M1 2025: Introduction to Lie groups and Lie algebras

Elisha FALBEL

April 3, 2026

Contents

1 Topological groups

In this first chapter we will introduce the basic properties of topological groups which we will use in the following chapter when we define Lie groups.

1.1 Topological groups

Definition 1.1. *A topological group is a group with a topology such that the group operations, product $G \times G \to G$ and inverse map $G \to G$*

$$(g,h) \to gh \text{ and } g \to g^{-1}$$

are continuous.

Remark 1.2. *Fixing an element $h \in G$ we let $L_h : G \to G$ be the composition of maps*

$$G \to G \times G \to G,$$

given by $g \to (h,g) \to hg$. It is clear that L_h is a homeomorphism with inverse map given by $L_{h^{-1}}$. Analogously, the map $R_h : G \to G$ defined by $h \to hg$ is a homeomorphism.

Example 1.3. *Any group may be seen as a topological group by defining the discrete topology where a basis of open sets is given by $(\{g\})_{g \in G}$.*

Example 1.4. *The set of rational numbers with its usual topology, $\mathbf{Q}$, is a totally disconnected group. Any neighbourhood of the identity is not connected.*

Example 1.5. *The additive group $\mathbb{R}$ and the multiplicative group $\mathbb{R}_+^*$ are connected topological groups.*

Example 1.6. *The group $\mathrm{U}(1) = \{ s \in \mathbb{C} \mid |z| = 1 \}$ is a compact group.*

Example 1.7. *A product of a finite number of topological groups is also a topological group. But also note that the direct product of a countable number of topological groups is a topological group with the product topology. In particular $\mathrm{U}(1)^{\mathbf{N}}$ is a compact topological group by Tichonoff theorem.*

1.2 General properties

Proposition 1.8. *A topological group is Hausdorff if and only if the set $\{e\}$ (where $e \in G$ is the identity element) is a closed set.*

Proof. Recall that a topological space is Hausdorff if and only if the diagonal in $X \times X$ is a closed set. By the definition of topological group, the map $\phi : G \times G \to G$ given by $(x, y) \to xy^{-1}$ is continuous and the diagonal of $G \times G$ is precisely $\phi^{-1}(e)$. □

Proposition 1.9. *If a subgroup $H \subset G$ in a topological group is open then it is also closed.*

Proof. Observe that the complement $G \setminus H$ is open. Indeed, for any $g \in G \setminus H$, the set $L_g H \subset G \setminus H$ is open. □

Proposition 1.10. *Let $G_0 \subset G$ be the connected component of the identity in a topological group G. Then G_0 is a normal subgroup of G.*

Proof. It is a subgroup because if $g \in G_0$ then $L_{g^{-1}} G_0$ is a connected subset containing e. Therefore $L_{g^{-1}} G_0 \subset G_0$. Also for any $g \in G$, gG_0g^{-1} contains e and is connected. Therefore $gG_0g^{-1} \subset G_0$ and we conclude that G_0 is a normal subgroup. □

Proposition 1.11. *Let $V \subset G$ be a neighborhood of the identity of a topological group G whose connected component containing the identity is G_0. Then*

$$\bigcup_{n=1}^{\infty} V^n \supset G_0.$$

Proof. Here V^n stands for all products $g_1 \cdots g_n$ with $g_i \in V$ for $1 \leq i \leq n$. Note that $\bigcup_{n=1}^{\infty} V^n$ is an open set because if $g = g_1 \cdots g_n \in V^n$ then $gV \in V^{n+1}$ is a neighborhood of g. Consider instead the symmetric neighborhood $W = V \cap V^{-1}$. Clearly, $\bigcup_{n=1}^{\infty} W^n$ is an open subgroup of G which, by proposition 1.8, is closed. Therefore it is a union of connected components. We have $\bigcup_{n=1}^{\infty} W^n \subset \bigcup_{n=1}^{\infty} V^n$ and therefore $G_0 \subset \bigcup_{n=1}^{\infty} V^n$. □

Remark 1.12. *If the neighborhood V is connected then clearly $G_0 = \bigcup_{n=1}^{\infty} V^n$. But there exists topological groups which are not locally connected.*

1.3 Example: The p-adic integers $\mathbb{Z}_p$

There is an important abelian topological group which is homeomorphic to a Cantor set. Let p be a prime and consider the infinite direct product with the product topology,

$$\prod_{i=1}^{\infty} \mathbb{Z}/p^i\mathbb{Z},$$

which is a totally disconnected Hausdorff compact abelian group. The group of p-adic integers $\mathbb{Z}_p$ is a subgroup of this group given by sequences $(x_1, x_2, \cdots)$ such that

$$\pi_i(x_{i+1}) = x_i,$$

where $\pi_i : \mathbb{Z}/p^{i+1}\mathbb{Z} \to \mathbb{Z}/p^i\mathbb{Z}$ is the canonical projection with kernel $p^i\mathbb{Z}/p^{i+1}\mathbb{Z}$.

This is an example of a Hausdorff, compact and totally disconnected group (a profinite group).

1.4 Example: Bounded operators on a Banach space

Let X be a Banach space and the set of bounded linear operators:

$$B(X) = \{ A \in X \to X \mid ||A|| = \sup_{x \neq 0} \frac{||A(x)||}{||x||} < \infty \}.$$

The set $B(X)$ itself is a Banach space. We consider now the set of invertible operators

$$\mathrm{GL}(X) = \{ A \in B(X) \mid A \text{ has a bounded inverse} \}.$$

Theorem 1.13. *The set of invertible operators* $\mathrm{GL}(X)$ *(with group operation the composition of operators) is an open subset of* $B(X)$ *which is a topological group.*

The proof is based on the following proposition.

Proposition 1.14. *Let X be a Banach space and $A : X \to X$ be a linear operator with* $||A|| < 1$. *Then $Id - A$ is invertible and*

$$(Id - A)^{-1} = \sum_{n=0}^{\infty} A^n.$$

Moreover $||(Id - A)^{-1}|| \leq \frac{1}{1 - ||A||}$.

Proof. The series $\sum_{n=0}^{\infty} A^n$ converges normally because $||A^n|| \leq ||A||^n$. This implies that

$$\lim_{N \to \infty} (Id - A) \left(\sum_{n=0}^{N} A^n \right) = \lim_{N \to \infty} (Id - A^{N+1}) = Id.$$

□

Back to the proof of the theorem:

Proof. We first verify that $\mathrm{GL}(X) \subset B(X)$ is an open subset. Let $A \in \mathrm{GL}(X)$ and consider the operators B such that $||A-B|| < \varepsilon$. Then

$$||Id - A^{-1}B|| = ||A^{-1}(A-B)|| \leq ||A^{-1}||||A-B|| < \varepsilon||A^{-1}||.$$

By the proposition $A^{-1}B$ is invertible with bounded inverse if $||\varepsilon||A^{-1}|| < 1$ and therefore B is invertible if it is contained in a small neighbourhood of A.

It remains to show that the map $A \to A^{-1}$ is continuous. We will use the estimate given in the proposition. First write

$$||B^{-1} - A^{-1}|| = ||B^{-1}(A-B)A^{-1}|| \leq ||B^{-1}||||(A-B)||||A^{-1}||.$$

We want to estimate this difference in terms of $||(A-B)||$ and $||A||$. It remains to estimate $||B^{-1}||$ in terms of $||(A-B)||$ and $||A||$. Observe that $B = A(Id - A^{-1}(A-B))$. Therefore if $||A^{-1}(A-B)|| < 1$, by the previous proposition,

$$||\left(Id - A^{-1}(A-B)\right)^{-1}|| \leq \frac{1}{1-||A^{-1}(A-B)||} \leq \frac{1}{1-||A^{-1}||||(A-B)||}.$$

This gives an estimate $||B^{-1}|| \leq \frac{1}{1-||A^{-1}||||(A-B)||}||A^{-1}||$. □

1.5 The definition of a Lie group

Definition 1.15. *A Lie group is a topological group which is a smooth manifold and such that the group operations, product* $G \times G \to G$ *and inverse map* $G \to G$

$$(g,h) \to gh \text{ and } g \to g^{-1}$$

are smooth.

In our definition, a manifold is always second countable and Hausdorff.

