addition map on tangent vectors.

Example 2.3.3. We can apply this machinery to find all connected commutative Lie groups G, i.e. the commutator is 0. Hence the Lie group G must have universal cover $\mathbb{R}^n$, with kernel a discrete subgroup $\mathbb{Z}^k$. It follows that $G = \mathbb{R}^{n-k} \times (S^1)^k$.

(We can actually use this to prove the fundamental theorem of algebra: if $[F : \mathbb{C}] > 1$, then $F^\times = \mathbb{R}^{2d} \setminus \{0\} \cong S^{2d-1} \times \mathbb{R}$, which is not commutative by the above result.)

2.4 Exponential map

There is a Lie algebra map from $\mathbb{R}$ (as a Lie algebra) to any other Lie algebra. Hence we have a Lie algebra map $\mathbb{R} \ni 1 \to \xi \in \mathfrak{g}$ that can be integrated to give a map $\exp\colon (\mathbb{R}, +) \to G$, which satisfies the differential equation $\partial_t e^{t\xi} = \xi e^{t\xi}$. In particular if $\mathfrak{g} \subset \mathfrak{gl}(V)$, then exp is exactly the matrix exponential.

Proposition 2.4.1. *$e^a e^b \neq e^{a+b}$ unless $[a, b] = 0$.*

Proof. If $[a, b] = 0$, then there is a Lie algebra homomorphism $\mathbb{R}^2 \ni 1 \mapsto (a, b) \in \mathfrak{g}$, which lifts to a Lie group homomorphism $(\mathbb{R}^2, +) \to G$. That this is a homomorphism gives $e^a e^b = e^{a+b}$. □

Proposition 2.4.2. *The exponential map $a \mapsto e^a$ is a diffeomorphism near e because $d\exp_e = \mathrm{id}$.*

Proposition 2.4.3 (Trotter product formula)**.** $e^{a+b} = \lim_{n\to\infty}(e^{a/n}e^{b/n})^n$.

Proof. Without loss of generality, we can arbitrarily scale $a + b$. So suppose a is very small, where $e^a = 1 + a + O(a^2)$. Then we are done. □

Remark. There is a formula due to Baker–Campbell–Hausdorff of $\ln(e^a e^b)$ in terms a convergent series involving only commutators. Then in a chart near the identity, multiplication is analytic in that chart. Hence a Lie group is actually a **real analytic manifold**.

What is the differential of the exponential map in general? This tells us when exp fails to be a diffeomorphism.

Theorem 2.4.4. *$d\exp(\xi)e^{-\xi} = F(\mathrm{ad}_\xi)d\xi$ where*

$$F(x) = (e^x - 1)/x = \sum_{k\geq 0} \frac{x^k}{(k+1)!}.$$

Proof. Assume $\mathfrak{gl} \subset \mathrm{GL}(n)$. Then

$$\exp(x) = 1 + x + x^2/2 + \cdots = \sum_{n\geq 0} x^n/n!$$

is the usual power series. When we differentiate, we must be careful because x is not necessarily commutative:

$$d(e^x) = \sum_{a\geq 0,\, b\geq 0} \frac{x^a\, dx\, x^b}{(a+b+1)!}.$$

Trick: write this series as a product, by noting that

$$\sum_{a\geq 0,\, b\geq 0} \frac{x^a\, dx\, x^b}{(a+b+1)!} = \int_0^1 e^{sx}\, dx\, e^{(1-s)x}\, ds \tag{2.1}$$

by observing that

$$\int_0^1 s^a(1-s)^b\, ds = \frac{a!b!}{(a+b+1)!}.$$

To extract an $\exp(x)$, we commute the ds term past the dx term (by conjugating the dx by e^{-sx}:)

$$\int_0^1 e^{sx}\,dx\,e^{(1-s)x}\,ds = \left(\int_0^1 ds\,e^{s\,\mathrm{ad}_x}(dx)\right)e^x.$$

Hence $F(x) = \int_0^1 e^{sx}\,dx$, which is indeed the expression we want. □

Remark. Equation (2.1) is a very general formula. Let X be a manifold and $v(x,t)$ be a time-dependent vector field on X. Let $G(t_0,t_1)\colon X \to X$ be the flow from time $t = t_0$ to $t = t_1$. If we vary the field, i.e. $v \mapsto v + \delta v$, what will happen to the flow? We don't know anything about G, but we can take the interval $[t_0, t_1]$ and partition it into $[t_{i/n}, t_{(i+1)/n}]$, which gives a product

$$G(t_0,t_1) = \cdots G(t_{1/n}, t_{2/n})G(t_0, t_{1/n}).$$

Taking the variation with respect to v, of course we get a sum:

$$\delta_v G(t_0,t_1) = \sum_{i=1}^n G(t_{(n-1)/n}, t_1)\cdots \delta_v G(t_{i/n}, t_{(i+1)/n})\cdots G(t_0, t_{1/n}).$$

But what is the flow $G(t, t+\epsilon)$ for a very short time? Well, it is just $G(t,t+\epsilon) = 1 + \epsilon v(x,t) + O(\epsilon^2)$. Hence if n is large,

$$dG(t_{i/n}, t_{(i+1)/n}) = dv(x,t)|t_{i/n} - t_{i+1}/n|.$$

Then for $n \to \infty$, we get a sum corresponding to the Riemann integral

$$\delta_v G(t_0,t_1) = \int_{t_0}^{t_1} G(t,t_1)\,dv\,G(t_0,t)\,dt.$$

Corollary 2.4.5. exp *is a local isomorphism if* $2\pi i k$ *for* $k \neq 0$ *is not an eigenvalue of the adjoint.*

Proof. exp is not a local isomorphism if the differential kills something, which happens if 0 is an eigenvalue of $F(\mathrm{ad}\,\xi)$, i.e. $2\pi i k$ is an eigenvalue of $\mathrm{ad}\,\xi$. □

Example 2.4.6. If $\mathrm{ad}(\xi)$ is nilpotent for every ξ, then exp is a covering. For example, take the Lie group consisting of upper triangular matrices. (Such Lie algebras are called **nilpotent**.)

Theorem 2.4.7 (Cartan)**.** *A closed subgroup $H \subset G$ of a Lie group G is a Lie subgroup, and the Lie algebra $\mathfrak{h}$ of H is*

$$\mathfrak{h} = \{\xi \in \mathfrak{g} : e^{t\xi} \in H\,\forall t\}.$$

Proof. Define $\mathfrak{h}$ this way; we will show it is the Lie algebra.

1. It is a linear subspace: $e^{a+b} = \lim_{n\to\infty}(e^{a/n}e^{b/n})^n$, and the right hand side lies in H for all n, so the limit lies in H because H is closed.

2. It is a Lie subalgebra (i.e. closed under bracket) because $\mathrm{Ad}(e^{t\xi}) = t\,\mathrm{ad}(\xi) + O(t^2)$ preserves H.

Write $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ where $\mathfrak{p}$ is the complementary linear subspace. Since exp is a local isomorphism, $G = e^{\mathfrak{h}}e^{\mathfrak{p}}$ locally (where $e^{\mathfrak{h}}$ and $e^{\mathfrak{p}}$ are submanifolds and we are taking their pointwise product).

Claim: $H = e^{\mathfrak{h}}$ locally. Suppose not. Then no matter how small we make our neighborhood, there exists $p_n \in \mathfrak{p}$ such that $p_n \to 0$ and $e^{p_n} \in H$. (If these points are not on $\mathfrak{p}$, of course we can "project" them onto $\mathfrak{p}$ by multiplying by elements of H.) But this is impossible, since then we can find a convergent subsequence among $p_n/\|p_n\|$ (where we literally take any norm), which we suppose converges to $\xi \in \mathfrak{p}$. Then

$$e^{t\xi} = \lim_{n\to\infty} e^{t(p_n/\|p_n\|)} = \lim_{n\to\infty} e^{p_n[t/\|p_n\|] + p_n\{t/\|p_n\|\}} \in H$$

since $e^{p_n[t/\|p_n\|]} \in H$ but $p_n\{t/\|p_n\|\} \to 0$. (Here $[x]$ denotes integral part and $\{x\}$ fractional part.) □

Example 2.4.8. We have the formula

$$\log \begin{pmatrix} e^a & c \\ 0 & e^b \end{pmatrix} = \begin{pmatrix} a & c\frac{a-b}{e^a-e^b} \\ 0 & b \end{pmatrix}$$

so there is a singularity when $a = b + 2\pi ik$ for $k \neq 0$. In other words, when there is a zero in exp, there is a singularity in log.

Proposition 2.4.9. *Let G be a compact Lie group, so that G has a Haar measure. Then the geodesics in this metric are $ge^{t\xi}$, i.e. $e^{t\operatorname{Ad}(g)\xi}g$. More generally, for any Lie group G,*

$$\begin{pmatrix} \text{left-invariant} \\ \text{metrics on } G \end{pmatrix} \cong \begin{pmatrix} \text{right-invariant} \\ \text{metrics on } G \end{pmatrix} \cong \begin{pmatrix} \text{metrics} \\ \text{on } \mathfrak{g} \end{pmatrix}.$$

Right translations act on left-invariant metrics via the Ad action on $\mathfrak{g}$. If G is compact, then this action preserves some metric on $\mathfrak{g}$ (because the set of metrics is convex).

2.5 Digression: classical mechanics

Example 2.5.1. Left-invariant metrics on SO(3) generalize Euler's equations for rigid bodies. The configuration space of a rigid body in $\mathbb{R}^3$ is $\mathbb{R}^3 \times \mathrm{SO}(3)$ (for center of mass and rotation). We can always work in a coordinate system where the center of mass is at rest, so only SO(3) remains. Given a rotation $g(t)$, we can view $\dot{g}$ as $\dot{g} = g\xi$ for some angular velocity vector $\xi \in \mathfrak{g}$, i.e. "in the body." Alternatively, we can find a vector ω such that $\omega g = \dot{g}$, where ω is some angular velocity in the space. Here the kinetic energy is the metric on $\mathfrak{g}$, i.e. some bilinear form on ξ, satisfying

$$\frac{1}{2}\|\dot{g}\|^2 = \frac{1}{2}\|g^{-1}\dot{g}\|^2 = \frac{1}{2}\|\xi\|^2.$$

The motion of the rigid body will be a geodesic under this metric. The Lagrangian here is $\int dt\, \|\dot{g}\|^2/2$. Note however that this is not the length of the geodesic, which is $\int dt\, \|\dot{g}\|/2$. It is better to integrate $\|\dot{g}\|^2$ even though length is reparametrization invariant.

Remark. More generally, Lagrangians are written $\int dt\, L(x(t), \dot{x}(t), t)$, and physical paths $x(t)$ are extremals of this functional. To find extremals, we vary $x \mapsto x + \delta x$, to get

$$\int dt\, (\partial_x L \delta x + \partial_{\dot{x}} L \delta \dot{x}) = \int dt\, (\partial_x L - \partial_t \partial_{\dot{x}} L)\, \delta x.$$

Since x is an extremal, this variation must vanish, i.e. $\partial_t \partial_{\dot{x}} L = \partial_x L$, the **Euler–Lagrange equation**. The description of classical mechanics in this manner allows us to easily work in moving coordinate systems.

Definition 2.5.2. We can rewrite the Lagrangian as a function $H(p, x)$ where p is now a cotangent vector by

$$H(p, x) = \max_{\dot{x}}(\langle p, \dot{x}\rangle - L(x, \dot{x}, t)).$$

The maximum is achieved when $p = \partial_{\dot{x}} L$. The equations

$$\dot{q} = \partial_p H, \quad \dot{p} = -\partial_q H$$

where $q \coloneqq x$ are **Hamilton's equations**. This says there is a Poisson algebra structure $\{p_i, q_j\} = \delta_{ij}$ on the space of functions, so that $\partial_t f(p, q) = \{H, f\}$. (Note: $\partial_t H = \{H, H\} = 0$, so energy is conserved.) Derivation of Hamilton's equations (noting that $\delta\dot{q} = 0$ because we are at an extremal for $\dot{q}$):

$$\begin{aligned} dH &= d\max_{\dot{q}}(\langle p, \dot{q}\rangle - L(q, \dot{q}, t)) \\ &= \dot{q}\delta p - \partial_q L \delta q - \partial_t L \delta t \\ &= \dot{q}\delta p - \dot{p}\delta q - \partial_t L \delta t. \end{aligned}$$

Hence we are done.

2.6 Universal enveloping algebra

Associated to a Lie algebra $\mathfrak{g}$ we will define an associative algebra $U\mathfrak{g}$ such that the category of finite-dimensional representations of $\mathfrak{g}$ is equivalent to the category of finite-dimensional representations of $U\mathfrak{g}$. Our goal is to find a basis for this algebra $U\mathfrak{g}$. First we recall some constructions in linear algebra.

Definition 2.6.1. For k any field and V a vector space over k, we can define the **tensor algebra** $T^*V \coloneqq \coprod_m T^m V$ where $T^m(V) \coloneqq V^{\otimes m}$. We can also define it using a universal property: it is the algebra with a map $V \to T^*V$ such that any other map $V \to A$ factors through T^*V.

Definition 2.6.2. From the tensor algebra, we get the **symmetric algebra** $S^*(V) = T^*(V)/I$, where I is the ideal generated by all elements of the form $x \otimes y - y \otimes x$ for any $x, y \in V$. If V has a basis $x_1, \dots, x_n$, then $S^*V \cong \mathbb{C}[x_1, \dots, x_n]$. In particular, the quotient map $\sigma\colon T^*(V) \to S^*(V)$ is injective on $T^0V = k$ and $T^1V = V$, since the generators of the ideal I are degree 2. By the universal property of the tensor algebra, $S^i(V) = \sigma(T^iV)$.

Definition 2.6.3. The **universal enveloping algebra** $U\mathfrak{g}$ of a Lie algebra $\mathfrak{g}$ is a pair $(i, U\mathfrak{g})$ where $U\mathfrak{g}$ is an associative algebra with unit, and $i\colon \mathfrak{g} \to U\mathfrak{g}$ satisfying the following universal property:

> for any associative algebra A with unit, any algebra homomorphism $\phi\colon \mathfrak{g} \to A$ with $\phi(x)\phi(y) - \phi(y)\phi(x) = \phi([x, y])$ factors through $i\colon \mathfrak{g} \to U\mathfrak{g}$.

As usual, with any definition via universal properties, $U\mathfrak{g}$ must be unique up to unique isomorphism. Its explicit construction, to show existence, is to take $U\mathfrak{g} \coloneqq T^*(\mathfrak{g})/J$ where J is the ideal generated by $x \otimes y - y \otimes x - [x, y]$ for all $x, y \in \mathfrak{g}$. Let $\pi\colon T^*(\mathfrak{g}) \to U\mathfrak{g}$ be the quotient map.

Remark. Note that elements in the ideal J are not homogeneous: $x \otimes y$ and $y \otimes x$ have degree 2, but $[x, y]$ has degree 1. So it is not obvious that $\pi|_{\mathfrak{g}}$ is injective, which was the case for the symmetric algebra. (Actually, it turns out $\pi|_{\mathfrak{g}}$ is injective, which we will prove later.) However it is clear that $\pi|_k$ is injective. In particular, at least $U\mathfrak{g}$ contains scalars and is non-empty.

Definition 2.6.4. There is a **filtration** on the tensor algebra, given by $T_m \coloneqq T^0 \oplus T^1 \oplus \cdots \oplus T^m$ (where the $T^i(V)$ are the graded components). We get an induced filtration $U_n \coloneqq \pi(T_n)$ on the universal enveloping algebra.

Definition 2.6.5. Whenever we have a filtration, we can consider the **associated graded algebra** $\mathrm{Gr} \coloneqq \mathrm{Gr}(U\mathfrak{g}) \coloneqq \bigoplus_{m \geq 0} \mathrm{Gr}^m$ where $\mathrm{Gr}^m \coloneqq U_m/U_{m-1}$. Clearly it has an algebra structure, because there is an induced multiplication

$$\mathrm{Gr}^m \times \mathrm{Gr}^n = U_m/U_{m-1} \times U_n/U_{n-1} \to U_{m+n}/U_{m+n-1} = \mathrm{Gr}^{m+n}\,.$$

So Gr is a graded associative algebra with unit 1. We have a surjective map $T^m \to U_m \to G^m = U_m/U_{m-1}$ for each graded component, so we get a surjective map $\phi\colon T^*(\mathfrak{g}) \to \mathrm{Gr}$.

Lemma 2.6.6. *ϕ is an algebra homomorphism, and $\phi(I) = 0$ where I is generated by $x \otimes y - y \otimes x$ for $x, y \in \mathfrak{g}$.*

Proof. That ϕ is an algebra homomorphism is easy, because it is induced by an algebra homomorphism. It suffices to check $\phi(I) = 0$. But $\pi(x \otimes y - y \otimes x) = \pi([x, y])$ by the construction of the universal enveloping algebra. Then because ϕ arises from $\pi\colon T^*(\mathfrak{g}) \to U(\mathfrak{g})$,

$$\phi(x \otimes y - y \otimes x) \in U_1/U_1 = 0.$$

□

Theorem 2.6.7 (Poincaré–Birkhoff–Witt (PBW))**.** *Since* $I \subset \ker(\phi\colon T^*(\mathfrak{g}) \to \mathrm{Gr})$, *we have an induced map* $T^*\mathfrak{g}/I \to \mathrm{Gr}(U\mathfrak{g})$. *This is an isomorphism of associative algebras, i.e.* $\mathrm{Gr}(U\mathfrak{g})$ *is just a polynomial algebra on the Lie algebra*

Corollary 2.6.8. *Let W be a subspace of $T^m\mathfrak{g}$, and suppose the map $T^m \to S^m\mathfrak{g}$ is an isomorphism on W. Then $\pi(W)$ is a complement to U_{m-1} in U_m.*

Proof. Consider the map from the graded piece:

$$T^m \xrightarrow{\pi} U_m \to Gr^m = U_m/U_{m-1}.$$

We have a different map $T^m \to S^m\mathfrak{g} \xrightarrow{\cong} \mathrm{Gr}^m$ (where the isomorphism is by PBW) which makes a commutative diagram. Since $W \subset T^m$ is sent isomorphically to $S^m\mathfrak{g}$, we know $W \cong \mathrm{Gr}^m = U_m/U_{m-1}$. Hence in U_m, we see Gr^m is a complement to U_{m-1}. □

Corollary 2.6.9. *The map $i\colon \mathfrak{g} \to U\mathfrak{g}$ is injective.*

Proof. This is trivial: take $S^1\mathfrak{g} = \mathfrak{g}$, and PBW says it maps isomorphically to $\mathrm{Gr}^1 = U_1/U_0$. □

Corollary 2.6.10. *Let $(x_1, x_2, \ldots)$ be a basis for the Lie algebra $\mathfrak{g}$. Then the elements*

$$x_{i(1)} \cdots x_{i(m)} \coloneqq \pi(x_{i(1)} \otimes \cdots \otimes x_{i(m)}) \quad m \in \mathbb{Z}_{\geq 0},\ i(1) \leq i(2) \leq \cdots \leq i(m)$$

form a basis for $U\mathfrak{g}$, along with 1.

Proof. Recall that $U\mathfrak{g}$ has a filtration $U_0 \subset U_1 \subset \cdots$. So if we can give a basis for every U_m/U_{m-1}, we can put them together to get a basis of the whole space $U\mathfrak{g}$. Let W be the subspace of T^m spanned by elements of the form $x_{i(1)} \otimes \cdots \otimes x_{i(m)}$. It satisfies the conditions of an earlier corollary, i.e. it is mapped isomorphically into S^m. By that corollary, the images of these elements form a basis for the complement of U_{m-1}. Putting these elements together, we get a basis for all of $U\mathfrak{g}$. □

Corollary 2.6.11. *Let $\mathfrak{h} \subset \mathfrak{g}$ be a Lie subalgebra. Extend a basis $(h_1, h_2, \ldots)$ of $\mathfrak{h}$ to an ordered basis $(h_1, h_2, \ldots, x_1, x_2, \ldots)$ of $\mathfrak{g}$. Then the map $U\mathfrak{h} \to U\mathfrak{g}$ is injective and $U\mathfrak{g}$ is a free $U\mathfrak{h}$-module with basis $\{x_{i(1)} \cdots x_{i(m)}\} \cup \{1\}$.*

Proof of PBW. We already know this map is surjective, so it suffices to prove injectivity. In other words, we must show that if $t \in T^m\mathfrak{g}$ such that $\pi(t) \in U_{m-1}$, then $t \in I$.

(Setup) Fix a basis $\{x_\lambda\}_{\lambda\in\Omega}$ of $\mathfrak{g}$. Write $S^*\mathfrak{g} = \mathbb{C}[z_\lambda]$ for $\lambda \in \Omega$. For each sequence $\Sigma = (\lambda_1, \ldots, \lambda_n)$ of indices, let

$$z_\Sigma \coloneqq z_{\lambda_1} \cdots z_{\lambda_n} \in S^m\mathfrak{g} x_\Sigma \qquad\qquad \coloneqq x_{\lambda_1} \otimes \cdots \otimes x_{\lambda_m} \in T^m\mathfrak{g}.$$

Write $\lambda \leq \Sigma$ to mean $\lambda \leq \mu$ for every $\mu \in \Sigma$.

Assume there exists a representation $\rho\colon \mathfrak{g} \to \mathrm{End}(S^*\mathfrak{g})$ satisfying:

1. $\rho(x_\lambda)z_\sigma = z_\lambda z_\sigma$ if $\lambda \leq \Sigma$;
2. $\rho(x_\lambda)z_\Sigma \equiv z_\lambda z_\Sigma \bmod S_m$ if $|\Sigma| = m$;
3. if we extend ρ to $\rho\colon T^*\mathfrak{g} \to \mathrm{End}(S^*\mathfrak{g})$, then $\ker\rho \supset J$.

We show the following result: if $t \in T_m \cap J$, written $t = t_m + t_{m-1} + \cdots$ where $t_i \in T^i\mathfrak{g}$ are the homogeneous components, then $t_m \in I$. The representation $\rho\colon \mathfrak{g} \to \mathrm{End}(S^*\mathfrak{g})$ extends to a representation $\rho\colon T^*\mathfrak{g} \to \mathrm{End}(S^*\mathfrak{g})$, so $\rho(t) = 0$ for $t \in T_m \cap J$. Then using property 2 above, the highest degree component of $\rho(t)1$ is determined by t_m, and is actually 0. Hence $t_m \in I$.

Now we proceed with the proof of PBW. Let $t \in T^m\mathfrak{g}$ and $\pi(t) \in U_{m-1}$. We want to show $t \in I$. If $\pi(t) \in U_{m-1} = \pi(T_{m-1})$, we know $\pi(t) = \pi(t')$ for $t' \in T_{m-1}$. Hence $\pi(t-t') = 0$, and we are in the situation of the preceding result: $t - t' \in T_m \cap J$, so we know the highest degree part of $t - t'$, i.e. t itself, lies in I. Hence $t \in I$.

Finally, we need to construct the representation $\rho\colon \mathfrak{g} \to \mathrm{End}(S^*\mathfrak{g})$. Equivalently, for every m, we need a map $f_m\colon \mathfrak{g} \otimes S^m \to S^*\mathfrak{g}$ satisfying the three properties we want:

1. $f_m(x_\lambda \otimes z_\Sigma) = z_\lambda z_\Sigma$ if $\lambda \leq \Sigma$ and $z_\Sigma \in S^m$;
2. $f_m(x_\lambda \otimes z_\Sigma) - z_\lambda z_\Sigma \in S^k$ for $k \leq m$ and $z_\Sigma \in S^k$;
3. $f_m(x_\lambda \otimes f_m(x_\mu \otimes z_\tau)) = f_m(x_\mu \otimes f_m(x_\lambda \otimes z_\tau)) + f_m([x_\lambda, x_\mu] \otimes z_\tau)$.

Just do it. We construct

$$f_m(x_\lambda \otimes z_{i(1)} \otimes \cdots \otimes z_{i(m)}) = z_\lambda \otimes z_{i(1)} \otimes \cdots, \quad \lambda \leq i(1).$$

If $i(1) < \lambda$, then we can swap two terms using the third property:

$$f_m(x_\lambda \otimes z_{i(1)} \otimes \cdots \otimes z_{i(m)}) = f_m(x_{i(1)} \otimes z_\lambda \otimes z_{i(1)} \otimes \cdots) + f_m([x\lambda, x_{i(1)}] \otimes z_{i(2)} \otimes \cdots)$$

which is well-defined because $[x_\lambda, x_{i(1)}]$ lies in $\mathfrak{g}$ and the remainder lies in S^{m-1}.

So we could use induction: if we defined f_{m-1}, we have defined f_m. Formally, induct on m. For $m = 0$ the construction is obvious. Now we use the commutator relation to push computations with f_m onto f_{m-1}. Explicitly we have $f_m(x_\lambda \otimes z_\Sigma) = z_\lambda z_\Sigma$ if $\lambda \leq \Sigma$. Otherwise if $\Sigma = (\mu, \tau)$ for $\mu < \lambda$, then

$$f_m(x_\lambda \otimes z_\Sigma) = f_m(x_\lambda \otimes f_{m-1}(x_\mu \otimes z_\tau))$$

Since $\mu < \lambda$, we know by the third property that this is equal to

$$f_m(x_\mu \otimes f_m(x_\lambda \otimes z_I)) + f_{m-1}([x_\lambda, x_\mu] \otimes z_\tau).$$

The hard part is to compute

$$f_m(x_\lambda \otimes z_\tau) = f_{m-1}(x_\lambda \otimes z_\tau) \equiv z_\lambda z_\tau \bmod S_{m-1}.$$

Hence now everything is well-defined, because we've pushed everything into lower degree.

Finally, the check that this construction satisfies the third property is a computation using the Jacobi identity for the bracket (which we haven't used yet). □

2.7 Poisson algebras and Poisson manifolds

A Poisson algebra A has two products: one as a commutative, associative algebra, and another as a Lie algebra. These products are compatible by the Leibniz rule

$$\{f, g_1 g_2\} = \{f, g_1\} g_2 + \{f, g_2\} g_1,$$

i.e. the bracket $\{f, -\}$ is a derivation for the commutative associative algebra. Recall that $\{-, -\}$ arises as the commutative limit of non-commutative algebras $*_\hbar$:

$$\{f, g\} = \lim_{\hbar \to 0} \frac{f *_\hbar g - g *_\hbar f}{\hbar}.$$

This limit is called the **classical limit**. The process in reverse is called **quantization** and is much more difficult.

Any commutative associative algebra can be thought of as a collection of functions on something. For example, if the ring of functions on a manifold has the structure of a Poisson algebra, we call it a **Poisson manifold**.

Example 2.7.1. Let $X = T^*M$. Then functions on X consist of pullbacks of functions on M, and also vector fields on M. We also have the algebra of differential operators of M whose lower-order bits are these two types of objects, where if coordinates on M are $(q_1, \ldots, q_n)$, then there is the commutation relation $[\partial_{q_i}, q_i] = \delta_{ij}$. If we denote $p_i \coloneqq \hbar \partial_{q_i}$ (by rescaling by $\hbar$ along fibers), then $[p_i, q_j] = \hbar \delta_{ij}$. The corresponding Poisson bracket is $\{p_i, q_j\} = \delta_{ij}$.

Remark. Consider the maximal ideal $\mathfrak{m}_x = \{f : f(x) = 0\}$ in the algebra of functions on X. Then $\{c, -\} = 0$ where c is a constant, but we also have

$$\{\mathfrak{m}_x^2, -\}|_x = 0,$$

since $\{f, -\}|_x$ is determined by the class of $f - f(x)$ in $\mathfrak{m}_x/\mathfrak{m}_x^2$, which is the cotangent space. Hence the Poisson bracket goes from differentials to functions, and therefore is a tensor.

Example 2.7.2. A Lie algebra $\mathfrak{g}$ is not a Poisson manifold, but its dual $\mathfrak{g}^*$ is. Functions on $\mathfrak{g}^*$ include constants $\Bbbk$, and linear functions $\mathfrak{g}$, and so on: $\Bbbk \oplus \mathfrak{g} \oplus S^2\mathfrak{g} \oplus \cdots$, denoted $S^\bullet \mathfrak{g}$. What is the non-commutative algebra whose limit is this? It is the universal enveloping algebra $U\mathfrak{g}_\hbar$, with a parameter $\hbar$: in the universal enveloping algebra $U\mathfrak{g}$, we had $\xi\eta - \eta\xi = [\xi, \eta]$, but for $U\mathfrak{g}_\hbar$ we define $\xi\eta - \eta\xi = \hbar[\xi, \eta]$, with $\hbar$ of degree 1.

Example 2.7.3. The intersection of the previous two examples is called the **Heisenberg Lie algebra**, where $[p_i, q_j] = e\delta_{ij}$, where e is a central element. (We can always mod out by central elements.)