Example 1.16. *1. The additive group* $\mathbb{R}^n$ *is a Lie group.*

Remark 1.17. *We will deal, in almost all sections of these notes, with Lie groups which are submanifolds of* $\mathrm{Gl}_n(\mathbb{R})$. *They are called linear Linear groups. We will not assume any knowledge of the general theory of differential manifolds for the main results.*

Proposition 1.18. *The group of invertible matrices of rank n,* $\mathrm{Gl}_n(\mathbb{R})$*, is a Lie group.*

Proof. We consider the group $\mathrm{Gl}_n(\mathbb{R})$ as a subset of the space of $n \times n$ matrices $\mathrm{M}_n(\mathbf{R}$ which is homeomorphic to $\mathbb{R}^{n^2}$. In fact,

$$\mathrm{Gl}_n(\mathbb{R}) = \{ A \in \mathrm{M}_n(\mathbf{R} \mid \det A \neq 0 \}$$

which is an open subset because the determinant map $\det : \mathrm{M}_n(\mathbf{R} \to \mathbb{R}$ is continuous (it is moreover smooth). The product of matrices and the inverse of a matrix are given by smooth maps on the coefficients of the matrices and therefore $\mathrm{Gl}_n(\mathbb{R})$ is a Lie group.

□

1.6 Examples

We will give several examples and concentrate on the fact that they are topological groups. We only discuss the smooth structure of $\mathrm{SL}(n,\mathbb{R})$ for the moment.

1.6.1 $\mathrm{U}(n,\mathbb{R})$

We define

$$\mathrm{U}(n,\mathbb{R}) = \{ g \in \mathrm{Gl}_n(\mathbb{R}) \mid \det g = 1 \}.$$

This is a closed subgroup of $\mathrm{Gl}_n(\mathbb{R})$ as the determinant is a continuous map. Moreover it is smooth and one can prove that $\mathrm{U}(n,\mathbb{R})$ is a submanifold of $\mathrm{Gl}_n(\mathbb{R})$ by showing that it is of rank one. We compute the differential of the determinant map:

$$D(\det)_g(H) = \det(g)\,\mathrm{tr}\big(g^{-1}H\big).$$

Indeed, since g is invertible, we write

$$\det(g + tH) = \det(g)\det(I + tg^{-1}H).$$

Using the expansion

$$\det(I + tX) = 1 + t\,\mathrm{tr}(X) + o(t),$$

the formula follows from

$$\det(g + tH) = \det(g)\big(1 + t\,\mathrm{tr}(g^{-1}H) + o(t)\big).$$

From this formula it is clear that det is a map of constant rank one.

1.6.2 $\mathrm{O}(p,q)$

Define the symmetric matrix

$$B_{pq} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}$$

where I_p and I_q are the identity matrix of dimension p and q respectively. Define

$$\mathrm{O}(p,q,\mathbb{R}) = \{\ g \in \mathrm{Gl}_n(\mathbb{R}) \mid g^T B_{pq} g = B_{pq}\ \}.$$

1.6.3 $\mathrm{Sp}(n,\mathbb{R})$

Define the skew-symmetric matrix

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

where I_n is the identity matrix of dimension n. The symplectic group is defined as

$$\mathrm{Sp}(n,\mathbb{R}) = \{\ g \in GL(2n,\mathbb{R}) \mid g^T J g = J\ \}.$$

Remark 1.19. *We may define more generally the group associated to a bilinear form. Consider a non-degenerate bilinear $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$. One defines*

$$\mathrm{O}(b,\mathbb{R}) = \{\ g \in \mathrm{Gl}_n(\mathbb{R}) \mid b(gu, gv) = b(u,v) \text{ for all } u, v \in \mathbb{R}^n\ \}.$$

1.6.4 The upper triangular matrices $T(n,\mathbb{R})$

The upper triangular matrices in $\mathrm{Gl}_n(\mathbb{R})$ form a solvable group.

1.6.5 $\mathrm{Gl}_n(\mathbb{C})$, $\mathrm{U}(p,q)$

One could repeat the definitions above using complex matrices and obtain the subgroup $\mathrm{Gl}_n(\mathbb{C}) \subset GL(2n,\mathbb{R})$. Also, one defines

$$\mathrm{U}(p,q) = \{\ g \in \mathrm{Gl}_n(\mathbb{C}) \mid \bar{g}^T B_{pq} g = B_{pq}\ \}.$$

1.7 Hilbert 5th problem

The problem, in modern language, is to characterize topological groups which are Lie groups. A characterization of locally compact topological groups which are Lie groups is obtained by imposing that they have a neigborhood of the identity with no subgroups different than $\{e\}$. (work by Gleason, Montgomery and Zippin). Note that, for the p-adic numbers, any open subset containing e contains the subgroup $p^N\mathbb{Z}_p$ for sufficiently large N.

1.8 Group actions and homogeneous spaces

A left (right) continuous group action $G \times M \to M$ ($M \times G \to M$) of a topological group on a Hausdorff topological space is a continuous map satisfying, for all $g, h \in G$ and $m \in M$,

- $em = m$ ($me = m$)
- $(gh)m = g(hm)$ ($m(gh) = (mg)h$)

Consider a left action $G \times M \to M$. The isotropy at a point $m \in M$ is the closed (because M is Hausdorff) subgroup

$$G_m = \{ g \in G \mid gm = m \}.$$

The action is said to be transitive if given two points m_1 and m_2, there exists an element $g \in G$ such that $m_2 = gm_1$. We say, in this case, that M is a homogeneous space. Note that this implies that the isotropy of the two points are conjugated. Indeed

$$G_{m_1} = gG_{m_2}g^{-1}.$$

Homogeneous spaces are identified to quotients of Hausdorff topological groups by closed subgroups. First, define the quotient in general:

Definition 1.20. *Given a subgroup H in a topological group G, define G/H as the quotient space where $g_1 \sim g_2$ if and only if $g_1 = g_2 h$ where $h \in H$. The map $\pi : G \to G/H$ is called the quotient map and we define $A \subset G/H$ to be an open set if and only if $\pi^{-1}(A) \subset G$ is open.*

We sometimes write $\pi(g) = [g]$.

Lemma 1.21. *The quotient map $\pi : G \to G/H$ is an open map.*

Proof. Let $U \subset G$ be open. Then $\pi(U)$ is open if and only if $\pi^{-1}(\pi(U))$ is open. But $\pi^{-1}(\pi(U)) = \bigcup_{h\in H} Uh$ which is an open set.

□

The Hausdorff condition on G and the closed condition on H guarantee that the quotient is Hausdorff:

Proposition 1.22. *Let G be a Hausdorff topological group.*

1. *Let $H \subset G$ be a closed subgroup. Then G/H is a Hausdorff topological space which is a homogeneous space for the action $G \times G/H \to G/H$ given by $(g, [h]) \to [gh]$.*

2. *If $G \times M \to M$ is a transitive action then for any $m \in M$, the map $G/I_m \to M$ defined by $[g] \to gm$ is a continuous bijection.*

Proof. First we show that G/H is a Hausdorff topological space. Consider the continuous map $\phi : G \times G \to G$ given by $(g,h) \to gh^{-1}$. We have that $\phi^{-1}(H)$ is closed. Let $x_1 = [g_1]$ and $x_2 = [g_2]$ be two distinct elements in G/H, that is $g_1 H \cap g_2 H = \emptyset$. This implies that (g_1, g_2) is in the complement of $\phi^{-1}(H)$. Therefore, there exits open subsets U_1 and U_2 containing respectively g_1 and g_2 with $U_1 \times U_2$ in that complement. Clearly $\pi(U_1)$ and $\pi(U_2)$ are open sets containing respectively $[g_1]$ and $[g_2]$ with empty intersection.

The map $G \times G/H \to G/H$ is an action and it is continuous: From the commutative diagram

$$\begin{array}{ccc} G \times G & \longrightarrow & G \\ \downarrow & & \downarrow \\ G \times G/H & \longrightarrow & G/H \end{array},$$

we conclude that if $A \subset G/H$ is an open subset its inverse image in $G \times G/H$ is an open subset.

In order to prove the second assertion, observe that $G \to M$ defined by $g \to gm$ is a continuous map which factors through the quotient G/I_m. The continuity of $G/I_m \to M$ follows therefore from the continuity of $G \to M$. Note that the map $G/I_m \to M$ might not be open and therefore it might not be a homeomorphism.

□

Remark 1.23. *If G and M are locally compact and second countable topological spaces then one may conclude that for a transitive action $G \times M \to M$ that G/I_m is homeomorphic to M. This is the case for Lie groups acting on manifolds. We will use this result in the sequel of these notes.*

Lemma 1.24. *Let G be a topological group and $H \subset G$ be a closed subgroup. Then, if H and G/H are connected then G is connected.*

Proof. By contradiction, suppose $G = U_1 \cup U_2$ with U_1 and U_2 disjoint open subsets. Then $\pi(U_1)$ and $\pi(U_2)$ are open sets. As G/H is connected, there is a point $[g] \in \pi(U_1) \cap \pi(U_2)$. But $gH \subset G$ is a connected subset in G and therefore, either $gH \subset U_1$ or else $gH \subset U_2$ and we conclude that $\pi(U_1) \cap \pi(U_2) = \emptyset$. □

1.8.1 Examples

1. The right or left action of a Hausdorff topological group on itself by the product $(g,h) \to gh$ is a transitive action with trivial isotropy.

2. One can identify $\mathrm{Gl}_n(\mathbb{R})$ with the set of basis $\mathscr{B}$ in $\mathbb{R}^n$. Indeed the set of column vectors of an element $g \in \mathrm{Gl}_n(\mathbb{R})$ define the bijection between $\mathrm{Gl}_n(\mathbb{R})$ and $\mathscr{B}$. It is convenient to define a right action $\mathscr{B} \times \mathrm{Gl}_n(\mathbb{R}) \to \mathrm{Gl}_n(\mathbb{R})$. A basis $b = (v_1, \cdots v_n)$ is written as a columns in a matrix and the action is then a multiplication of matrices

$$\big(b, g\big) \to (v_1, \cdots v_n)g.$$

This action is transitive with trivial isotropy.

3. Analogously we identify $\mathrm{O}(n,\mathbb{R})$ with the set of orthonormal basis $\mathscr{B}^o$. To each element $g \in \mathrm{O}(n)$ we consider the basis formed by the column vectors. Clearly this is a bijection and one may define a right action as before $\mathscr{B}^o \times \mathrm{Gl}_n(\mathbb{R}) \to \mathrm{Gl}_n(\mathbb{R})$. It is transitive with trivial isotropy. One can define positive basis as those basis corresponding to matrices in $S\mathrm{O}(n,\mathbb{R})$.

4. The action of $\mathrm{SO}(n)$ on S^n:

   Define the unit sphere in $\mathbb{R}^n$ $(n \geq 2)$ as the set of vectors of norm one:

$$S^n = \{\, x \in \mathbb{R}^n \mid |x| = 1 \,\}.$$

   The action $\mathrm{SO}(n) \times S^n \to S^n$ is defined as

$$(g, x) \to gx.$$

   It is clearly a continuous action. It is a transitive action because $\mathrm{SO}(n)$ acts transitively on positive basis and therefore, in particular, on vectors of norm one (one can complete it to a positive orthonormal basis).

   The isotropy at the point $x = (1, 0, \cdots, 0)^T$ is the subgroup $\mathrm{SO}(n-1)$ (for $n \geq 2$). Observe that, by induction, starting with the known result that $\mathrm{SO}(2)$ is connected and that S^n is connected for $n \geq 1$ one obtains from the last lemma that $\mathrm{SO}(n)$ is connected.

2 The general linear group $\mathrm{Gl}_n(\mathbb{R})$

In this section we prove two decompositions theorems, the polar decomposition and the Gram decomposition which have a generalization to Lie groups as the Cartan and KAN decomposition respectively.

2.0.1 The polar decomposition

Definition 2.1. *We define P_n to be the set of positive definite symmetric matrices of dimension n.*

Recall that a matrix A is said to be positive definite if $u^T Au > 0$ for all $u \in \mathbb{R}^n$. It is easy to verify that P_n is a convex cone in the vector space defined by the symmetric matrices. One can prove that it is also an open subset in the set of symmetric matrices.

Theorem 2.2. *For every $g \in \mathrm{Gl}_n(\mathbb{R})$ there exists a unique decomposition $g = kp$ where $k \in \mathrm{O}(n)$ and $p \in P_n$. Moreover, the map $\mathrm{Gl}_n(\mathbb{R}) \to \mathrm{O}(n) \times P_n$ is a homeomorphism.*

Proof. We first prove existence. The idea is to work with $g^T g$ which is symmetric and therefore diagonalizable. If the decomposition exists then $g^T g = p^T k^T kp = p^2$. From the fact that $g^T g$ is positive definite we may write $g^T g = hDh^{-1}$, where $h \in \mathrm{O}(n)$ and D is diagonal with all entries strictly positive. We define then

$$p = h\sqrt{D}h^{-1},$$

where $\sqrt{D}$ is the diagonal matrix with entries the square roots of the entries of D. Clearly $p \in P_n$. We verify that $gp^{-1} \in \mathrm{O}(n)$ by the computation

$$kk^T = (gp^{-1})^T gp^{-1} = p^{-1}p^2p^{-1} = I_n.$$

In order to prove unicity, suppose $g = k_1p_1$. But ${p_1}^2 = p^2 = g^T g$, which implies that p_1 and p commute with $g^T g$ and therefore have the same eigenspaces. On each eigenspace p_1 and p have the same eignevalue by definition (namely, the square root of the eigenvalue of $g^T g$). This proves that $p_1 = p$ and therefore $k_1 = k$.

We now prove that the decomposition defines a homeomorphism. Clearly, the map $\mathrm{Gl}_n(\mathbb{R}) \leftarrow \mathrm{O}(n) \times P_n$ is continuous. To prove that $\mathrm{Gl}_n(\mathbb{R}) \to \mathrm{O}(n) \times P_n$ is continuous we consider a convergent sequence $g_n = k_n p_n$, $g_n \to g$. We have to prove that convergence $k_n \to k$ and $p_n \to p$ holds satisfying moreover $g = kp$. But $\mathrm{O}(n)$ is compact and therefore there exists a convergence subsequence $k_i \to k$. This implies that $p_i = {k_i}^{-1} g_i$ converges. Therefore $k_i p_i \to g$ and we conclude that $kp = g$. All subsequences of k_n have a subsequence converging to the same matrix k. This implies that k_n converges to k and the result follows. □

Corollary 2.3. *For every $g \in \mathrm{Gl}_n(\mathbb{R})$ there exists a unique decomposition $g = k_1Dk_2$ where $k_1, k_2 \in \mathrm{O}(n)$ and D is a positive definite diagonal matrix.*

Proof. By the previous theorem $g = kp$ where $k \in \mathrm{O}(n)$ and p positive symmetric. We may then write $p = k'Dk_2$ where D is positive and diagonal and $k', k_2 \in \mathrm{O}(n)$. □

2.0.2 The Gram decomposition

The Gram decomposition is related to the Gram-Schmidt orthonormalization process and will be generalized to general Lie groups as the KAN decomposition.

Theorem 2.4. *For every $g \in \mathrm{Gl}_n(\mathbb{R})$ there exists a unique decomposition $g = kt$ where $k \in \mathrm{O}(n)$ and $t \in T^+(n,\mathbb{R})$ (that is, triangular matrices with positive diagonal coefficients). Moreover, the map $\mathrm{Gl}_n(\mathbb{R}) \to \mathrm{O}(n) \times T^+(n,\mathbb{R})$ is a homeomorphism.*

Proof. The existence is precisely the Gram-Schmidt orthonormalization process. Recall that one may identify an element $g \in \mathrm{Gl}_n(\mathbb{R})$ to a basis $b \in \mathscr{B}$ of $\mathbb{R}^n$. The basis b is formed by the column vectors of g. The product gh is a change of basis where the columns are the new basis vectors.

The Gram-Schmidt orthonormalization process starts with a basis and transforms it into an orthonormal basis through a product by a triangular matrix. Indeed, if $b = (v_1, \cdots, v_n)$ then the new basis is constructed by first defining $v_1' = \frac{v_1}{|v_1|}$ and then, by recurrence, having defined v_i', we define v_{i+1}' to be

$$v_{i+1} - \sum_{k=1}^{i} (v_{i+1}, v_i) v_i$$

divided by its norm. That gives $gt^{-1} = k$ for a certain upper triangular matrix t with positive coefficients in the diagonal and k an orthogonal matrix.

Unicity follows from the fact that if $g = kt = k't'$ we obtain $k'^{-1}k = t't^{-1}$. But $\mathrm{SO}(n) \cap T^+(n,\mathbb{R}) = \{Id\}$.