Fix H a function on X, called the **Hamiltonian**. Then **Hamilton's equation** says

$$\frac{d}{dt} f = \{H, f\}.$$

As discussed, $\{H, -\}$ is a derivation of a commutative product, i.e. a vector field on X, which specifies dynamics. (Not every dynamical system is Hamiltonian though.) For example, the geodesic flow we discussed earlier on is an example of Hamiltonian dynamics, with $X = T^*M$ and $H(p, q) = (1/2)\|p\|^2$. (Of course, this corresponds to the Lagrangian formulation

$$\frac{1}{2}\int_{t_0}^{t_1} L(q, \dot{q}, t)\, dt, \quad L(q, \dot{q}, t) := \|\dot{q}(t)\|^2,$$

since $H(p, q) = \max_{\dot{q}}(\langle p, q\rangle - L(q, \dot{q}))$.) The Legendre transform is the classical limit of the Fourier transform.

Lemma 2.7.4. *The following are equivalent:*

1. $\{H, G\} = 0$ *for some function* G*;*
2. G *is preserved by the flow of* H*;*
3. H *is preserved by the flow of* G.

If $H = (1/2)\|\xi\|^2$, then we get geodesics in a left-invariant metrics. Then H is preserved by left translations by G, but there is $\dim G$ worth of flows. We call preserved quantities **integrals**, so there are $\dim G$ many integrals. For a rigid body, we write the phase space $T^*\,\mathrm{SO}(3)$ as either $\mathfrak{g} \times G$ (with coordinates (ω, g)) or $G \times \mathfrak{g}$ (with coordinates (g, ξ)), and it turns out these integrals are precisely the angular momentum ω.

So we understand ω, and we want to look at the time-evolution of ξ. By general principles,

$$\frac{d}{dt}\xi = \frac{1}{2}\{\|\xi\|^2, \xi\}.$$

We know the Poisson bracket $\{\xi_1, \xi_2\} = -[\xi_1, \xi_2]$ (the minus sign is because the ξ are left invariant). Hence we re-interpret $\{\|\xi\|^2, \xi\}$ as a bracket on T^*G as $\{\xi, \|\xi\|^2\}$ a bracket on $\mathfrak{g}^*$:

$$\frac{d}{dt}\xi = \frac{1}{2}\{\|\xi\|^2, \xi\} = \{\xi, \frac{1}{2}\|xi\|^2\}.$$

Because the metric is both left and right invariant, ξ is Ad-invariant, fixed by the action of G, i.e. $\{\eta, \|\xi\|^2\} = 0$ for every $\eta \in \mathfrak{g}$. Hence ξ is a constant.

2.8 Baker–Campbell–Hausdorff formula

In a neighborhood of the identity, $\exp\colon \mathfrak{g} \to G$ is a diffeomorphism. What does multiplication look like in this chart? In other words, what is $\log(e^X e^Y)$? We know the first-order terms are $X + Y$.

Warmup: start with a matrix Lie group, where $e^X = 1 + X + X^2/2 + \cdots$ and $\log(1+X) = X - X^2/2 + \cdots$. Then

$$\begin{aligned}\log(e^X e^Y) &= \log(1 + X + Y + X^2/2 + XY + Y^2/2 + \cdots) \\ &= X + Y + (X^2/2 + XY + Y^2/2 - (X+Y)^2/2) + \cdots = X + Y + [X,Y]/2 + \cdots.\end{aligned}$$

Let $\mathfrak{g}$ be the free Lie algebra generated by variables X and Y. Then it is graded by the number of generators: $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_1 \oplus \cdots$, where for example $\mathfrak{g}_3$ contains $[x,[x,y]]$ and $[y,[x,y]]$. What is the dimension of $\mathfrak{g}_n$? The universal enveloping algebra $U\mathfrak{g}$ is a free associative algebra and is also graded. If we take $\sum_{d \ge 0} t^d \dim(U\mathfrak{g})_d$ to be the generating function of the dimensions, it is equal to $(1-2t)^{-1}$. From this we can compute the dimensions of the grading on $\mathfrak{g}$.

Consider $\exp\colon \mathfrak{g} \to \widehat{U\mathfrak{g}}$ (completion with respect to the grading) given by $X \mapsto \sum_{n \ge 0} X^n/n!$. This is an isomorphism between series $0 + \cdots$ in $\widehat{U\mathfrak{g}}$ and series $1 + \cdots$ in $\widehat{U\mathfrak{g}}$. (Sidenote: completion means we take a series to converge if the degree of its terms goes to infinity.) Then we will show $\log(e^X e^Y)$ lies in $\widehat{\mathfrak{g}}$, i.e. that all the terms in the resulting series involve only (nested) commutators.

Suppose G is finite. Then it has a group algebra

$$\mathcal{A} \coloneqq \mathbb{C}G \cong \bigoplus_{\text{irreps } V} \operatorname{End}(V).$$

The map from G to $\mathbb{C}G$ does not remember the group, e.g. think when G is abelian. How can we reconstruct the group from the group algebra? Well, there is a (coassociative) diagonal map

$$G \xrightarrow{\Delta} G \times G, \quad g \mapsto (g,g)$$

which is a group homomorphism. By linearity, this extends to an algebra homomorphism $\mathcal{A} \xrightarrow{\Delta} \mathcal{A} \to \mathcal{A}$. This map remembers the multiplication on irreps $V_1 \oplus V_2 = \sum m_{12}^i V_i$. Hence the group is the set of solutions in $\mathcal{A}$ to $\Delta(x) = x \oplus x$, which is a non-linear equation. (Elements x satisfying this equation are called **group-like**.)

Definition 2.8.1. Such an algebra $\mathcal{A}$ with a coassociative comultiplication is called a **bialgebra**. A bialgebra is a **Hopf algebra** if in addition it has an anti-automorphism $S\colon \mathcal{A} \to \mathcal{A}$ called the **antipode**. In our case, we take $S(g) \coloneqq g^{-1}$.

Let G be a Lie group. Then take $\mathcal{A} = \mathbb{C}G$, i.e. finite linear combinations, which can be viewed as measures with finite support (where multiplication is precisely convolution). Define a map

$$\Delta\colon U\mathfrak{g} \to U\mathfrak{g} \otimes U\mathfrak{g}, \quad X \mapsto X \otimes 1 + 1 \otimes X.$$

This is the differential of $\Delta\colon G \to G \otimes G$. We can sanity-check:

$$[\Delta(X), \Delta(Y)] = [X \otimes 1 + 1 \otimes X, Y \otimes 1 + 1 \otimes Y] = [X,Y] \otimes 1 + 1 \otimes [X,Y] = \Delta([X,Y]).$$

Hence we have a Hopf algebra structure on $U\mathfrak{g}$.

Proposition 2.8.2. *If k is a field of characteristic 0, then the set of* **primitive elements**

$$\{\textit{solutions to } \Delta y = y \otimes 1 + 1 \otimes y\} \subset U\mathfrak{g}$$

is equal to $\mathfrak{g}$.

Remark. This is no longer true in characteristic p, since

$$\Delta(X^p) = \Delta(X)^p = (X \otimes 1 + 1 \otimes X)^p = X^p \otimes 1 + 1 \otimes X^p$$

shows that X^p is also primitive.

Proof. Filter $U\mathfrak{g}$ by degree (as in PBW). Denote the associated graded by $\operatorname{Gr} U\mathfrak{g}$, which is just $S\mathfrak{g}$, the symmetric algebra. View $S\mathfrak{g}$ as the polynomial algebra on $\mathfrak{g}^*$. If y is primitive, then the top degree term of y is primitive for $S\mathfrak{g}$. But comultiplication on $S\mathfrak{g}$ is just $\Delta\colon \mathbb{C}[\mathfrak{g}^*] \to \mathbb{C}[\mathfrak{g}^* \times \mathfrak{g}^*] = \mathbb{C}[\mathfrak{g}^*] \otimes \mathbb{C}[\mathfrak{g}^*]$. In other words,

$$y(\lambda + \mu) = y(\lambda) + y(\mu), \quad \lambda, \mu \in \mathfrak{g}^*.$$

Hence the top degree term of y is additive, and therefore linear. So y itself is linear, and therefore $y \in \mathfrak{g}$. (This is where we need characteristic 0: in characteristic p, it is not true that if a polynomial is additive, it is linear.) □

Lemma 2.8.3. *An element $X \in \mathfrak{g}$ is primitive if and only if $e^X := 1 + Y$ is group-like. In other words, $\Delta X = X \otimes 1 + 1 \otimes X$ if and only if $\Delta e^X = e^X \otimes e^X$.*

Proof. This is a statement about a 1-dimensional Lie algebra $\mathfrak{g}$ generated by X. Then $U\mathfrak{g}$ really just is polynomials on $\mathfrak{g}^*$, and $e^{a+b} = e^a e^b$. □

Theorem 2.8.4. $\log(e^X e^Y) \in \mathfrak{g}$.

Proof. If we have a Lie algebra $\mathfrak{g}$ freely generated by X, Y, then X and Y are primitive. By the lemma, e^X and e^Y are group-like. Then their product $e^X e^Y$ is group-like, since

$$\Delta(g_1 g_2) = \Delta(g_1)\Delta(g_2) = (g_1 \otimes g_1)(g_2 \otimes g_2) = (g_1 g_2) \otimes (g_1 g_2).$$

But then $\log(e^X e^Y)$ is primitive, by the lemma. □

So how do we actually write $\log(e^X e^Y)$ as a sum of (nested) commutators? Consider the map $\Phi\colon U\mathfrak{g} \to \hat{\mathfrak{g}}$ which takes a monomial in $U\mathfrak{g}$ and replaces the (free) multiplication with the Lie bracket, e.g.

$$xyx^3 \mapsto [[[[x, y], x], x], x].$$

Another example: $[x, y] \in \mathfrak{g}_2$ goes to $[x, y] - [y, x] = 2[x, y] \in \hat{\mathfrak{g}}$.

Lemma 2.8.5. *An element $A \in \mathfrak{g}_k \subset \mathfrak{g} \subset U\mathfrak{g}$ satisfies $\Phi(A) = kA$. In particular, A can be written in terms of (nested) commutators.*

Hence, using this lemma, we can convert the expression in $U\mathfrak{g}$ for $\log(e^X e^Y)$ into a sum of (nested) commutators, sometimes called the **Baker–Campbell–Hausdorff series** in Dynkin form. This series has a radius of convergence 1.

Corollary 2.8.6. *Lie groups are actually real analytic.*

Chapter 3

Compact Lie groups

Example 3.0.1. Some examples of compact Lie groups: $S^1 = \mathbb{R}/\mathbb{Z}$, $\mathrm{SU}(n)$, $U(n)$, $O(n, \mathbb{R})$. Some examples of non-compact Lie groups: $\mathrm{GL}(n, \mathbb{R})$, $\mathrm{SL}(n, \mathbb{R})$, $O(n, \mathbb{C})$.

If G is a compact Lie group, then it has the following nice properties.

1. G has a left and right invariant measure μ_{Haar}, which is finite. (This comes from the fact that any homomorphism $G \to (\mathbb{R}_{>0}, *)$ is trivial.)
2. (Averaging) Using this measure, we can take a vector to another vector fixed by the action of the group G:
$$v \mapsto \int_G g \cdot v \, \mu(dg);$$
3. (Complete reducibility) Any complex finite-dimensional representation V of G has a positive definite Hermitian metric, and therefore $V = \bigoplus V_i$ where the V_i are irreducible.
4. G has a left and right invariant Riemannian metric, which induces a positive-definite bilinear form $(\cdot, \cdot)$ on $\mathfrak{g}$ which is **invariant**, i.e. $(\mathrm{Ad}(g)\xi, \mathrm{Ad}(g)\eta) = (\xi, \eta)$. This can be differentiated to give $([\xi, \gamma], \eta) = (\xi, [\gamma, \eta])$. Equivalently, $\mathrm{ad}(\gamma)$ is skew-symmetric.

Proposition 3.0.2. *If $\mathfrak{g}$ has a positive-definite invariant metric, then the universal cover $\hat{G}$ of its Lie group is $\mathbb{R}^n$ times some compact Lie group.*

Proof. First, apply complete reducibility to the adjoint representation of G on $\mathfrak{g}$, to get $\mathfrak{g} = \bigoplus_i \mathfrak{g}_i$ where the $\mathfrak{g}_i$ are simple. A simple Lie algebra can either be $\mathbb{R}$ or a simple non-abelian Lie algebra. So it suffices to show that if $\mathfrak{g}$ is simple non-abelian with positive-definite invariant metric $(\cdot, \cdot)$, then $\hat{G}$ is compact.

Given $\xi \in \mathfrak{g}$, the exponential $e^{t\xi}$ is a geodesic. Claim: there is some constant c such that it fails to be a minimal geodesic for $\|t\xi\| > c$. We know $\mathrm{ad}(t\xi)$ is skew-symmetric, so its eigenvalues are purely imaginary. By rescaling ξ, which gives us the constant c, we can make sure its eigenvalues are not a subset of $(-2\pi i, 2\pi i)$. (Not all its eigenvalues can be zero, otherwise it commutes with everything.) Hence the volume of $\hat{G}$ is bounded. □

3.1 Peter–Weyl theorem

We now look at a generalization of Fourier's theorem, which says that there is an isometry
$$L^2(\mathbb{R}/\mathbb{Z}, dx) \cong \widehat{\bigoplus}_k \mathbb{C}e^{2\pi i k x}.$$
(Here $\widehat{\bigoplus}$ means to take the direct sum of the subspaces first, and then to take the completion.) From the perspective of Lie theory, the summands $\mathbb{C}e^{2\pi i k x}$ are 1×1 irreducible representations of G.

Definition 3.1.1. If V is a representation of G, then there is a function

$$\phi_{\ell,v}(g) \coloneqq \ell(g \cdot v), \quad v \in V, \, \ell \in V^*$$

called a **matrix element**. (We will prove soon that matrix elements are orthogonal.)

Theorem 3.1.2 (Peter–Weyl)**.** *If V ranges over all irreducible complex representations of G, then*

$$L^2(G, \mu_{Haar}) = \widehat{\bigoplus}_V (V^* \otimes V, (A, B) \coloneqq (\operatorname{tr} A^* B)/\dim)$$

where $V^* \otimes V$ are the matrix elements.

Remark. There is an action of $G \times G$ on $L^2(G, \mu_{\text{Haar}})$ by left and right translation:

$$(L_g f)h \coloneqq f(g^{-1}h), \quad (R_g f)h = f(hg).$$

What are the left and right actions of G on matrix elements? Well,

$$(L_g \phi_{\ell,v})h = \ell(g^{-1}hv) = \phi_{g\ell,v}, \quad (R_g \phi_{\ell,v})h = \ell(hgv) = \phi_{\ell,gv}.$$

Hence the embedding $V^* \otimes V \to \{\text{matrix elements}\}$ is $(G \times G)$-equivariant. In fact, matrix elements of V are precisely functions that transform in a representation V under R_g. The space $V^* \otimes V = \operatorname{End}(V)$ has a natural Hermitian form $(A, B) \coloneqq \operatorname{tr} A^* B$, i.e. the elementary matrices E_{ij} are orthonormal.

Theorem 3.1.3. *Matrix elements of inequivalent irreducible representations are orthogonal. Matrix elements ϕ_{ij} of a representation V are orthogonal and*

$$\|\phi_{ij}\|^2_{L^2(G)} = \frac{1}{\dim V}.$$

Hence $\|\sum \phi_{ii}\| = 1$.

Proof. Let V, W be irreducible representations of G, and let $A \colon V \to W$ be any operator. Then $\bar{A} \coloneqq \int gAg^{-1} \colon V \to W$ commutes with all $g \in G$. Schur's lemma says that:

1. if $W \neq V$, then $\bar{A} = 0$;
2. if $W = V$, then $\bar{A} = \lambda I$ where $\lambda = \operatorname{tr} A/\dim V$.

If we choose an invariant Hermitian form for V then $g^{-1} = (\bar{g})^T$ (i.e. $g \in U(V)$). Taking $A = E_{ij}$, the integral becomes

$$\left(\int g E_{ij} g^{-1} \, d\mu(g) \right)_{kl} = (\phi_{\ell j}, \phi_{ki})_{L^2}.$$

□

Hence we have shown that

$$\bigoplus_{\text{irreps } V} (V^* \otimes V, \| \cdot \|^2/\dim V) \to L^2(G, \mu)$$

is an injection, and the left hand side is $(G \times G)$-equivariant. The image consists of ***G*-finite vectors** in $L^2(G)$, i.e. vectors that transform in a finite-dimensional representation. A rephrasing Peter–Weyl is that the image of this map is dense.

Lemma 3.1.4. *Peter–Weyl is equivalent to showing that G has a faithful linear representation.*

Proof. If W is a faithful linear representation, then $G \subset \operatorname{GL}(W)$. Polynomials of $\operatorname{GL}(W)$ are just matrix elements of $W^{\otimes n}$, which decomposes as $\bigoplus V_{i,n}$ where $V_{i,n}$ are irreps. But Stone–Weierstrass says polynomials are dense in continuous functions, and continuous functions are dense in L^2. □

Hence we have proved Peter–Weyl for all the compact groups we have seen; it is an easy consequence of Stone–Weierstrass.

3.2 Compact operators

Let V be a Banach space (though we will work with Hilbert spaces only). Recall that the unit ball $\{v : \|v\| \le 1\}$ is compact if and only if $\dim V < \infty$.

Definition 3.2.1. An operator $A\colon V \to V$ is **compact** if it sends bounded sets to pre-compact sets, i.e. sets whose closures are compact.

Example 3.2.2. A map $A\colon \mathbb{C}^n \to \mathbb{C}^n$ is an $n \times n$ matrix. We have $(Av)^i = \sum_j a_{ij} v^j$, which we can write as $[Af](i) = \int a(i,j) f(j)$ with the counting measure, on basis vectors $\{1, \ldots, n\}$. But we can replace $\{1, \ldots, n\}$ with (X, ν) where ν is a measure. So we consider maps

$$K \coloneqq f(x) \mapsto \int_X K(x,y) f(y).$$

Then $K\colon L^2(X) \to C(X) \subset L^2(X)$ and takes bounded sets to pre-compact sets; we know pre-compact sets (in $C(X)$ with the sup norm) are precisely those whose functions are uniformly bounded and equi-continuous, so this is not hard to check. For example,

$$|Kf_n(x_1) - Kf_n(x_2)| \le \int |K(x_1,y) - K(x_2,y)||f_n(y)|\, dy \le C \int |f_n(y)|^2\, dy.$$

Another proof of the same fact: use that an operator A is compact if and only if it is the limit of finite rank operators in the operator norm. Such maps are called **integral operators** and are a primary example of compact operators.

Remark. Here is the more general situation. Suppose we have a functor F from topological spaces to algebras that behaves well with respect to pushforwards and pullbacks. Then $F(X \times X)$ acts on $F(X)$ via

$$Af \coloneqq (p_1)_*(A \cdot p_2^*(f)),$$

called a **Fourier–Mukai kernel**.

Theorem 3.2.3 (Spectral theorem for compact self-adjoint operators)**.** *If $K = K^*$ is compact, then $V = \widehat{\bigoplus}_i \mathbb{C} v_i$ such that $K v_i = \lambda_i v_i$, and $\lim_{i\to\infty} |\lambda_i| \to 0$. In general,*

$$K = \sum_i \lambda_i (f_i, \cdot) e_i$$

with $|\lambda_i| \to 0$, where $\|e_i\| = \|f_i\| = 1$.

Example 3.2.4. Let $X = G$, and consider the operator K which is the average of left shifts by $g \in G$:

$$K \coloneqq \int k(g) L_g\, dg, \quad (L_g f)(h) \coloneqq f(g^{-1}h).$$

Here k is some continuous function on G which we think of as a weight. Explicitly,

$$[Kf](h) = \int k(g) f(g^{-1}h)\, dg = \int k(hg^{-1}) f(g)\, dg.$$

So if we declare $K(h,g) \coloneqq k(hg^{-1})$, we have obtained an integral operator. We can make it self-adjoint by imposing $k(g^{-1}) = \overline{k(g)}$. Hence by the spectral theorem, if λ_i and v_i are the eigenvalues and eigenvectors, respectively, of K, then

$$L^2(G) = \widehat{\bigoplus}_i \mathbb{C} v_i$$

consists of summands which are clearly finite-dimensional for non-zero eigenvalues. (This comes from $\lim_{i\to\infty} |\lambda_i| = \infty$, so every non-zero eigenvalue can appear only a finite number of times.)

This is how we finish off the proof of Peter–Weyl! Note that K commutes with the right-action of G. Hence G acts on the right on $\widehat{\bigoplus}_i \mathbb{C} v_i$, and every vector corresponding to $\lambda \neq 0$ is G-finite. For $\lambda = 0$, choose a sequence k_n such that $k_n \to \delta_e$ and $k_n(g^{-1}) = \overline{k_n(g)}$. Then $\int k_n(hg^{-1}) f(g)\, dg \to f(h)$ shows that f is zero.

3.3 Complexifications

Definition 3.3.1. The **finite part** of $\widehat{\bigoplus}_V \operatorname{End}(V)$ is $\bigoplus_V \operatorname{End}(V)$. We denote it by $L^2(G)_{\text{fin}}$.

Consider $L^2(\mathrm{SU}(n))$. Its finite part $L^2(\mathrm{SU}(n))_{\text{fin}}$ is precisely $\mathbb{C}[\mathrm{SL}(n,\mathbb{C})]$, since the complexification of $\mathrm{SU}(n)$ is $\mathrm{SL}(n,\mathbb{C})$.

Definition 3.3.2. Given a 1-connected compact Lie group G with Lie algebra $\mathfrak{g}$, its **complexification** $G_{\mathbb{C}}$ is the 1-connected complex Lie group with Lie algebra $\mathfrak{g}_{\mathbb{C}} \coloneqq \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$.

Hence there is a correspondence between:

1. finite-dimensional complex representations of G;
2. finite-dimensional complex representations of $\mathfrak{g}$ (by Lie's theorem);
3. finite-dimensional complex representations of $\mathfrak{g}_{\mathbb{C}}$;
4. finite-dimensional complex representations of $G_{\mathbb{C}}$ (by Lie's theorem again).

Clearly G sits in $G_{\mathbb{C}}$ as a totally real submanifold. Matrix elements of $G_{\mathbb{C}}$ are complex analytic, and matrix elements of G are real analytic. The map between the two is by restriction and by analytic continuation.

While in general $L^2(G)$ is not an algebra (the product of two L^2 functions is not necessarily L^2 anymore), matrix elements are analytic and therefore form an algebra:

$$\operatorname{End}(V) \otimes \operatorname{End}(V') \subset \operatorname{End}(V \otimes V').$$

This algebra is finitely generated. (It also clearly has no zero divisors.) So we can make the analytic variety $G_{\mathbb{C}}$ algebraic by producing this finitely generated algebra which separates points. In other words, $G_{\mathbb{C}}$ is automatically a linear algebraic group. Also, because finite-dimensional complex representations of compact G are semisimple, the same holds for finite-dimensional complex representations of $G_{\mathbb{C}}$.

Let G be a linear algebraic group, i.e. a closed subgroup of $\mathrm{GL}(N,\Bbbk)$ for $\Bbbk$ algebraically closed. It is fairly easy to show that if G is reductive, then the category of representations of G is semisimple, and also that the analogue $\Bbbk[G] = \bigoplus_V V^* \otimes V$ of Peter–Weyl holds. Reductive Lie groups arise as complexifications of Lie groups.

3.4 Symmetric spaces

Let G be a compact Lie group, and H a Lie subgroup. We know $L^2(G/H) = \bigoplus_{\text{irreps } V} V^* \otimes V^H$. In general, we can ask: what can we say about V^H?

Definition 3.4.1. Let X be a compact (for simplicity) Riemannian manifold. We call X **symmetric** if for every point $x \in X$, there exists an isometry s_x which fixes x and acts by -1 on $T_x X$.

Remark. Since every isometry preserves geodesics, to specify an isometry it suffices to specify its action on a point and on the tangent bundle.

Example 3.4.2. The spheres S^n are clearly symmetric. We can also mod by $\{\pm 1\}$ to get $\mathbb{RP}^n$. In fact, any compact Lie group G is symmetric: the isometry around the origin is $g \mapsto g^{-1}$.

Suppose any two points on X are connected by a geodesic. Pick two points x, y and let $(x+y)/2$ denote the midpoint on the geodesic connecting them. What is $\tau_{x \to y} \coloneqq s_{(x+y)/2} s_x$? It preserves the geodesic, and on the geodesic it will be a translation by the length from x to y. It is therefore true that the group of isometries acts transitively. Hence $X = \operatorname{Isom}(X)/\operatorname{Stab}_x$.

How do we pick out the stabilizer? Note that $\operatorname{Stab}_x \subset \operatorname{Isom}(X)^{s_x}$. By the example below, we see this may not be an equality.

Example 3.4.3. Take $S^{n-1} = \mathrm{SO}(n)/\mathrm{SO}(n-1)$ with $x = e_1$. Then s_x is $\mathrm{diag}(1, -1, -1, \ldots, -1)$. But then

$$\mathrm{SO}(n)^{s_x} = \left\{ \begin{pmatrix} * & 0 & 0 & \cdots \\ 0 & & & \\ 0 & & * & \\ \vdots & & & \end{pmatrix} \right\} = O(n-1) \neq \mathrm{SO}(n-1).$$

In fact, we see that $\mathrm{Stab}_x \supset \mathrm{Isom}(X)_0^{s_x}$, the connected component of the identity. In general, the following proposition is true.

Proposition 3.4.4. $G^s \supset \mathrm{Stab}_x \supset (G^s)_0$.

Proof. Any isometry that commutes with reflection by s_x takes x to a fixed point of s_x. □

Let G be a compact Lie group with an automorphism $s\colon G \to G$ of order 2. Then G^s, the collection of fixed points of s, may not be connected, but we can choose a subgroup H such that $G^s \supset H \supset (G^s)_0$. (Keep in mind the example of the sphere, where $G^s = O(n-1)$ and $(G^s)_0 = \mathrm{SO}(n-1)$.) Then s descends to $X = G/H$, and the identity 1 is an isolated fixed point. So we have shown that symmetric spaces are precisely the quotients of compact Lie groups G by a subgroup H such that $G^s \supset H \supset (G^s)_0$ where $s^2 = 1$ is an involution.

Example 3.4.5. If $X = G$ is a compact Lie group, then at least $G \times G$ acts transitively via $(g_1, g_2) \cdot x = g_1 x g_2^{-1}$. The stabilizer Stab_1 of the identity is precisely the diagonal $\Delta(G)$. On $G \times \mathrm{G}$, there is an involution that permutes factors. It descends to $x \mapsto x^{-1}$ on X. In this case, the stabilizer Stab_1 is precisely the fixed points $(G \times G)^s$.

Example 3.4.6. The complex Grassmannian $\mathrm{Gr}(k, n, \mathbb{C})$ can be written as $U(n)/(U(k) \times U(n-k))$. Of course, $U(k) \times U(n-k)$ is the matrix commuting with $\mathrm{diag}(1, 1, \ldots, 1, -1, -1, \ldots, -1)$. It follows that the complex Grassmannian is a symmetric space. In the real case, we can write $\mathrm{Gr}(k, n, \mathbb{R})$ as $\mathrm{SO}(n)/S(O(k) \times O(n-k))$. Alternatively, we can also quotient by $\mathrm{SO}(k) \times \mathrm{SO}(n-k)$ to get the oriented Grassmannian, a double cover of $\mathrm{Gr}(k, n, \mathbb{R})$.

Example 3.4.7. Equip $\mathbb{R}^{2n}$ with a symplectic form $\omega = \sum_{i=1}^n dp_i \wedge dq_i$. A **Lagrangian subspace** is an n-dimensional subspace $L \subset \mathbb{R}^{2n}$ such that $\omega|_L = 0$. It is easy to see that n is the maximal dimension for which $\omega|_L = 0$ can happen, since ω is non-degenerate. The space of all Lagrangian subspaces is called the **Lagrangian Grassmannian** $L\,\mathrm{Gr}(2n)$.

This is a homogeneous space, but the way to see this is interesting. Think of $\mathbb{R}^{2n} \cong \mathbb{C}^n$ via $z_i := p_i + \sqrt{-1} q_i$. Then ω is proportional to the imaginary part of the Hermitian form $(z, w) := \sum_i \bar{z}_i w_i$. By definition, the unitary group $U(n)$ preserves the Hermitian form, and therefore preserves, separately, its real and imaginary parts. Hence $U(n)$ preserves ω, and is in fact transitive on $L\,\mathrm{Gr}(2n)$. The stabilizer of a point is $O(n)$, since it is precisely the stabilizer of $\mathbb{R}^n \subset \mathbb{C}^n$, i.e. where $\mathrm{im}\, z = 0$. Note that $O(n) = U(n)^s$ where s is complex conjugation $g \mapsto \bar{g}$. Alternatively, we can also take $U(n)/\mathrm{SO}(n)$ to get the double cover consisting of oriented Lagrangian subspaces.

Theorem 3.4.8 (Gelfand lemma)**.** *If $X = G/H$ is a symmetric space, then $\dim V^H \in \{0, 1\}$ for any irrep V.*

Proof. We know $L^2(H) = \bigoplus_W W \otimes W^*$ where the sum is over irreps W. Inside the sum is the trivial representation $\mathbb{C} \cdot 1$. Therefore there exists a projector $P\colon f(h) \mapsto \int_H f(h)\, dh$ where dh is the normalized Haar measure. This is analogous to the Fourier case:

$$f(t) = \sum_k \hat{f}(k) e^{2\pi i k t}, \quad \hat{f}(k) = \int_0^1 f(t) e^{-2\pi i k t}\, dt$$

extracts $\hat{f}(k)$. In our projector, we are just extracting the coefficient associated to the trivial representation.