A proof that the map $\mathrm{Gl}_n(\mathbb{R}) \to \mathrm{O}(n) \times T^+(n,\mathbb{R})$ is a homeomorphism may be given in a similar way to the proof that the polar decomposition defines a homeomorphism. One can also observe that the Gram-Schmidt procedure gives an explicit continuous map from $\mathrm{Gl}_n(\mathbb{R})$ to $T^+(n,\mathbb{R})$.

□

3 The exponential map

3.1 Exponential of a matrix

Let $\mathrm{M}_n(K)$ be space of $n \times n$ matrices with coefficients in the field K ($K = \mathbb{R}$ or $\mathbb{C}$). We consider it as a normed vector space with norm, for $X \in \mathrm{M}_n(K)$, defined by

$$||X|| = \sup_{u \neq 0} \frac{|X(u)|}{|u|},$$

where $|u|$ is the usual norm in $\mathbb{R}^n$ or the hermitian norm in $\mathbb{C}^n$. In the following we will consider mainly $K = \mathbb{R}$.

Definition 3.1. *For $X \in \mathrm{M}_n(K)$ we define*

$$\exp(X) = \sum_{k=0}^{\infty} \frac{X^k}{k!}.$$

From Cauchy-Schwartz inequality and the property $||XY|| \leq ||X||||Y||$ we obtain that

$$||\sum_{k=0}^{N} \frac{X^k}{k!}|| \leq \sum_{k=0}^{N} \frac{||X^k||}{k!} \leq \sum_{k=0}^{N} \frac{||X||^k}{k!}$$

We conclude that the series is normally convergent on compact sets.

Theorem 3.2. *Let $X \in \mathrm{M}_n(\mathbb{R})$ be a fixed matrix. Then the map* $\exp : \mathbb{R} \to \mathrm{M}_n(\mathbb{R})$ *defined by*

$$t \to \exp tX$$

is analytic and is the unique solution to the differential equation

$$\frac{dY}{dt} = XY$$

with initial condition $Y(0) = Id$.

Proof. This is a classic theorem in the theory of linear differential equations. Here we just verify that the exponential satisfies the equation. Write $\exp tX = \sum_{k=0}^{\infty} \frac{t^k X^k}{k!}$ and observe that the series

$$\sum_{k=0}^{\infty} \frac{k t^{k-1} X^k}{k!} = X \sum_{k=0}^{\infty} \frac{t^k X^k}{k!}$$

is also normally convergent. We obtain then

$$\frac{d \exp tX}{dt} = X \exp tX.$$

□

Proposition 3.3. *The exponential map satisfies the following properties*

1. *For any* $g \in \mathrm{Gl}_n(\mathbb{R})$ *and* $X \in \mathrm{M}_n(\mathbb{R})$
$$g(\exp X)g^{-1} = \exp(gXg^{-1}).$$

2. *If* $X, Y \in \mathrm{M}_n(\mathbb{R})$ *are two commuting matrices then*
$$\exp(X+Y) = \exp X \exp Y.$$

3. *For any* $X \in \mathrm{M}_n(\mathbb{R})$
$$\exp(-X) = (\exp X)^{-1}.$$

Proof. The first property follows from the fact that $(gXg^{-1})^n = gX^n g^{-1}$ and the definition of the exponential.

The second property may be proven using the unicity in theorem 3.2 describing the exponential via a differential equation. Indeed,

$$\frac{d}{dt}\exp tX \exp tY = X \exp tX \exp tY + (\exp tX) Y \exp tY = (X+Y)\exp tX \exp tY.$$

Therefore $\exp tX \exp tY$ satisfies the same equation as $\exp t(X+Y)$ with same initial condition.

The third property is a consequence of 2. and the formula $\exp(X-X) = Id$.

□

3.2 Computations

1. If X is diagonalizable then one writes
$$X = g\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & & 0 \\ 0 & 0 & \lambda_n \end{pmatrix} g^{-1}$$
and use the first property of the proposition above to get
$$\exp X = g\begin{pmatrix} e^{\lambda_1} & 0 & 0 \\ 0 & & 0 \\ 0 & 0 & e^{\lambda_n} \end{pmatrix} g^{-1}.$$

2. If $X = \lambda Id + N$ is a Jordan block, with

$$N = \begin{pmatrix} 0 & 1 & 0 \\ . & . & . \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

we obtain

$$\exp tX = \exp(t(\lambda Id + N)) = \exp(t\lambda Id)\exp tN,$$

where, because the matrix N satisfies $N^n = 0$, we may compute using a finite sum

$$\exp tN = \sum_{k=0}^{n-1} \frac{t^k N^k}{k!}$$

and

$$\exp(t\lambda Id) = \begin{pmatrix} e^{t\lambda} & 0 & 0 \\ 0 & & 0 \\ 0 & 0 & e^{t\lambda} \end{pmatrix}$$

Proposition 3.4 (Jacobi's formula)**.** *For any* $X \in M_n(K)$ *we have*

$$\det(\exp X) = e^{\mathrm{tr}(X)}.$$

Proof. Any matrix in $M_n(\mathbb{C})$ may be transformed into an upper triangular matrix by conjugation. We write

$$X = gYg^{-1}$$

where Y is upper triangular. Then

$$\det(\exp X) = \det(\exp Y) = e^{\mathrm{tr}(Y)} = e^{\mathrm{tr}(X)}.$$

□

This shows that the the image of the exponential map acting on $M_n(\mathbb{R})$ is contained in $GL^+(n,\mathbb{R})$, the invertible matrices with positive determinant.

Example 3.5. *The map* $exp : M_2(\mathbb{R}) \to GL_2^+(\mathbb{R})$ *is neither surjective nor injective.*

Compute first

$$\exp\begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix} = \sum_{k=0}^{\infty} \frac{1}{k!}\begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}^k = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} -\theta^2 & 0 \\ 0 & -\theta^2 \end{pmatrix} + \frac{1}{3!}\begin{pmatrix} 0 & -\theta^3 \\ \theta^3 & 0 \end{pmatrix} + \cdots$$

$$= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

1. The exponential map is not injective because

$$\exp\begin{pmatrix} 0 & 2\pi \\ -2\pi & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

2. It is not surjective: One can always trigonalize a complex matrix A over $\mathbb{C}$. That is, write, for $\lambda_1, \lambda_2 \in \mathbb{C}$ and $g \in \mathrm{GL}_n(\mathbb{C})$,

$$A = g\begin{pmatrix} \lambda_1 & * \\ 0 & \lambda_2 \end{pmatrix} g^{-1}.$$

We have then

$$\exp A = g\begin{pmatrix} e^{\lambda_1} & * \\ 0 & e^{\lambda_2} \end{pmatrix} g^{-1}.$$

In particular, if a matrix B has oposite signs real eigenvalues it is not in the image of the exponential map.

3.3 $\exp : Sym_n \to P_n$

Theorem 3.6. *The exponential map restricted to the real symmetric matrices* $\exp : Sym_n \to P_n$ *is homeomorphism onto the cone of positive definite symmetric matrices.*

Proof. Surjectivity: Let $p \in P_n$. One can diagonalize the matrix by conjugating by an orthogonal element:

$$p = k\begin{pmatrix} \lambda_1 & 0 & 0 \\ \cdot & \cdot & \cdot \\ 0 & 0 & \lambda_n \end{pmatrix} k^{-1}$$

with positive diagonal coefficients. Therefore

$$p = \exp\left(k\begin{pmatrix} \ln\lambda_1 & 0 & 0 \\ \cdot & \cdot & \cdot \\ 0 & 0 & \ln\lambda_n \end{pmatrix} k^{-1} \right).$$

Injectivity: Observe first that the eigenspaces of a symmetric matrix are the same as those of its exponential. Moreover, the eigenvalues of the exponential matrix are the exponentials of the eigenvalues of the matrix. Suppose now that $\exp(X)$ and $\exp(Y)$ have the same eigenspaces and eigenvalues. By the observation, this implies $X = Y$.