Consider $L^2(H\backslash G/H)$, i.e. functions invariant under the H-action on both the left and the right. This is just $PL^2(G)P$ by the definition of the projector P. Similarly, the same applies for $C(H\backslash G/H)$, the space of left and right invariant continuous functions on G. Hence $L^2(H\backslash H/H) = \widehat{\bigoplus}_V (V^*)^H \otimes V^H$ since we take invariants on both sides. But each term is just $\mathrm{End}(V^H)$. The statement that $\dim V^H \in \{0,1\}$ for every V is equivalent to the statement that $\bigoplus_V \mathrm{End}(V^H)$ is commutative. But this algebra is commutative iff its completions are commutative, i.e. $C(H\backslash G/H)$ is commutative. So it suffices to prove $C(H\backslash G/H)$ is commutative.

Fact: if an algebra A has an anti-automorphism σ, i.e. a linear map such that $\sigma(ab) = \sigma(b)\sigma(a)$, such that $\sigma = 1$, then A is commutative. This is stupidly obvious but is apparently somewhat deep. Take $A = C(H\backslash G/H) = C(H\backslash X)$. We will define such an anti-automorphism σ on A by first defining it on G. Define it to be $\sigma\colon g \mapsto s(g^{-1}) = s(g)^{-1}$ (since s is a group automorphism), so that it is an anti-automorphism of G and therefore of $C(G)$ and therefore of A. Now we show it is the identity on A. Given g near the identity in X, we can write it as $g = \tau_{x\to y}h$. Then

$$\sigma(g) = \sigma(h)\sigma(\tau_{x\to y}) = \sigma(h)\tau_{x\to y}.$$

Hence $\sigma(g) \in HgH$, i.e. applying σ does not change the two-sided coset. It follows that σ is the identity on $A = C(H\backslash G/H)$. □

Remark. It was important for H to be compact because we needed to integrate over H, but not so important for G to be compact. Indeed, there are non-compact symmetric spaces like the Lobachevsky plane.

Corollary 3.4.9. $L^2(X) = \widehat{\bigoplus}_{\dim V^H = 1} V$.

Corollary 3.4.10. *G-invariant operators (of any nature) in $L^2(X)$ commute.*

Proof. Such operators commute with G and preserve the decomposition of $L^2(X)$, and therefore act by scalars in each V. So of course they commute. □

Chapter 4

Subgroups and subalgebras

4.1 Solvable and nilpotent Lie algebras

Let F be any field (of any characteristic, and not necessarily algebraically closed). Throughout, let L denote the Lie algebra, finite dimensional over the field F.

Definition 4.1.1. Define the following sequence of ideals:

$$L^{(1)} := L, \quad L^{(2)} := [L^{(1)}, L^{(1)}], \quad L^{(3)} := [L^{(2)}, L^{(2)}], \quad \cdots.$$

We say L is **solvable** if $L^{(n)} = 0$ for some n.

Example 4.1.2. A basic example is the Lie algebra L of upper triangular matrices inside $\mathfrak{gl}(n, F)$. It is easy to check that L is solvable.

Proposition 4.1.3. *1. If L is solvable, then so are all the subalgebras and homomorphic images of L.*

2. If $I \subset L$ is a solvable ideal such that L/I is solvable, then L is also solvable.

3. If $I, J \subset L$ are solvable ideals, then $I + J$ is also solvable.

Proof. (1) is obvious. (2) follows by noting that L/I is solvable implies $(L/I)^{(n)} = 0$ for some n, i.e. $L^{(n)} \subset I$ for some n. But I is solvable, so L is therefore also solvable. (3) follows from the isomorphism $(I + J)/J \to I/(I \cap J)$. Since I is solvable, $I/(I \cap J)$ is solvable by (1). But J is also solvable, so by (2), $I + J$ is also solvable. □

Definition 4.1.4. By (3) in the preceding proposition, there must exist a unique maximal solvable ideal in L, called the **radical** $\operatorname{rad} L$ of L. We say L is **semisimple** if $\operatorname{rad} L = 0$.

Remark. For any L, it follows that $L/\operatorname{rad}(L)$ is semisimple.

Definition 4.1.5. Define another sequence of ideals:

$$L^1 := L, \quad L^2 := [L^1, L^1], \quad L^3 := [L^1, L^2], \quad \cdots.$$

We say L is **nilpotent** if $L^n = 0$ for some n.

Example 4.1.6. The Lie algebra of strictly upper triangular matrices in $\mathfrak{gl}(n, F)$ is nilpotent.

Remark. It is easy to see that $L^{(i)} \subset L^i$. Hence nilpotent implies solvable. The converse is not true.

Proposition 4.1.7. *1. If L is nilpotent, then so are all the subalgebras and homomorphic images of L.*

2. *If $L/Z(L)$ is nilpotent, so is L.*

3. *If L is nilpotent and non-zero, then $Z(L) \neq 0$.*

Proof. (1) is obvious. (2) comes from $(L/Z(L))^i = 0$ implying $L^i \subset Z(L)$, so that $L^{i+1} = 0$. (3) comes from $0 = L^n = [L, L^{n-1}]$ implying $0 \neq L^{n-1} \subset Z(L)$. □

Remark. Note that L is nilpotent iff for some n, $\operatorname{ad} x_1 \operatorname{ad} x_2 \cdots \operatorname{ad} x_n(y) = 0$ for every $x_1, \ldots, x_n \in L$. In particular, $(\operatorname{ad} x)^n = 0$. So $\operatorname{ad} x \in \mathfrak{gl}(L)$ is a nilpotent matrix.

Theorem 4.1.8 (Engel)**.** *L is nilpotent if and only if all elements of L are* ad*-nilpotent, i.e.* $\operatorname{ad} x$ *is a nilpotent matrix for all $x \in L$.*

Remark. Question: given a nilpotent matrix $X \in \mathfrak{gl}(V)$, is the adjoint $\operatorname{ad} X$ also nilpotent? Yes, because $(\operatorname{ad} X)Y = XY - YX$ is nilpotent. However, the converse is not true: take $X = I$, which is not nilpotent, but $\operatorname{ad} X = 0$.

Theorem 4.1.9. *Let L be a subalgebra of $\mathfrak{gl}(V)$ (with $\dim V < \infty$). If L consists of nilpotent endomorphisms and $V \neq 0$, then there exists a non-zero vector $v \in V$ such that $Lv = 0$.*

Proof. Use induction on the dimension of L. The base cases $\dim L = 0, 1$ are obvious. So take $\dim L \geq 2$, and let $0 \neq K \subsetneq L$ be a subalgebra. By the previous remark, since every element in K is nilpotent, the adjoint action of K on L is also nilpotent. The adjoint action of K on L/K (which is well-defined because the action preserves K) is also nilpotent. Hence there is a homomorphism $K \to \mathfrak{gl}(L/K)$. By the induction hypothesis, there exists a non-zero element $x + K \in L/K$ such that $(\operatorname{ad} K)(x + K) = 0$, i.e. $[K, x] \subset K$ with $x \notin K$. Hence the normalizer $N_L(K)$ contains x, and therefore $K \subsetneq N_L(K)$. So if we take K to be a maximal proper subalgebra of L, then $N_L(K) = L$ because of the maximality of K, and $\dim L/K = 1$. Write $L = K + F \cdot z$ for some $z \in L \setminus K$. Define

$$W = \{v \in V : K \cdot v = 0\},$$

which is non-zero because x exists. It suffices now to find an element in W annihilated by z. We have

$$xzv = [x, z]v + zxv = 0 + zxv$$

since $x \in N_L(K)$. Then z commutes with the K action, and therefore we can find $v \in W$ such that $zv = 0$. □

Proof of Engel's theorem. Consider the map $L \xrightarrow{\operatorname{ad}} \mathfrak{gl}(L)$. By hypothesis, the operators $\operatorname{ad} x$ are nilpotent for every $x \in L$. Hence by the preceding theorem, there exists $v \in L$ such that $(\operatorname{ad} x)v = 0$ for all $x \in L$. Engel's theorem follows by induction on the dimension of L, using that $\dim L/Z(L) < \dim L$ and that $L/Z(L)$ nilpotent implies L nilpotent. □

Corollary 4.1.10. *Let $L \subset \mathfrak{gl}(V)$. If L consists of nilpotent endomorphisms, then there exists a flag (V_i) in V such that $X \cdot V_i \subset V_{i-1}$ for all i and all $X \in L$. In other words, there exists a basis of V such that all the matrices of L are strictly upper triangular.*

Proof. Using the theorem, find $v \in V$ such that $Lv = 0$. Take $V_1 = Fv$. Now induct to find a flag on V/V_1 which can be lifted back to V. □

From now on, assume *char* $F = 0$, and $F = \bar{F}$ is algebraically closed. We would like an analogue of Engel's theorem for solvable Lie algebras.

Theorem 4.1.11. *If $L \subset \mathfrak{gl}(V)$ is solvable (with $\dim V < \infty$), then V contains a common eigenvector for L.*

Proof. Again, induct on $\dim L$. We first find a ideal $K \subset L$ of codimension 1. Note that $[L, L] \neq L$, and is therefore a proper subalgebra. Let K be the pre-image of a codimension 1 subspace in $L/[L, L]$. Such a subspace is an ideal because $L/[L, L]$ is abelian. Hence K is a codimension 1 ideal in L. Now by the induction hypothesis, there exists an eigenvector $v \in V$ for K with associated character $\lambda\colon K \to F$ (i.e. $xv = \lambda(x)v$). Fix such a character λ, and define

$$W \coloneqq \{w \in V : xw = \lambda(x)w \;\forall x \in K\}.$$

Since $v \in W$, we know $W \neq 0$. Finally, we show L preserves W. Pick $x \in L$, $w \in W$, and $y \in K$. Then

$$yxw = [y, x]w + xyw = \lambda([y, x])w + \lambda(y)xw$$

since $[y, x] \in K$ (because K is an ideal). So if we can show $\lambda([y, x]) = 0$, then $xw \in W$. Let n be the smallest integer such that $w, xw, x^2w, \ldots, x^nw$ are linearly dependent. Define $W_i \coloneqq Fw + Fxw + \cdots + Fx^{i-1}w$ and $W_0 \coloneqq 0$, and $W_n \coloneqq W_{n+1} \coloneqq \cdots$. Check by induction (using commutators to push terms into W_i) that for all $y \in K$, we have

$$yW_i \subset W_i, \quad yx^iw \cong \lambda(y)x^iw \bmod W_i.$$

Hence $\operatorname{tr}_{W_n} y = n\lambda(y)$, because the first equation says y is an upper triangular matrix, and the second equation says the diagonal of y consists of only $\lambda(y)$. Now we have

$$n\lambda([y, x]) = \operatorname{tr}_{W_n}[y, x] = 0$$

because $\operatorname{tr}_{W_n}[y, x]$ is just the trace of two matrices. Because *char* $F = 0$, we can divide by n to get $\lambda([y, x]) = 0$. Hence write $L = K + Fz$, and find an eigenvector in W for z. Then we are done. □

Corollary 4.1.12 (Lie)**.** *If $L \subset \mathfrak{gl}(V)$ is solvable (with $\dim V < \infty$), then L stabilizes some flag (V_i) in V. In other words, the matrices of L, relative to some basis, are upper triangular.*

Proof. Obvious. □

Corollary 4.1.13. *If L is solvable, then there exists a chain of ideals of L $0 \subset L_1 \subset \cdots \subset L_n = L$ such that $\dim L_i = i$.*

Proof. Apply the preceding corollary to the adjoint representation $L \xrightarrow{\text{ad}} \mathfrak{gl}(L)$. □

Corollary 4.1.14. *If L is solvable, then $x \in [L, L]$ implies $\operatorname{ad} x$ is nilpotent. In particular, $[L, L]$ is nilpotent.*

Proof. Consider the adjoint representation $L \xrightarrow{\text{ad}} \mathfrak{gl}(L)$. Then $\operatorname{ad} L$ consists of upper triangular matrices, and $\operatorname{ad}[L, L] = [\operatorname{ad} L, \operatorname{ad} L]$ consists of strictly upper triangular matrices. By Engel's theorem, $[L, L]$ is nilpotent. □

Remark. Conversely, if $[L, L]$ is nilpotent, then L is solvable. This is because $L/[L, L]$ is commutative and therefore solvable, and $[L, L]$ is nilpotent and therefore solvable.

Theorem 4.1.15 (Cartan)**.** *Let $L \subset \mathfrak{gl}(V)$ (with $\dim V < \infty$). If $\operatorname{tr} xy = 0$ for all $x \in [L, L]$ and $y \in L$, then L is solvable.*

Lemma 4.1.16. *Let $A \subset B$ be two subspaces of $\mathfrak{gl}(V)$. Set*

$$M :- \{x \in \mathfrak{gl}(V) : [x, B] \subset A\}.$$

Suppose $x \in A$ satisfies $\operatorname{tr} xy = 0$ for all $y \in M$. Then x is nilpotent.

Proof. This is a statement from Humphrey's book. We will skip the proof. □

Proof of Cartan's theorem. We know that L is solvable iff $[L, L]$ is nilpotent. Hence it suffices to prove $[L, L]$ is nilpotent. By Engel's theorem, it suffices to show $\mathrm{ad}[L, L]$ is nilpotent. Apply the lemma: let $A = [L, L]$, and $B = L$, so that $M = \{x \in \mathfrak{gl}(V) : [x, L] \subset [L, L]\}$. In particular, $M \supset L$. For $z \in M$, we have $\mathrm{tr}([x, y]z) = \mathrm{tr}(x[y, z])$, but $[y, z] \in L$ so by hypothesis, this trace vanishes. Hence we can apply the lemma, and we are done. □

Corollary 4.1.17. *Let L be a Lie subalgebra such that $\mathrm{tr}(\mathrm{ad}\, x\, \mathrm{ad}\, y) = 0$ for all $x \in [L, L]$ and $y \in L$. Then L is solvable.*

4.2 Parabolic and Borel subgroups

Definition 4.2.1. A variety X is **complete** if for any other variety Y, the projection $X \times Y \xrightarrow{\mathrm{pr}_2} Y$ is a closed morphism.