The map is a homeomorphism: clearly $Sym_n \to P_n$ is a continuous map. It remains to show that the inverse map is continuous. Suppose $p_n \in P_n$ is a sequence converging to p. One needs to show that the sequence s_n (defined by $\exp(s_n) = p_n$) converges to s

with $\exp s_n = p$. For that sake, write $p_n = k_n d_n k_n^{-1}$ where d_n is diagonal and therefore $s_n = k_n \ln(d_n) k_n^{-1}$ where $\ln(d_n)$ is the diagonal matrix with coefficients equal to the logarithm of the diagonal coefficients of d_n. As $\mathrm{O}(n)$ is compact there exists a convergent subsequence of k_n. For any such convergent subsequence we obtain that d_n converges and $k_n d_n k_n^{-1} \to p$. we conclude that s_n converges and by injectivity therefore $s_n \to s$.

□

Remark 3.7. *From theorem 2.2 and the previous theorem we conclude that the map*

$$\mathrm{O}(n) \times Sym_n \to \mathrm{Gl}_n(\mathbb{R})$$

is a homeomorphism.

3.4 The exponential is a local diffeomorphism at the origin

The image of the null matrix by the exponential map is the identity matrix. There is a diffeomorphism if one restricts to a sufficiently small open set containing the null matrix to an open subset containing the identity.

Theorem 3.8. *The exponential map* $\exp : \mathrm{M}_n(\mathbb{R}) \to \mathrm{Gl}_n(\mathbb{R})$ *is a diffeomorphism on a neighborhood of the null matrix onto its image.*

Proof. We need to show that the differential is an isomorphism and apply the local inversion theorem. We compute the differential of exp at the origin:

$$D\exp_0(H) = \lim_{t\to 0} \frac{\exp(0+tH) - \exp(0)}{t} = \lim t \to 0 H \exp tH = H.$$

Therefore $D\exp_0 = Id$ which is an isomorphism.

□

Remark 3.9. *The exponential is not a global diffeomorphism as example 3.5 shows.*

Remark 3.10 (The logarithm)**.** *The fact that the exponential is a local homeomorphism at the origin allows us to define its inverse function which we also call the logarithm,* log*, defined on a neighborhood of the identity of* $\mathrm{Gl}_n(\mathbb{R})$ *with values in* $\mathrm{M}_n(\mathbb{R})$ *such that* $\exp \circ \log = Id$. *But note that* log *is not defined for all element in* $\mathrm{Gl}_n(\mathbb{R})$. *An explicit series which compute the logarithm of* $U \in \mathrm{Gl}_n(\mathbb{R})$ *with* $||U - Id|| < 1$ *is*

$$\log U = \sum_{k=1}^{\infty} \frac{(-1)^k}{k} (U - Id)^{-1}.$$

In order to compute the differential of the exponential at a matrix A we introduce the adjoint map:

Definition 3.11. *Let $A \in \mathrm{M}_n(\mathbb{R})$ be a matrix. The adjoint map*

$$\mathrm{ad}_A : \mathrm{M}_n(\mathbb{R}) \to \mathrm{M}_n(\mathbb{R})$$

is defined by $\mathrm{ad}_A(X) = AX - XA$.

We may now describe the differential of the exponential map at A:

Theorem 3.12. *Let $A \in \mathrm{M}_n(\mathbb{R})$. Then*

$$D\exp_A(X) = \exp(A) \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!} (\mathrm{ad}_A)^k X.$$

Remark 3.13. *Remembering the formula $\sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!} z^k = \frac{1-\mathrm{e}^z}{z}$, one may write formally* $D\exp_A(X) = \exp(A)(\frac{Id - \mathrm{e}^{-\mathrm{ad}_A}}{\mathrm{ad}_A})(X)$.

The behaviour of products of exponentials near the origin is given in the following:

Proposition 3.14. *Let $X, Y \in \mathrm{M}_n(\mathbb{R})$. Then*

1. $\exp(tX)\exp(tY) = \exp\left(t(X+Y) + \frac{t^2}{2}[X,Y] + O(t^3)\right)$,

2. $\exp(tX)\exp(tY)\exp(-tX)\exp(-tY) = \exp\left(t^2[X,Y] + O(t^3)\right)$.

Proof. 1. We compute

$$\exp(tX)\exp(tY) = \left(I + tX + \frac{t^2}{2}X^2 + O(t^3)\right)\left(I + tY + \frac{t^2}{2}Y^2 + O(t^3)\right)$$

$$= I + t(X+Y) + \frac{t^2}{2}(X^2 + 2XY + Y^2) + O(t^3).$$

On the other hand

$$\exp\left(t(X+Y) + \frac{t^2}{2}[X,Y] + O(t^3)\right) = Id + t((X+Y) + \frac{t}{2}[X,Y]) + \frac{t^2}{2}(X+Y)^2 + O(t^3)$$

$$= I + t(X+Y) + \frac{t^2}{2}(XY - YX + X^2 + XY + YX + Y^2) + O(t^3)$$

which gives the same expression as above and therefore proves the first assertion.

2. Use the first result to obtain

$$\exp(tX)\exp(tY)\exp(-tX)\exp(-tY) =$$
$$\exp\left(t(X+Y)+\frac{t^2}{2}[X,Y]+O(t^3)\right)\exp\left(-t(X+Y)+\frac{t^2}{2}[X,Y]+O(t^3)\right).$$

Use it again so that

$$= \exp\left(t^2[X,Y]+O(t^3)\right).$$

□

Example 3.15. *The Heisenberg group* Heis^3 *is a nilpotent subgroup of* $\mathrm{GL}_3(\mathbb{R})$ *given by matrices of the form*

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix},$$

where $x, y, z \in \mathbb{R}$. *Consider the following matrices in* $M(3,\mathbb{R})$*:*

$$X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad Z = [X,Y] = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We compute

$$\exp(tX) = \begin{pmatrix} 1 & t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \exp(tY) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \exp(tZ) = \begin{pmatrix} 1 & 0 & t^2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

A computation shows hat in this case the first formula in the previous proposition is exact:

$$\exp(tX)\exp(tY) = \begin{pmatrix} 1 & t & t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} = \exp\left(t(X+Y)+\frac{t^2}{2}[X,Y]\right).$$

Observe that the exponential map is a diffeomorphism between upper triangular matrices with null diagonal with the Heisenberg group.

Corollary 3.16. *Let* $X, Y \in \mathrm{M}_n(\mathbb{R})$. *Then*

$$\lim_{k\to\infty}\left(\exp(\frac{1}{k}X)\exp(\frac{1}{k}Y)\right)^k = \exp(X+Y),$$

and

$$\lim_{k\to\infty}\left(\exp(\frac{1}{k}X)\exp(\frac{1}{k}Y)\exp(-\frac{1}{k}X)\exp(-\frac{1}{k}Y)\right)^{k^2} = \exp([X,Y]).$$

Proof. From the previous proposition we have

$$\left(\exp(\frac{1}{k}X)\exp(\frac{1}{k}Y)\right)^k = \left(\exp\left(\frac{1}{k}(X+Y)+O(\frac{1}{k^2})\right)\right)^k = \exp k\left(\frac{1}{k}(X+Y)+O(\frac{1}{k^2})\right),$$

which implies the first formula.

The second formula follows from

$$\exp(\frac{1}{k}X)\exp(\frac{1}{k}Y)\exp(-\frac{1}{k}X)\exp(-\frac{1}{k}Y) = \exp\left(\frac{1}{k^2}[X,Y]+O(\frac{1}{k^3})\right).$$

□

3.5 Exponential of linear transformations

The exponential of a linear transformation $A : V \to V$ of a finite dimensional vector space is defined as for the exponential of a matrix through an infinite series using the operator norm for a fixed norm $|\cdot|$ in V (that is, $||A|| = \sup_{v\neq 0} \frac{|Av|}{|v|}$).

Definition 3.17. *Let V be a finite dimensional normed vector space. Denote by* $\mathrm{End}(V)$ *the vector space of its linear operators and* $\mathrm{Gl}(V)$ *the group of its isomorphisms. Define*

$$\mathrm{Exp} : \mathrm{End}(V) \to \mathrm{Gl}(V)$$

by, for $A \in \mathrm{End}(V)$,

$$\mathrm{Exp}(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

The convergence of the series follows because for any linear maps A and B, $||A \circ B|| \leq ||A||||B||$ as for the series defining the exponential of matrices. Note, moreover, that the convergence is independent on the norm on V because all norms on a finite dimensional space are equivalent. Choosing a basis of the vector space, one can associate a matrix M_A to the linear map A and one obtains

$$M_{\mathrm{Exp}(A)} = \exp(M_A) = \sum_{k=0}^{\infty} \frac{1}{k!} {M_A}^k.$$

Remark 3.18. *We may define the exponential map on bounded operators on Banach spaces using the same series as above.*

4 Linear Lie groups and their Lie algebra

4.1 Linear Lie groups

Definition 4.1. *A linear Lie group is a closed subgroup of* $\mathrm{Gl}_n(\mathbb{R})$.