Proposition 4.2.2. *Let X be complete. Then:*

1. *a closed subvariety of X is also complete;*
2. *if Y is complete, then so is the product $X \times Y$;*
3. *if $\phi\colon X \to Y$ is a morphism, then $\phi(X)$ is closed and complete;*
4. *if X is a subvariety of Y, then X is closed;*
5. *if X is irreducible, then $k[X] = k$;*
6. *if X is affine, then X is finite;*
7. *a projective variety is complete.*

Definition 4.2.3. G is **solvable** if there exists a series of subgroups $\{1\} = G_0 \le G_1 \le \cdots \le G_n = G$ such that G_{j-1} is normal in G_j an G_j/G_{j-1} is abelian. G is **nilpotent** if there exists n such that $(x_1, (x_2, \dots, (x_n, y)) \cdots) = e$ for all $x_1, \dots, x_n, y \in G$, where $(x, y) := xyx^{-1}y^{-1}$.

Definition 4.2.4. A closed subgroup P is **parabolic** if G/P is complete.

Example 4.2.5. Let $G = \mathrm{GL}(n, \Bbbk)$. Take P to be the block-diagonal matrices with a $k \times k$ block and a $(n-k) \times (n-k)$ block. Then P is a parabolic subgroup, since G/P is just the Grassmannian $\mathrm{Gr}(n, k)$, which is projective and therefore complete.

Lemma 4.2.6. *If P is parabolic, then G/P is projective.*

Proof. We already know G/P is quasi-projective by construction. We also know it is complete. Hence G/P is a closed subset of a projective variety, and therefore projective. □

Lemma 4.2.7. *Let $Q \subset P \subset G$ be parabolic subgroups of G. Then $Q \subset G$ is also parabolic.*

Proof. We need to show G/Q is complete, i.e. for any variety Z, the projection $G \times Z \to G/Q \times Z \to Z$ is closed. (Fact: a map $X \to Y$ between G-varieties gives an open map $X \times Z \to Y \times Z$.) Equivalently, we must show that $A \subset G \times X$ closed such that $(g, x) \in A$ implies $(gQ, x) \in A$. Consider

$$\begin{array}{ccc} P \times G \times X & \xrightarrow{\alpha} & G \times X \\ \uparrow & & \uparrow \\ \alpha^{-1}A & \longrightarrow & A. \end{array}$$

Then something happens. (?) □

Lemma 4.2.8. *If $P \subset G$ is parabolic, then any $Q \supset P$ is parabolic. Also, P is parabolic if and only if $P^0 \subset G^0$ is parabolic (connected components).*

Proof. Clearly $G/P \to G/Q$ is surjective. But G/P is complete, so the image G/Q is also complete. The second claim uses the fact that G/G^0 is finite, so $G^0 \subset G$ is automatically parabolic. This holds for any G, so in particular $P^0 \subset P$ is parabolic. If $P \subset G$ is parabolic, $P^0 \subset G$ is also parabolic. The map $G^0/P^0 \subset G/P^0$ is closed, so since closed subvarieties of complete varieties are complete, G^0/P^0 is complete, and therefore $P^0 \subset G^0$ is parabolic. Conversely, if $P^0 \subset G^0$ is parabolic, we know $G^0 \subset G$ is parabolic, so by transitivity, $P^0 \subset G$ is parabolic. But $P^0 \subset P \subset G$, so by the first part of the lemma, $P \subset G$ is also parabolic. ☐

Proposition 4.2.9. *A connected group G contains a non-trivial parabolic subgroup if and only if G is not solvable.*

Proof. Fact: if G acts on X, then there exists a closed orbit in X. (If G is a unipotent group, then every orbit is closed.) Put $G \subset \mathrm{GL}(V)$ for $\dim V$ sufficiently large. In particular, G acts on $\mathbb{P}V$. Then there exists a closed orbit O_x, which bijects with G/G_x. Since O_x is closed, it is projective and therefore complete. Then the stabilizer G_x is parabolic.

If $G_x = G$, then consider the action of G on $\mathbb{P}(V/kx)$. By the same argument, we can find another parabolic subgroup. Hence there are two cases:

1. there exists a non-trivial parabolic subgroup, i.e. at some point we stop, with $G_x \neq G$;
2. there does not exist a non-trivial parabolic subgroup, i.e. $G_x = G$ at each step, and therefore G is contained within upper triangular matrices. But upper triangular matrices are solvable, and subgroups of solvable groups are solvable, so G is solvable.

Conversely, assume G is connected and solvable, and we want to show G has no non-trivial parabolic subgroup. Assume $P \subset G$ is a maximal parabolic subgroup. Consider (G, G), which is also connected. Define $Q = P \cdot (G, G)$, which is also connected, and contains the parabolic subgroup P and is therefore parabolic.

1. If $Q = P$, then $(G, G) \subset P$ (and is a normal subgroup). Then G/P is affine, and therefore finite. But it is also connected, so $P = G$.
2. If $Q = G$, then $G(G/P) = P(G, G)/P \cong (G, G)/((G, G) \cap P)$. But $(G, G) \cap P \subset (G, G)$ is parabolic. By induction on $\dim G$, we can descend to working with (G, G), and hence $P = G$.

Hence there is no non-trivial parabolic subgroup $P \subset G$. ☐

Theorem 4.2.10 (Borel's fixed point theorem)**.** *Let G be a connected solvable linear algebraic group. Let X be a complete G-variety. Then there exists a point $x \in X$ fixed by G.*

Remark. If G acts on V, then G also acts on $\mathbb{P}V$. If there is a line $L \in \mathbb{P}V$ fixed by G, then there is an eigenvector for the group G.

Example 4.2.11. Note that in characteristic 0, a Lie group G is solvable if and only if its Lie algebra $\mathfrak{g}$ is solvable. In characteristic non-zero, the converse is false: $\mathfrak{g}$ solvable does not imply G solvable. For example, the Lie algebra $\mathfrak{sl}(2, F)$ is solvable over a field of characteristic 2, because it has the standard basis $\{e, f, h\}$ satisfying $[h, e] = 2e$, $[h, f] = 2f$, and $[e, f] = h$, which is nilpotent. They both act on $\mathbb{P}(F^2)$, but $\mathfrak{sl}(2, F)$ does not have a fixed point.

Proof of Borel's fixed point theorem. Since G acts on X, there exists a closed orbit $O_x \cong G/G_x$. We assumed G is complete, so O_x is also complete. Hence G_x is a parabolic subgroup. But G is connected and solvable, so by the proposition either $G_x = G$ or $G_x = \{e\}$. Hence either x is the desired fixed point, or we get a contradiction. ☐

Definition 4.2.12. A **Borel subgroup** of G is a closed connected solvable subgroup of G which is maximal among all subgroups with these properties.

Example 4.2.13. Take $\mathrm{GL}(n)$. Then the subgroup of all upper triangular matrices is a Borel subgroup.

Theorem 4.2.14.

1. *$P \subset G$ if parabolic if and only if P contains a Borel subgroup.*
2. *Any Borel subgroups are parabolic.*
3. *Any two Borel subgroups are conjugate.*

Proof. (1) Assume P is parabolic. Take any Borel subgroup B. Then B acts on G/P by left multiplication, so by Borel's fixed point theorem, there exists $gP \in G/P$ fixed by B. Then $g^{-1}Bg \in P$ is a Borel subgroup, by definition. Conversely, assume G is not solvable. Then there exists a parabolic subgroup $P \subset G$. Then pick a Borel set $B \subset P$ (by the forward direction). By induction on $\dim G$, we get B is parabolic in P. Since P is parabolic in G, it follows that B is parabolic in G.

(2) Easy, using the forward direction of (1).

(3) Apply Borel's fixed point theorem. □

Theorem 4.2.15 (Lie–Kolchin)**.** *Let G be a closed connected and solvable subgroup of GL_n. Then there exists some $x \in \mathrm{GL}_n$ such that xGx^{-1} is a subset of the upper triangular matrices.*

4.3 Maximal tori

Theorem 4.3.1 (Kolchin)**.** *Let V be a vector space over F, and let G be any subgroup of $\mathrm{GL}(V)$ that consists of unipotent elements (i.e. all eigenvalues are 1). Then G has a fixed point.*

Proof. We are solving the linear equation $g \cdot v = v$, so we can assume $F = \bar{F}$. We can also assume V is irreducible. Finally, we can assume the image of the group algebra $F[G]$ in $\mathrm{End}(V)$ is all of $\mathrm{End}(V)$, by Burnside. It suffices to show $g - 1 = 0$ for all $g \in G$. Compute

$$\mathrm{tr}((g-1)g') = \mathrm{tr}\, gg' - \mathrm{tr}\, g' = \dim V - \dim V = 0.$$

On the other hand, matrices of the form $(g-1)g'$ span $\mathrm{End}(V)$. Since $\mathrm{tr}(ab)$ is non-degenerate, it follows that $g - 1 = 0$ for all $g \in G$. □

An important use of fixed point theorems in Lie theory is to show conjugacy of certain kinds of subgroups.

1. If G is an arbitrary Lie group, then all maximal compact Lie subgroups are conjugate.
2. If K is a compact Lie group, then all maximal connected abelian subgroups (maximal tori) are conjugate.
3. If G is a connected linear algebraic group over $\Bbbk = \bar{\Bbbk}$, then all connected solvable groups (i.e. Borel subgroups) are conjugate.

The general argument goes as follows: if $H, H' \subset G$ are two subgroups of a certain kind, and we want to prove $gH'g^{-1} \subset H$. The subgroup H is the stabilizer of 1 in G/H. So $gH'g^{-1} \subset H$ iff H' fixes a point in G/H, namely $g^{-1}H$.

For example, to show (2), we need a torus $T' \cong (S^1)^m$ to have a fixed point on K/T. Clearly we can write $(S^1)^m$ as the closure of a single orbit, because we can pick an irrational orbit. So this is really a question about whether an operator $g \in T'$ acting on K/T has a fixed point. The Lefschetz fixed point theorem says that for $g \in \mathrm{Diff}(M)$ with M a manifold,

$$\sum_{x \in M^g} (-1)^x = \sum_{i=0}^{\dim M} (-1)^i \,\mathrm{tr}\, g|_{H^i(M,\mathbb{C})}.$$

In particular, if $g \in \mathrm{Diff}(M)_0$, then since $\operatorname{tr} g|_{H^i(M,\mathbb{C})}$ depends only on the isotopy class of g, it behaves the same as the identity, i.e.

$$\sum_{x \in M^g} (-1)^x = \sum_{i=0}^{\dim M} (-1)^i \dim H^i(M, \mathbb{C}) = \chi(M).$$

So if the Euler characteristic $\chi(M)$ is non-zero, then g must have a fixed point.

How do we prove Lefschetz's fixed point theorem? Consider the diagonal $\Delta \subset M^2$. If Γ is the graph of G, then it is $(1 \times G)\Delta$ where G acts on the second coordinate. We have $\sum_{x \in M^g} (-1)^x = \Delta \cap \Gamma$. But the Künneth formula says

$$[\Delta] = \sum_i \alpha_i \otimes \alpha^i \in H^{\text{middle}}(M^2, \mathbb{C})$$

where $\{\alpha^i\}$ and $\{\alpha_i\}$ are Poincaré duals. So the class $[\Gamma]$ of the graph is just $\sum \alpha_i \otimes g(\alpha^i)$. But now after applying the pairing, this sum is just the trace of the matrix corresponding to g.

So it suffices to show $\chi(K/T)$ is non-zero, since we know it is a compact manifold. For example, let $K = U(n)$ and T be the diagonal matrices inside. Then $M = K/T$ is the space of complete flags, since $U(n)$ acts on orthonormal frames up to rescaling. Then M^T is just the coordinate flags, which consists of S_n, the symmetric group, acting on the standard flag. Hence $|M^T| = \chi(M) = |S_n| \neq 0$. In general, let $N(T) \coloneqq \{g \in K : gTg^{-1} = T\}$ be the normalizer. Then $W = N(T)/T$ is called the **Weyl group**.

Lemma 4.3.2. *T is the connected component in $N(T)$, so W is actually a discrete group.*

Proof. There is a map $N(T) \to \mathrm{Aut}(T)$ given by $g \mapsto (t \mapsto gtg^{-1})$. But $\mathrm{Aut}(T)$ is a discrete group, since these automorphisms come from its universal cover, which is a lattice. The connected component of $N(T)$ is therefore mapped to the connected component of $\mathrm{Aut}(T)$, which is just the identity. Hence $N(T)_0 = C(T)_0$. But T is maximal connected abelian, so $C(T)_0 = T$. ☐

Theorem 4.3.3. *$\chi(K/T) = |W|$, which in particular is non-zero.*

Proof. Consider $M = K/N(T)$. Then $K/T \to M$ is a covering of degree $|W|$. Hence it suffices to prove $\chi(M) = 1$. We do this by computing the fixed points of T on M, and then applying the Lefschetz fixed point theorem. But T fixes a point iff $gTg^{-1} = T$ modulo $N(T)$, so there is only one fixed point. To get the index $(-1)^T$ of this fixed point, consider the action of T on $T_1 M = \mathrm{Lie}(K)/\mathrm{Lie}(T)$. This is just a torus acting on a vector space, so each (rotation) action is non-trivial (i.e. all weights are non-zero). Hence we have one fixed point with index 1, since the index of the origin under rotations is 1. Hence $\chi(M) = 1$. ☐

Remark. We really require characteristic 0 here; it turns out not all maximal tori are conjugate in $\mathrm{SL}(n, \mathbb{Q}_p)$ or $\mathrm{SL}(n, \mathbb{Z}_p)$.

4.4 More Borel subgroups

Let G be either a complex Lie group or an algebraic group. To use fixed point theory, we assume $k = \bar{k}$.

Theorem 4.4.1 (Borel)**.** *All Borel subgroups are conjugate.*

Example 4.4.2. Take $G = \mathrm{GL}(n)$. Then every Borel subgroup B is conjugate to the subgroup of upper diagonal matrices, by Lie's theorem. Actually, we can deduce Lie's theorem from the fixed point theorem: G/B is the space of complete flags $0 \subset F_1 \subset F_2 \subset \cdots \subset F_n = \mathbb{C}^n$. This space is projective, because it is a closed subspace of the Grassmannian. So every solvable subgroup will preserve a flag, and therefore is upper triangular in the corresponding basis.

Proof. The idea is to fix one Borel subgroup B_0 and show that G/B_0 is projective. Then any other Borel subgroup B will have a fixed point on $G/B_0 = M$, so that $gBg^{-1} \subset B_0$.

Choose a B_0 of maximal dimension, i.e. $\dim B_0 = \max_B \dim B$. Choose an embedding $G \subset \mathrm{GL}(n)$ (to be made more precise later). Consider the action of G on $\mathrm{Fl}(n)$, the space of flags. A Borel subgroup B

acting on $\mathrm{Fl}(n)$ will have some fixed point F_0, where F_0 is a flag. So consider the orbit $G \cdot F_0 \subset \mathrm{Fl}(n)$. It is closed, because it is of minimal dimension: $\dim G \cdot F_0 = \dim G - \dim \mathrm{Stab}_G F_0$, and $\mathrm{Stab}_G F_0$ is solvable, and we chose B_0 maximal. Hence $M = G \cdot F_0$ is projective, and $M^B \neq \emptyset$ for any connected solvable B. So there exists g such that $gBg^{-1} \subset (\mathrm{Stab}\, F_0)_0$. But $(\mathrm{Stab}\, F_0)_0$ is solvable and connected and contains B_0. By maximality of B_0, we have $(\mathrm{Stab}\, F_0)_0 = B_0$. We can actually make $\mathrm{Stab}\, F_0 = B_0$ by using Chevalley's theorem to find an embedding $G \subset \mathrm{GL}(n)$ and a vector e_1 such that $B_0 = \mathrm{Stab}_G(\mathbb{C}e_1)$. □

Remark. We say $P \subset G$ is **parabolic** if G/P is projective. These G/P are called **homogeneous** projective varieties. Note that G/P is projective iff P contains a Borel subgroup. It is a fact that there are only finitely many such P in G up to conjugacy.

Proposition 4.4.3. *The connected component of the normalizer $N(B) = \{g \in G : gBg^{-1} \subset B\}$ of a Borel subgroup is equal to B itself.*

Proof. We know $N(B)_0 \subset B$, because otherwise adding $g \in N(B)_0 \setminus B$ into B creates a bigger connected solvable subgroup. Now we show $N(B)/B$ is trivial. Every Borel subgroup fits into an exact sequence $1 \to U \to B \to T \to 1$ where U is unipotent and T is semisimple. (Think of T as the diagonal and U as the strictly upper triangular entries.) Consider the action of T on $G/N(B)$, which is the space of all Borel subgroups. Then $[B] \in G/N(B)$ is an isolated fixed point of T. We know $T_{[B]}G/N(B) = \mathfrak{g}/\mathfrak{b}$, where $\mathfrak{b}$ is the Lie algebra of B, and 0 is the unique fixed point. Hence the variety $G/N(B)$ is a vector space plus something (the "boundary") of codimension one. Then $\pi_1(G/N(B)) = 0$. Hence there is a fibration

$$N(B)/B \to G/B \to G/N(B)$$

which is a priori a finite cover, i.e. $N(B)/B$ is finite. But $G/N(B)$ is 1-connected, so G/B is connected, and therefore $N(B)/B$ is trivial. □

Theorem 4.4.4 (B-B decomposition)**.** *Let M be projective and smooth inside $\mathbb{P}(V)$. Let $T \subset \mathrm{GL}(V)$ be a torus acting on M. Then the fixed point locus $M^T = \bigcup_i F_i$ is also smooth, where the F_i are connected components.*

Definition 4.4.5. In the situation of the theorem, given a generic 1-parameter subgroup $\sigma\colon \mathrm{GL}(1) \to T$, define the **attracting manifold**

$$\mathrm{Attr}(F_i) \coloneqq \{m \in M : \lim_{z \to 0} \sigma(z)m \in F_i\}.$$

A map $\Bbbk^\times \to M$ extends uniquely to $\mathbb{P}^1 \to M$ since M is projective, and $\lim_{z\to 0} \sigma(z)m$ is just this additional point.

Example 4.4.6. Let $M = \mathrm{GL}(n)/B$, and T the diagonal. Then

$$M^T = \{g \in G : gTg^{-1} \subset B\}/B = \{g \in G : gTg^{-1} \subset T\}/T = N_G(T)/T = W,$$

the Weyl group, because the normalizer of T inside B is $N_B(T) = T$. Take $w = 1 \in W$. The torus T acts on the equivalence class of $\mathrm{diag}(t_1, \ldots, t_n)$ by t_i/t_j for $i > j$ on the (i,j)-th entry. If we take a 1-parameter subgroup such that $t_i/t_j \to 0$, the attracting manifold $\mathrm{Attr}(1)$ is precisely the group N_-, the lower triangular B along with 1's along the diagonal. (We know via Gaussian elimination that $\mathrm{GL}(n) = \bigsqcup_w N_- wB$.)

Theorem 4.4.7. *$M = \bigsqcup_i \mathrm{Attr}(F_i)$, and $\mathrm{Attr}(F_i) \to F_i$ is an affine linear bundle.*

Remark. This gives a decomposition of an algebraic variety into pieces, each of which is a vector bundle over a simpler algebraic variety. This equality is actually structure-preserving. For example, the Hodge structure on M is equivalent to the Hodge structures on the $\mathrm{Attr}(F_i)$, shifted appropriately.

Theorem 4.4.8 (Borel)**.** *Let G be an algebraic group over $\Bbbk = \bar{\Bbbk}$. We can ask for tori $T \cong \prod \mathrm{GL}(1, \Bbbk)$. Then all maximal tori are conjugate.*

Proof sketch. Since T is commutative, in particular solvable, and connected, there exists B such that $T \subset B$. All B are conjugate, so it is enough to show that all $T \subset B$ are B-conjugate. In fact, they are conjugate under $U \subset B$, the unipotent radical, by induction on $\dim U$. □

4.5 Levi–Malcev decomposition

Theorem 4.5.1 (Levi–Malcev)**.** *Any Lie algebra* $\mathfrak{g}$ *decomposes as a semidirect sum* $\mathfrak{g} = \mathfrak{r} \oplus \mathfrak{g}_{ss} \bigoplus_i \mathfrak{g}_i$ *where* $\mathfrak{r}$ *is solvable, called the* **radical***, and* $\mathfrak{g}_{ss} := \bigoplus_i \mathfrak{g}_i$ *is a sum of simple non-abelians. (We have* $[\mathfrak{g}_{ss}, \mathfrak{r}] \subset \mathfrak{r}$*.)*

Remark. Solvable Lie algebras have non-trivial moduli, but simple Lie algebras are **rigid**, i.e. they have no non-trivial deformations.

Remark. We will construct $\mathfrak{r}$ as the maximal solvable ideal in $\mathfrak{g}$. We must show it is uniquely determined. This is because if $\mathfrak{r}_1, \mathfrak{r}_2 \subset \mathfrak{g}$ are solvable, then $\mathfrak{r}_1 + \mathfrak{r}_2$ are also solvable.

Proof of Levi–Malcev. The radical $\mathfrak{r}$ of $\mathfrak{g}$ fits into a short exact sequence $0 \to \mathfrak{r} \to \mathfrak{g} \to \mathfrak{g}_{ss} \to 0$, where $\mathfrak{g}_{ss}$ is semisimple. It remains to show $\mathfrak{g}_{ss}$ is a sum of simples. This we do using Cartan's theorem below. □

Definition 4.5.2. A Lie algebra $\mathfrak{g}$ is **semisimple** if its radical is zero.

Definition 4.5.3. If $\mathfrak{g}$ is a Lie algebra and $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is a linear representation, define

$$(a, b)_\rho := \operatorname{tr}(\rho(a)\rho(b)).$$

This is invariant in the sense that

$$\operatorname{ad} \mathfrak{g} \subset \mathfrak{so}(\mathfrak{g}, (\cdot,\cdot)_\rho), \quad \text{i.e.}(a, [b, c])_\rho = ([a, b], c)_\rho.$$

The **Killing form** is $(\cdot,\cdot)_{\mathrm{ad}}$.

Theorem 4.5.4 (Cartan)**.** $\mathfrak{g}$ *is semisimple iff the Killing form is non-degenerate.*

Corollary 4.5.5. $\mathfrak{g}$ *is semisimple iff* $\mathfrak{g} = \bigoplus \mathfrak{g}_i$ *where* $\mathfrak{g}_i$ *are simple.*

Proof. Let $\mathfrak{g}_1 \subset \mathfrak{g}$ be a simple ideal. Then $\mathfrak{g}_1^\perp$ is also an ideal: if $\xi \in \mathfrak{g}_1^\perp$, then

$$(\mathfrak{g}_1, [b, \xi]) = ([\mathfrak{g}_1, b], \xi) = 0$$

since $[\mathfrak{g}_1, b] \subset \mathfrak{g}_1$. Since $\mathfrak{g}_1$ is simple, $\mathfrak{g}_1 \cap \mathfrak{g}_1^\perp$ is $\mathfrak{g}_1$ or 0. The former cannot happen because the Killing form is non-degenerate. □

Proof of Cartan's theorem. If the Killing form is degenerate, then $\mathfrak{g}^\perp \subset \mathfrak{g}$ is a non-zero ideal, on which Killing form is identically zero. In particular, $(a, [b, c])_{\mathrm{ad}} = 0$. Hence by the following theorem, $\mathfrak{g}^\perp$ is solvable, so $\mathfrak{g}$ is not semisimple.

Conversely, suppose the radical $\mathfrak{r}$ is non-zero. Then by taking enough commutators, we get an abelian ideal $\mathfrak{a}$. For any $y \in \mathfrak{g}$ and any $a \in \mathfrak{a}$,

$$(\operatorname{ad}(y)\operatorname{ad}(a))^2 x \subset \operatorname{ad}(y)\operatorname{ad}(a)\operatorname{ad}(y)\mathfrak{a} \subset \operatorname{ad}(y)\operatorname{ad}(a)\mathfrak{a} = 0.$$

Hence $\operatorname{tr}(\operatorname{ad}(y)\operatorname{ad}(a)) = 0$. So $\mathfrak{a} \subset \mathfrak{g}^\perp$, and the Killing form is degenerate. □

Theorem 4.5.6. *Let* $\mathfrak{g} \subset \mathfrak{gl}(V)$ *be a Lie subalgebra. Then*

$$\operatorname{tr}([a, b]c) = 0 \in (\wedge^3 \mathfrak{gl}(V)^*)^{\mathrm{GL}(V)}$$

identically iff $\mathfrak{g}$ *is solvable.*

Remark. The space $(\wedge^3 \mathfrak{gl}(V)^*)^{\mathrm{GL}(V)}$ is one-dimensional, because given a 3-form on the tangent space $\mathfrak{gl}(V)$ of $\mathrm{GL}(V)$, we can extend it to a left and right invariant element of $\Omega^3 \mathrm{GL}(V)$. In particular, it restricts to $\Omega^3 U(V)$. It is a general principle that $H^3(\mathrm{GL}(V))$ is 1-dimensional, coming from H^3 of its maximal compact $U(V)$, and is represented by an invariant form.

Remark. Given $X \in \mathfrak{gl}(V)$, take its Jordan decomposition $X = X_s + X_n$ where X_s is semisimple and X_n is nilpotent such that $[X_s, X_n] = 0$. Fact: both X_s and X_n are polynomials in X. In particular, in a linear representation, a tensor is preserved by X iff it is preserved by X_s and X_n.

Lemma 4.5.7. *If $X \in \mathfrak{g}$ where $\mathfrak{g}$ is the Lie algebra of an algebraic group, then $X_s, X_n \in \mathfrak{g}$.*

Definition 4.5.8. Let $\mathfrak{g}_{\text{alg}}$ be the intersection of all Lie algebras of algebraic groups that contain $\mathfrak{g}$. It is the Lie algebra of $\bar{G}$, the Zariski closure of G, which sits in the chain of inclusions $\mathrm{GL}(V) \supset \bar{G} \supset G$.

Proposition 4.5.9. $[\mathfrak{g}, \mathfrak{g}] = [\mathfrak{g}_{alg}, \mathfrak{g}_{alg}]$.

Proof. Consider $\{x \in \mathfrak{gl}(V) : [x, \mathfrak{g}] \subset [\mathfrak{g}, \mathfrak{g}]\}$. It is the Lie algebra of the group $\{h : h\mathfrak{g}h^{-1} \in [\mathfrak{g}, \mathfrak{g}]\}$. Hence $\mathfrak{g}_{\text{alg}}$ is contained in it, i.e. $[\mathfrak{g}, \mathfrak{g}_{\text{alg}}] \subset [\mathfrak{g}, \mathfrak{g}]$. □

Proposition 4.5.10. *Suppose $A \subset B \subset \mathrm{End}(V)$, and*

$$\mathfrak{g} = \{x : [x, B] \subset A\} = \mathrm{Lie}\{g : gBg^{-1} \equiv B \bmod A\}$$

Then for any $x \in \mathfrak{g}^\perp$, with respect to $(x, y) := \mathrm{tr}(x, y)$, we have $x_s = 0$.

Proof. Firstly, $x_s \in \mathfrak{g}$, since $\mathfrak{g}$ is algebraic. If $e_1, \ldots, e_n$ is an eigenbasis with eigenvalues λ_i, then E_{ij} are eigenvectors of $\mathrm{ad}(x_s)$ with eigenvalues $\lambda_i - \lambda_j$. If we can find a function f on the set $\{\lambda_i - \lambda_j\}$ such that $f(\lambda_i - \lambda_j) = \mu_i - \mu_j$, then the operator $\mathrm{ad}(y) = f(\mathrm{ad}(x_s))$ where $y = \mathrm{diag}(\mu_1, \ldots, \mu_n)$. But $y \in \mathfrak{g}$ and hence $\sum \mu_i \lambda_i = \mathrm{tr}\, yx = 0$ (since $x \in \mathfrak{g}^\perp$). Consider the $\mathbb{Q}$-vector space V spanned by λ_i in $\mathbb{C}$. We must show $\dim_{\mathbb{Q}} V = 0$. Suppose not. Then there exists a non-zero linear function $V \xrightarrow{\mu} \mathbb{Q}$. Now apply μ to $\sum_i \mu_i \lambda_i$, to get $\sum_i \mu_i^2$, which is 0 iff every $\mu_i = 0$. □

Proof of theorem. If $\mathfrak{g}$ is solvable, it consists of upper triangular matrices, and clearly $\mathrm{tr}([a, b]c) = 0$ when a, b, c are upper triangular. Conversely, consider the short exact sequence $0 \to Z(\mathfrak{g}) \to \mathfrak{g} \to \mathrm{ad}\,\mathfrak{g} \to 0$. Then $\mathfrak{g}$ is solvable iff $\mathrm{ad}\,\mathfrak{g}$ is solvable. Let

$$\tilde{\mathfrak{g}} := \{w : [w, \mathfrak{g}] \subset [\mathfrak{g}, \mathfrak{g}]\}.$$

If $w \in \tilde{\mathfrak{g}}$, then $\mathrm{tr}\, w[y, z] = \mathrm{tr}[w, y]z$. But $[w, y] = [x, y]$ for some $x, y \in \mathfrak{g}$, by the definition of $\tilde{\mathfrak{g}}$. Hence $\mathrm{tr}[w, y]z = \mathrm{tr}[x, y]z = 0$. So $[y, z] \in (\tilde{\mathfrak{g}})^\perp$, i.e. $[y, z]_s = 0$, and $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent. □

Chapter 5

Semisimple theory

5.1 Roots and weights

Example 5.1.1. Consider $\mathfrak{g} = \mathfrak{sl}(n)$, which has a subalgebra of diagonal matrices

$$\mathfrak{h} \coloneqq \{\operatorname{diag}(a_1, \dots, a_n) : \sum_i a_i = 0\}$$

called the **Cartan subalgebra**. We can ask how $\mathfrak{g}$ decomposes under $\operatorname{ad} \mathfrak{h}$. It will decompose as

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{i \neq j} \mathbb{C} E_{ij}$$

where E_{ij} is an eigenvalue of weight $\alpha_{ij} \coloneqq a_i - a_j \in \mathfrak{h}^*$, i.e. $[h, E_{ij}] = \alpha_{ij}(h) E_{ij}$.