Remark 4.2. *In particular, a discrete subgroup of* $\mathrm{Gl}_n(\mathbb{R})$ *is a Lie group.*

Example 4.3. *The topological group* $\mathrm{GL}(n,\mathbb{C})$ *may be embedded as a closed subgroup of* $\mathrm{GL}(2n,\mathbb{R})$*: Let* $\iota : \mathrm{M}_n(\mathbb{C}) \to \mathrm{M}_{2n}(\mathbb{R})$ *be defined by*

$$A + iB \to \begin{pmatrix} A & -B \\ B & A \end{pmatrix}.$$

One can check that a vector $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^{2n}$ *satisfies*

$$\begin{pmatrix} A & -B \\ B & A \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0$$

if and only if $(A + iB)(v_1 + iv_2) = 0$*. Therefore, restricting the domain of* ι*, we obtain the injective homomorphism with closed image (check it!)* $\iota : \mathrm{GL}_n(\mathbb{C}) \to \mathrm{GL}_{2n}(\mathbb{R})$.

4.2 One parameter subgroups

Definition 4.4. *A one parameter subgroup of a topological subgroup G is a continuous homomorphism*

$$\gamma : \mathbb{R} \to G.$$

Example 4.5. *Let* $X \in \mathrm{M}_n(\mathbb{R})$ *be a matrix. The map* $\gamma : \mathbb{R} \to \mathrm{Gl}_n(\mathbb{R})$ *defined by*

$$t \to \exp(tX)$$

is a one parameter subgroup.

The next theorem shows that all one parameter subgroups arise as the above example.

Theorem 4.6. *Let* $\gamma : \mathbb{R} \to G$ *be a one parameter subgroup. Then* γ *is analytic and*

$$\gamma(t) = \exp(tX),$$

where $X = \gamma'(0)$.

Proof. First suppose that $\gamma : \mathbb{R} \to \mathrm{Gl}_n(\mathbb{R})$ is of class C^1. Then

$$\gamma'(t) = \lim_{s\to 0} \frac{\gamma(t+s)-\gamma(t)}{s} = \lim_{s\to 0} \frac{\gamma(t)\gamma(s)-\gamma(t)}{s} = \lim_{s\to 0} \frac{\gamma(s)-\gamma(0)}{s}\gamma(t) = \gamma'(0)\gamma(t).$$

By the existence and unicity theorem 3.2 we obtain that $\gamma(t) = \exp(tX)$, where $X = \gamma'(0)$.

It remains to show that γ is of class C^1. For that sake we show that its convolution with a convenient smooth function is equal to γ multiplied by an invertible matrix. As the convolution is smooth, γ itself will be smooth.

Let α be a real smooth function of compact support and define the convolution (which is a smooth function with values in $\mathrm{M}_n(\mathbb{R})$)

$$f(t) = \int_{\mathbb{R}} \alpha(t-s)\gamma(s)ds.$$

We have

$$f(t) = \int_{\mathbb{R}} \alpha(s)\gamma(t-s)ds = \int_{\mathbb{R}} \alpha(s)\gamma(-s)ds\gamma(t).$$

Therefore $f(t) = B\gamma(t)$ where B is a fixed matrix. One can choose α so that B be an invertible matrix. Indeed, choose $\alpha \geq 0$ with compact support sufficiently close to the origin such that $\int_{\mathbb{R}} \alpha(s)ds = 1$. Then $\int_{\mathbb{R}} \alpha(s)\gamma(-s)ds$ is close to the identity and therefore invertible. Indeed,

$$|| \int_{\mathbb{R}} \alpha(s)\gamma(-s)ds - I|| \leq \int_{\mathbb{R}} \alpha(s)||\gamma(-s) - I||$$

and choosing the support of α conveniently we obtain that $\int_{\mathbb{R}} \alpha(s)\gamma(-s)ds$ is invertible. □

4.3 The Lie algebra of a linear Lie group

Definition 4.7. *A Lie algebra over a field K is a vector space $\mathfrak{g}$ over K equipped with a bilinear map (called Lie bracket)*

$$[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g},$$

such that

1. $[X,Y] = -[X,Y]$ *(skewsymmetry),*
2. $[X,[Y,Z]] + [Z,[X,Y]] + [Y,[Z,X]] = 0$ *(Jacobi identity).*

A Lie sub-algebra of $\mathfrak{g}$ is a vector subspace $\mathfrak{h}$ such that, for all $X, Y \in \mathfrak{h}$, $[X,Y] \in \mathfrak{h}$.

Example 4.8. *The vector space* $\mathrm{M}_n(K)$ *is a Lie algebra with the Lie bracket defined for all* $X, Y \in \mathrm{M}_n(K)$ *by* $[X, Y] = XY - YX$.

To any Lie group one can associate a Lie algebra. It is the tangent space at the identity element of the group. We will first give the definition in the case of linear Lie groups which does not involve the knowledge of the tangent space to a manifold.

Definition 4.9. *Let* $G \subset \mathrm{Gl}_n(\mathbb{R})$ *be a linear group. Define*

$$\mathrm{Lie}(G) = \{\, X \in \mathrm{M}_n(\mathbb{R}) \mid \exp(tX) \in G \textit{ for all } t \in \mathbb{R} \,\}.$$

Proposition 4.10. *Let* $G \subset \mathrm{Gl}_n(\mathbb{R})$ *be a linear group. Then* $\mathrm{Lie}(G) \subset \mathrm{M}_n(\mathbb{R})$ *is a Lie subalgebra.*

Proof.

1. $\mathrm{Lie}(G)$ is a vector space: It is clear that $X \in \mathrm{Lie}(G)$ implies that $tX \in \mathrm{Lie}(G)$ for all $t \in \mathbb{R}$. Suppose now that $X, Y \in \mathrm{Lie}(G)$. Then, for all $t \in \mathbb{R}$,

$$\lim_{k\to\infty} \left(\exp(\frac{t}{k}X) \exp(\frac{t}{k}Y) \right)^k = \exp t\,(X+Y)\,.$$

As the group G is closed, this implies that $X + Y \in \mathrm{Lie}(G)$.

2. Suppose again that $X, Y \in \mathrm{Lie}(G)$. Then $[X, Y] \in \mathrm{Lie}(G)$. Indeed, for all $t \in \mathbb{R}$,

$$\lim_{k\to\infty} \left(\exp(\frac{t}{k}X) \exp(\frac{t}{k}Y) \exp(-\frac{t}{k}X) \exp(-\frac{t}{k}Y) \right)^{k^2} = \exp\,(t[X,Y]) \in G.$$

□

Remark 4.11. *One can verify that the following diagram commutes.*

$$\begin{array}{ccc} \mathrm{M}_n(\mathbb{C}) & \xrightarrow{\ \iota\ } & \mathrm{M}_{2n}(\mathbb{R}) \\ \downarrow{\scriptstyle \exp} & & \downarrow{\scriptstyle \exp} \\ \mathrm{GL}_n(\mathbb{C}) & \xrightarrow{\ \iota\ } & \mathrm{GL}_{2n}(\mathbb{R}) \end{array}$$

This allows considering $\mathrm{Lie}(G)$ *of a linear Lie group defined as a subgroup in* $\mathrm{GL}_n(\mathbb{C})$ *in* $\mathrm{M}_n(\mathbb{C})$.

Proposition 4.12. *Here are some examples of Lie algebras of linear Lie groups*

1. *If* $\Gamma \subset \mathrm{GL}_n(\mathbb{R})$ *is a discrete subgroup, then* $\mathrm{Lie}(\Gamma) = \{0\}$.