Definition 5.1.2. The **roots** of $\mathfrak{g}$ are the elements $\alpha \in \mathfrak{h}^*$ which are non-zero weights of $\operatorname{ad} \mathfrak{h}$. So the above decomposition can be written as

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha} \mathfrak{g}_\alpha$$

where $\mathfrak{g}_\alpha$ is the eigenspace corresponding to α.

Proposition 5.1.3. *Let V be a representation of $\mathfrak{g}$, so that $V = \bigoplus_\alpha V_\alpha$. Then $\mathfrak{g}_\alpha V_\beta \subset V_{\alpha+\beta}$.*

Proof. Let $e \in \mathfrak{g}_\alpha$ and $v \in V_\beta$. Then compute

$$hev = [h, e]v + ehv = \alpha(h)ev + \beta(h)ev.$$ □

Corollary 5.1.4. *For every root α, there is also a root $-\alpha$.*

Proof. The proposition shows $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$. Then $\operatorname{ad}(\mathfrak{g}_\alpha) \operatorname{ad}(\mathfrak{g}_\beta) \mathfrak{g}_\gamma \subset \mathfrak{g}_{\gamma+\alpha+\beta}$. Since there are only finitely many roots, $\operatorname{ad}(\mathfrak{g}_\alpha) \operatorname{ad}(\mathfrak{g}_\beta)$ is nilpotent unless $\alpha = -\beta$. Hence the trace of this operator is 0, i.e. $\mathfrak{g}_\alpha \perp \mathfrak{g}_\beta$ with respect to the Killing form unless $\alpha = -\beta$. So there is a non-degenerate pairing between $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ given by the Killing form. □

Remark. We have a map $\mathrm{SL}_2 \to \operatorname{Ad}(G)$ given by

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \mapsto E_{ij}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \mapsto E_{ji}.$$

In SL_2, let the matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ map to $s_\alpha \in \operatorname{Ad}(G)$. This will be a permutation of the roots, but at the same time also a linear transformation $\beta \mapsto \beta - \ell_\alpha(\beta)\alpha$ where ℓ_α is some linear function.

Definition 5.1.5. A **root system** is a finite collection of non-zero vectors spanning a vector space such that for every α there exists a linear transformation of the form $\beta \mapsto \beta - \ell_\alpha(\beta)\alpha$, where $\ell_\alpha(\beta) \in \mathbb{Z}$, that preserves the root system and sends α to $-\alpha$, i.e. $\ell_\alpha(\alpha) = 2$.

Remark. These conditions are stronger than they seem. Since a root system is finite, the permutation group on the vectors in the root system is finite. In particular, the group W generated by the linear transformations s_α is finite, and therefore compact. So it preserves a positive definite inner product $(\cdot, \cdot)$. Under this inner product,

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

a reflection. Such groups generated by reflections can be classified: these are the **crystallographic groups**.

Definition 5.1.6. Let $\mathfrak{g}$ be a Lie algebra. A subalgebra $\mathfrak{h} \subset \mathfrak{g}$ is a **Cartan subalgebra** if $\mathfrak{h}$ is nilpotent and the normalizer of $\mathfrak{h}$ is $\mathfrak{h}$ itself.

Definition 5.1.7. Let V be a representation of $\mathfrak{h}$, e.g. the adjoint action on $\mathfrak{g}$. By Lie's theorem, $h \in \mathfrak{h}$ goes to a upper triangular matrix with $\alpha_i(h)$ on the diagonal. Call the $\alpha_i \in (\mathfrak{h}/[\mathfrak{h}, \mathfrak{h}])^* \subset \mathfrak{h}^*$ the **weights** of V. Write V_α for the generalized eigenspace of a weight α, i.e.

$$V_\alpha \coloneqq \{v \in V : (h - \alpha(h))^i v = 0 \text{ for some } i\}.$$

Clearly V_α is invariant under $\mathfrak{h}$.

Remark. Applying this definition to the adjoint representation, we get $\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\alpha \neq 0} \mathfrak{g}_\alpha$. We will show that $\mathfrak{g}_0 = \mathfrak{h}$. The proposition we showed earlier gives $\mathfrak{g}_\alpha V_\beta \subset V_{\alpha+\beta}$.

Definition 5.1.8. The **rank** of $\mathfrak{g}$ is the minimal number of zero eigenvalues of $\operatorname{ad} x$ for $x \in \mathfrak{g}$. Equivalently, it is the maximum size of a minor in $\operatorname{ad} x$ (over the field of rational functions in x) that is not identically zero. We say $x \in \mathfrak{g}$ is **regular** if $\operatorname{ad} x$ has this generic rank.

Remark. The set of regular elements $x \in \mathfrak{g}$ is a Zariski open set, since it is given by the condition that at least one of the minors is non-zero.

Proposition 5.1.9. *Let x be regular and consider*

$$\mathfrak{g} = \mathfrak{g}_0^x \oplus \bigoplus_{\alpha \neq 0} \mathfrak{g}_\alpha^x.$$

Then $\dim \mathfrak{g}_0^x = \operatorname{rank} \mathfrak{g}$ *and* $\mathfrak{h} \coloneqq \mathfrak{g}_0^x$ *is a Cartan subalgebra.*

Proof. We know $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$, from the result that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$. So we can restrict $\operatorname{ad} \mathfrak{h}$ to $\mathfrak{h}$. Then $\operatorname{ad}(y)|_\mathfrak{h}$ is nilpotent for every $y \in \mathfrak{h}$, because otherwise $\operatorname{ad}(y)$ will have fewer zero eigenvalues than x, since

$$\operatorname{ad}(y) = \operatorname{ad}(y)|_\mathfrak{h} \oplus \operatorname{ad}(y)|_{\mathfrak{g}/\mathfrak{h}}.$$

Hence $\mathfrak{h}$ is nilpotent, by definition. Now suppose some element z is in the normalizer of $\mathfrak{h}$, i.e. $[x, z] \in \mathfrak{h}$. By the nilpotence of $\mathfrak{h}$, we know $\operatorname{ad}(x)^N z = 0$ for $N \gg 0$. Hence $z \in \mathfrak{h}$, by the definition of $\mathfrak{h}$. □

Remark. Let $\mathfrak{h}$ be an arbitrary Cartan subalgebra. Then $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \neq 0} \mathfrak{g}_\alpha$ and define

$$\mathfrak{h}_{\text{reg}} \coloneqq \{h : \alpha(h) \neq 0 \ \forall \alpha\}$$

so that for all $x \in \mathfrak{h}_{\text{reg}}$, $\mathfrak{g}_0^x = \mathfrak{h}$.

Proposition 5.1.10. *Let $\mathfrak{g}$ be a simple Lie algebra.*

1. *The Cartan subalgebra $\mathfrak{h}$ is commutative and consists of* ad*-semisimple elements.*

2. *The Killing form restricted to $\mathfrak{h}$ is non-degenerate.*

Proof. We know $\mathfrak{h}$ nilpotent implies $([x,y],z)=0$ for any $x,y,z\in\mathfrak{h}$. Since z is arbitrary, and the Killing form is non-degenerate, $[x,y]=0$ for all $x,y\in\mathfrak{h}$.

In the decomposition $\mathfrak{g}=\mathfrak{h}\oplus\bigoplus_{\alpha\in\mathfrak{h}^*}\mathfrak{g}_\alpha$, we know $\mathfrak{g}_\alpha\perp\mathfrak{g}_\beta$ (with respect to the Killing form) unless $\alpha+\beta=0$. So $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ are dual, leaving $\mathfrak{h}$ in the direct sum. Hence the Killing form is also non-degenerate on $\mathfrak{h}$. □

5.2 Root systems

Definition 5.2.1. A **root system** $\Delta\subset\mathbb{R}^n\setminus\{0\}$ is a finite subset of non-zero vectors such that for any $\alpha\in\Delta$, the reflection

$$r_\alpha(\beta)\coloneqq\beta-2\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha$$

preserves Δ and $\langle\alpha,\beta\rangle\coloneqq 2(\alpha,\beta)/(\alpha,\alpha)$ is an integer. We say Δ is

1. **reducible** if $\Delta=\Delta_1\oplus\Delta_2$, and
2. **reduced** if $2\alpha\notin\Delta$ for any $\alpha\in\Delta$.

Example 5.2.2 (Root systems for $n=1$)**.** Suppose $\alpha\in\Delta$. Then $-\alpha\in\Delta$ as well. Take another vector $\beta\in\Delta$. Then $2(\alpha,\beta)/(\alpha,\alpha)$ must be an integer, i.e. $2\beta/\alpha\in\mathbb{Z}$. So there is only one reduced root system, called A_1, given by $\{\pm\alpha\}$, and one non-reduced root system $\{\pm\alpha,\pm2\alpha\}$. It turns out Lie algebras always have reduced root systems, so $\{\pm\alpha\}$ corresponds to $\mathfrak{sl}(2)$.

Example 5.2.3 (Root systems for $n=2$)**.** Suppose there is a vector β forming an angle θ with α, and this is the smallest θ formed by any vector with α. Then

$$\langle\alpha,\beta\rangle\langle\beta,\alpha\rangle=4\frac{(\alpha,\beta)^2}{(\alpha,\alpha)(\beta,\beta)}=4\cos^2\theta$$

must be an integer. So there are five possibilities.

1. ($\theta=\pi/2$) This is exactly $A_1\oplus A_1$, and corresponds to the root system D_2.
2. ($\theta=\pi/3$) Here $\langle\alpha,\beta\rangle=\langle\beta,\alpha\rangle=1$, so α,β are equal length with angle $\pi/3$ between them. By applying reflections, we get the root system A_2, corresponding to $\mathfrak{sl}(3)$.
3. ($\theta=\pi/4$) Here $\langle\alpha,\beta\rangle\langle\beta,\alpha\rangle=2$, so there is a choice of factorization.
   (a) If we pick $\langle\alpha,\beta\rangle=1$ and $\langle\beta,\alpha\rangle=2$, then β is $\sqrt{2}$ longer than α. By applying reflections, we get the root system B_2, corresponding to $\mathfrak{so}(2n+1)$.
   (b) Alternatively, if we pick $\langle\alpha,\beta\rangle=2$, then we get the root system C_2, corresponding to $\mathfrak{sp}(2n)$.
4. ($4\cos^2\theta=3$) This gives the exceptional root system G_2.

Take $e\in\mathfrak{g}_\alpha$. Via the Killing form, $\mathfrak{g}_{-\alpha}=\mathfrak{g}_\alpha^*$. We know $[e,f]\in\mathfrak{h}$. To know which element in $\mathfrak{h}$, it is enough to pair it using the Killing form:

$$([e,f],h)=(e,[f,h])=(e,\alpha(h)f)=2\frac{\alpha(h)}{(\alpha,\alpha)}.$$

If we identify $\mathfrak{h}\cong\mathfrak{h}^*$ via the Killing form, we can think of α as an element in $\mathfrak{h}$, so that $([e,f],h)=2(\alpha,h)/(\alpha,\alpha)$.

Definition 5.2.4. Write $h_\alpha\coloneqq 2\alpha/(\alpha,\alpha)$, also sometimes denoted $\alpha^\vee$.

Proposition 5.2.5. *The elements e, f, h_α form a copy of $\mathfrak{sl}(2)$, and up to scalars, h_α is the same vector regardless of the choice of e and f.*

Proof. We just computed $[e, f] = h_\alpha$, and we know that

$$[h_\alpha, e] = \alpha(h_\alpha)e = 2\frac{(\alpha,\alpha)}{(\alpha,\alpha)}e = 2e, \quad [h_\alpha, f] = -2f. \qquad \square$$

Corollary 5.2.6. *The dimension of $\mathfrak{g}_\alpha$ is 1, and if $\alpha \in \Delta$, then $n\alpha \notin \Delta$ for $n \neq \pm 1$.*

Proof. Consider the action of $\mathfrak{sl}(2)_\alpha \coloneqq \operatorname{span}\{e, f, h_\alpha\}$ on $\mathbb{C}h_\alpha \oplus \bigoplus_{n \in \mathbb{Z}_{\neq 0}} \mathfrak{g}_{n\alpha}$. Then e is a raising operator and f is a lowering operator, i.e. $[e, \mathfrak{g}_{n\alpha}] \subset \mathfrak{g}_{(n+1)\alpha}$, and similarly for f. But this whole thing is a finite-dimensional $\mathfrak{sl}(2)$-module with a 1-dimensional space of weight 0 (with respect to h_α) and with even weights. By the representation theory of $\mathfrak{sl}(2)$, this representation is irreducible. But it contains $\mathfrak{sl}(2)$, and is therefore equal to $\mathfrak{sl}(2)$. $\square$

Similarly, take $\beta \notin \mathbb{Z}\alpha$, and look at $\bigoplus_{n \in \mathbb{Z}} \mathfrak{g}_{\beta + n\alpha}$. Then e raises, f lowers, and h_α acts by the scalar $\langle \beta, \alpha \rangle$ on $\mathfrak{g}_\beta$. By the corollary, each $\mathfrak{g}_{\beta + m\alpha}$ has dimension either 0 or 1.

Corollary 5.2.7. *This representation is irreducible, $\langle \beta, \alpha \rangle \in \mathbb{Z}$, and for any $\beta \in \Delta$, the vector $r_\alpha(\beta) \coloneqq \beta - \langle \beta, \alpha \rangle \alpha$ is also in Δ.*

Proof. Any finite-dimensional $\mathfrak{sl}_2$ representation has weight spaces symmetric across the origin. But each weight space here has dimension either 0 or 1, so this representation cannot split. Also, $\langle \beta, \alpha \rangle$ is the scalar that h_α acts by on $\mathfrak{g}_\beta$, and we know for finite-dimensional representations that this is an integer. Finally, $r_\alpha(\beta)$ is precisely the weight corresponding to reflecting β across the origin. $\square$

We have shown that the set of weights of $\operatorname{ad}(\mathfrak{h})$ is a root system. It remains to show that it is reduced.