2. $\mathrm{Lie}(\mathrm{SL}(n,\mathbb{R})) = \{\, X \in \mathrm{M}_n(\mathbb{R}) \mid \mathrm{tr}\, X = 0 \,\}$,

3. $\mathrm{Lie}(\mathrm{SO}(n,\mathbb{R})) = \left\{\, X \in \mathrm{M}_n(\mathbb{R}) \mid X^T = -X \,\right\}$,

4. $\mathrm{Lie}(\mathrm{Sp}(n,\mathbb{R})) = \left\{ \begin{pmatrix} A & B \\ C & -A^T \end{pmatrix} \in \mathrm{M}_{2n}(\mathbb{R}) \mid A \in \mathrm{M}_n(\mathbb{R}),\, B, C \in Sym_n \right\}$,

5. $\mathrm{Lie}(\mathrm{U}(n)) = \left\{ X \in \mathrm{M}_n(\mathbb{C}) \mid \overline{X}^T = -X \right\}$.

Proof. We prove the first two assertions and leave to the reader the other two.

1. From Jacobi identity $\det(\exp(tX)) = e^{t\,\mathrm{tr}\, X}$. Therefore, $\mathrm{tr}\, X = 0$ if and only if $\det(\exp(tX)) = 1$ for all t.

2. From the formula $(\exp tX)^T = \exp(tX^T)$ we obtain that $X \in \mathrm{Lie}(\mathrm{SO}(n,\mathbb{R}))$ if and only if $\exp(tX^T) = \exp(-tX)$ for all $t \in \mathbb{R}$. As the exponential is a diffeomorphism at the origin we obtain that this is equivalent to $X = -X^T$.

□

Example 4.13 (SL(2, ℝ) in more detail)**.** *A standard basis of* $\mathrm{Lie}(\mathrm{SL}(2,\mathbb{R}))$ *is the following:*

$$E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Note that $[E,F] = H$, $[H,E] = 2E$ *and* $[H,F] = -2F$.

Example 4.14 (Affine transformations of $\mathbb{R}^n$)**.** *Consider the group* $\mathrm{Aff}(n)$ *defined as a closed subgroup of* $\mathrm{GL}(n+1,\mathbb{R})$*:*

$$\mathrm{Aff}(n) = \left\{ \begin{pmatrix} A & v \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}(n+1,\mathbb{R}) \mid A \in \mathrm{Gl}_n(\mathbb{R}),\ v \in \mathbb{R}^n \right\}.$$

The affine group is a semi-direct product $\mathbb{R}^n \rtimes \mathrm{Gl}_n(\mathbb{R})$, *meaning that there exits an exact sequence*

$$\{0\} \to \mathbb{R}^n \to \mathrm{Aff}(n) \to \mathrm{Gl}_n(\mathbb{R}) \to \{e\}$$

which splits (there exists a homomorphism $h : \mathrm{Gl}_n(\mathbb{R}) \to \mathrm{Aff}(n)$ *which, composed with the last surjection is the identity).*

The action $\mathrm{Aff}(n) \times \mathbb{R}^n \to \mathbb{R}^n$ *is given by*

$$\left(\begin{pmatrix} A & v \\ 0 & 1 \end{pmatrix}, x \right) = \begin{pmatrix} A & v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = Ax + v.$$

The computation of the Lie algebra gives

$$\mathrm{Lie}(\mathrm{Aff}(n)) = \left\{ \begin{pmatrix} M & n \\ 0 & 0 \end{pmatrix} \in \mathrm{M}_{n+1}(\mathbb{R}) \mid M \in \mathrm{M}_n(\mathbb{R}),\ n \in \mathbb{R}^n \right\}.$$

4.4 The adjoint maps

Definition 4.15. *1. A morphism $A : \mathfrak{g} \to \mathfrak{h}$ between Lie algebras is a linear map such that, for all $X, Y \in \mathfrak{g}$,*

$$[AX, AY]_{\mathfrak{h}} = [X, Y]_{\mathfrak{g}},$$

where $[\cdot,\cdot]_{\mathfrak{g}}$ and $[\cdot,\cdot]_{\mathfrak{h}}$ are the Lie brackets of $\mathfrak{g}$ and $\mathfrak{h}$ respectively. We define $\mathrm{Aut}(\mathfrak{g}) \subset \mathrm{GL}(\mathfrak{g})$ to be the group of automorphisms (bijective morphisms) of $\mathfrak{g}$.

2. *A derivation $\mathrm{D} : \mathfrak{g} \to \mathfrak{g}$ of a Lie algebra $\mathfrak{g}$ is a linear map satisfying*

$$\mathrm{D}([X, Y]) = [\mathrm{D}(X), Y] + [X, \mathrm{D}(Y)].$$

We define $\mathrm{Der}(\mathfrak{g}) \subset \mathrm{End}(\mathfrak{g})$ to be the derivations of $\mathfrak{g}$.

Example 4.16. *Let G be a linear Lie group, $\mathfrak{g}$ its Lie algebra and $g \in G$ a fixed element. The map*

$$\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g},$$

defined by $\mathrm{Ad}_g(X) = g^{-1}Xg$ is an automorphism of $\mathfrak{g}$. One easily verifies that for any $h, h \in G$ and $X \in \mathfrak{g}$,

$$\mathrm{Ad}_{gh} X = \mathrm{Ad}_g \circ \mathrm{Ad}_h(X).$$

The map $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ defined by $\mathrm{ad}_X(Y) = [X, Y]$ is a derivation. Observe that for all $X, Y, Z \in \mathfrak{g}$

$$ad_{[X,Y]}(Z) = [\mathrm{ad}_X, \mathrm{ad}_Y](Z).$$

This example justifies the following:

Definition 4.17. *Let G be a linear Lie group and $\mathfrak{g}$ be its Lie algebra.*

1. *The adjoint representation of the group is the homomorphism*

$$\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g}),$$

given by $\mathrm{Ad}(g) = \mathrm{Ad}_g$.

2. *The adjoint representation of the Lie algebra is the homomorphism*

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

given by $\mathrm{ad}(X) = \mathrm{ad}_X$.

We refer sometimes to Ad_g as the *adjoint map* defined by g and sometimes to Ad as the adjoint map. We also say that the derivation ad_X is the adjoint map (yet another adjoint!) defined by X.

A first relation between those maps is given by the following:

Proposition 4.18. *Let G be a linear Lie group, $\mathfrak{g}$ be its Lie algebra and $X \in \mathfrak{g}$. Then*

$$\frac{d}{dt}\mathrm{Ad}_{\exp(tX)}\big|_{t=0} = \mathrm{ad}_X .$$

Remark 4.19. *Here, $\mathrm{Ad}_{\exp(tX)} \in \mathrm{GL}(\mathfrak{g})$ for all $t \in \mathbb{R}$. The derivative at $t = 0$ is, therefore, an element in $\mathrm{End}(\mathfrak{g})$.*

Proof. We write, for $Y \in \mathfrak{g}$,

$$\frac{d}{dt}\mathrm{Ad}_{\exp(tX)}\big|_{t=0}(Y) = \frac{d}{dt}\exp(tX)\,Y\exp(-tX)\big|_{t=0} = XY - YX.$$

□

Recall from 3.5 that we define $\mathrm{Exp} : \mathrm{End}(\mathfrak{g}) \to \mathrm{GL}(\mathfrak{g})$ by a power series, as $\exp : \mathrm{M}_n(\mathbb{R}) \to \mathrm{Gl}_n(\mathbb{R})$, applied to linear maps instead of matrices.

Proposition 4.20. *For all $X \in \mathfrak{g}$,*

$$\mathrm{Exp}(\mathrm{ad}_X) = \mathrm{Ad}_{\exp X} .$$

Remark 4.21. *That is better seen with the commuting diagram*

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\ \mathrm{ad}\ } & \mathrm{End}(\mathfrak{g}) \\ \Big\downarrow{\scriptstyle \exp} & & \Big\downarrow{\scriptstyle \mathrm{Exp}} \\ G & \xrightarrow{\ \mathrm{Ad}\ } & \mathrm{GL}(\mathfrak{g}) \end{array}$$

Proof. Both sides of the equality are one parameter groups with values in $\mathrm{GL}(\mathfrak{g})$. It suffices to compute the derivatives at the origin and show they are equal:

$$\frac{d}{dt}\mathrm{Exp}(t\,\mathrm{ad}_X)\big|_{t=0} = \mathrm{ad}_X \circ \mathrm{Exp}(t\,\mathrm{ad}_X)\big|_{t=0} = \mathrm{ad}_X$$

and, on the other hand, from the previous proposition,

$$\frac{d}{dt}\mathrm{Ad}_{\exp tX} = \mathrm{ad}_X .$$

□

4.4.1 The differential of the exponential map

We may now prove theorem 3.12. That is, the formula

$$D\exp_A(H) = \exp(A)\sum_{k=0}^{\infty}\frac{(-1)^k}{(k+1)!}(\mathrm{ad}_A)^k H = \frac{1-\mathrm{Exp}(-\mathrm{ad}_A)}{\mathrm{ad}_A}H.$$

Proof. We need to compute for $H = X'(0)$ where $X(t)$ is a path in $\mathrm{M}_n(\mathbb{R})$ such that $A = X(0)$,

$$\exp(-A)\frac{d}{dt}\exp(X(t))_{t=0}$$

Introduce the function in two variables $\Gamma(s,t) = \exp(-sX(t))\frac{d}{dt}\exp(sX(t))$. We have

$$\exp(-A)\frac{d}{dt}\exp(X(t))_{t=0} = \Gamma(1,0) = \int_0^1 \frac{\partial\Gamma(s,0)}{\partial s}ds.$$

Now we compute

$$\begin{aligned}\frac{\partial\Gamma(s,t)}{\partial s} &= \exp(-sX(t))(-X(t))\frac{d}{dt}\exp(sX(t)) + \exp(-sX(t))\frac{d}{dt}(X(t)\exp(sX(t)))\\ &= \exp(-sX(t))X'(t)\exp(sX(t)) = Ad_{\exp(-sX(t))}X'(t) = \mathrm{Exp}\left(-\mathrm{ad}_{sX(t)}(X'(t))\right)\end{aligned}$$

We conclude with the following:

$$\begin{aligned}\int_0^1 \frac{\partial\Gamma(s,0)}{\partial s}ds &= \int_0^1 \mathrm{Exp}(-\mathrm{ad}_{sA}(H))ds = \int_0^1 \sum_{k=0}^{\infty}\frac{s^k}{k!}(-\mathrm{ad}_A)^k H ds\\ &= \sum_{k=0}^{\infty}\frac{1}{(k+1)!}(-\mathrm{ad}_A)^k H = \frac{1-\mathrm{Exp}(-\mathrm{ad}_A)}{\mathrm{ad}_A}H.\end{aligned}$$

□

4.5 A linear Lie group is a submanifold of $\mathrm{Gl}_n(\mathbb{R})$

We first give the following definition of a submanifold of dimension m in $\mathbb{R}^N$:

Definition 4.22. *We say a subset $G \subset \mathbb{R}^N$ is an embedded submanifold of dimension m if for every $x \in G$ there exists an open subset $O \subset \mathbb{R}^N$ containing the origin, a diffeomorphism $\Phi : O \to W$ where $W \subset \mathbb{R}^N$ such that $\Phi(0) = x$ and a vector subspace $V \subset \mathbb{R}^N$ of dimension m such that*

$$\Phi(O \cap V) = G \cap W.$$

The tangent space at $x \in G$ is

$$T_xG = D(\Phi)_0(V).$$

Example 4.23. *$SO(2)\times SO(2)$ can be seen as a closed topological subgroup of $SO(4,\mathbb{R})\subset$ $\mathrm{GL}(4,\mathbb{R})$. Writing $SO(2)$ as the complex numbers of absolute value one we may describe one parameter subgroups $\gamma_{a,b}:\mathbb{R}\to SO(2)\times SO(2)$ given by*

$$t\to(e^{iat},e^{ibt})$$

where $a,b\in\mathbb{R}$. If $\frac{a}{b}\notin\mathbf{Q}$ then $\gamma_{a,b}(\mathbb{R})$ is not an embedded submanifold.

The following theorem is the main step to prove that a linear Lie group is an embedded submanifold. For a convenient map Φ, It says that the condition of definition 4.22 is satisfied for $x = Id$.

Theorem 4.24. *Let $G\subset \mathrm{Gl}_n(\mathbb{R})$ be a linear Lie subgroup and $\mathfrak{g}=\mathrm{Lie}(G)$ be its Lie algebra. Choose $\mathfrak{h}\subset \mathrm{M}_n(\mathbb{R})$ such that $\mathrm{M}_n(\mathbb{R})=\mathfrak{g}\oplus\mathfrak{h}$. Then, there exists a neighborhood O of the origin in $\mathrm{M}_n(\mathbb{R})$ such that the restriction of the map $\Phi:\mathfrak{g}\oplus\mathfrak{h}\to\mathrm{Gl}_n(\mathbb{R})$ given by $X+Y\to\exp(X)\exp(Y)$ is a diffeomorphism onto its image and satisfies*

$$\Phi(O\cap\mathfrak{g})=G\cap\Phi(O).$$

Proof. In order to prove that the map Φ is a local diffeomorphism we compute its differential:

$$D\Phi_0(X+Y)=X+Y.$$

That is, $D\Phi_{0)}=Id$ which is an isomorphism and, therefore, there exists a neighborhood of 0, $O\subset\mathfrak{g}\oplus\mathfrak{h}$, such that $\Phi_{|_O}$ is a diffeomorphism onto its image.

We want next to prove that $\Phi(O\cap\mathfrak{g})=G\cap\Phi(O)$. We need two lemmas in order to finish the proof:

Lemma 4.25. *Let G be a linear Lie group and consider a sequence $g_k\in G$ converging to the identity. Let $X_k\in \mathrm{M}_n(\mathbb{R})$ converging to 0 such that $g_k=\exp(X_k)$. Then the accumulation points of the sequence $\frac{X_k}{|X_k|}$ are in $\mathrm{Lie}(G)$.*

Proof. Note that $X_k\to 0$. Suppose

$$\frac{X_k}{|X_k|}\to Y\in\mathrm{M}_n(\mathbb{R}).$$

We have to show that $\exp(tY)\in G$ for all $t\in\mathbb{R}$. Observe that

$$\exp\left(\frac{t}{|X_k|}X_k\right)=\exp\left(\left(\frac{t}{|X_k|}-\left[\frac{t}{|X_k|}\right]\right)X_k\right)\exp\left(\left[\frac{t}{|X_k|}\right]X_k\right),$$

where $[s]$ is the integer part of $s \in \mathbb{R}$. For a fixed t we have that, as $X_k \to 0$,

$$\lim_{k\to\infty} \exp\left(\left(\frac{t}{|X_k|} - \left[\frac{t}{|X_k|}\right]\right) X_k\right) = Id.$$

The second factor is a power of an element of G for all k. Therefore it converges, because G is closed, to an element in G. We conclude that $\exp(tY) \in G$. □

Lemma 4.26. *Let* $\mathfrak{h}$ *be a subspace of* $\mathrm{M}_n(\mathbb{R})$ *such that* $\mathrm{M}_n(\mathbb{R}) = \mathfrak{g} \oplus \mathfrak{h}$. *Then there exist a neighborhood* V *of the origin in* $\mathfrak{h}$ *such that* $\exp(V) \cap G = \{Id\}$.

Proof. Suppose, by contradiction, that there exists $X_k \in \mathfrak{h}$ such that $X_k \to 0$ and $\exp(X_k) = g_k \in G$. By the previous lemma, any convergent subsequence $\frac{X_k}{|X_k|} \to Y$ converges to an element in $\mathfrak{g}$ (there is always a convergent subsequence by compactness of the set of vectors of norm one). Therefore $Y \in \mathfrak{g} \cap \mathfrak{h} = \{0\}$. A contradiction because $||Y|| = 1$.

□

Now we return to the proof of the theorem: It remains to prove that $\Phi(O \cap \mathfrak{g}) = G \cap \Phi(O)$. Clearly $\Phi(O \cap \mathfrak{g}) \subset G \cap \Phi(O)$. In order to prove the other inclusion, suppose, by contradiction, that there exists $g \in G \cap \Phi(O)$ such that $g = \exp(X)\exp(Y)$ and $Y \neq 0$. Then $\exp(-X)g = \exp(Y)$, a contradiction with the previous lemma choosing O of the form

$$\{\, X + Y \in \mathfrak{g} \oplus \mathfrak{h} \mid X \in U,\ Y \in V \,\}.$$

□

Corollary 4.27. *A linear Lie group is an embedded submanifold of* $\mathrm{Gl}_n(\mathbb{R})$. *Its Lie algebra is identified to the tangent space at the identity.*

Proof. We use the neighborhood O constructed in the previous theorem: there exists a neighborhood $0 \in O \in \mathrm{M}_n(\mathbb{R})$ and a diffeomorphism $\Phi : O \to W$ satisfying

$$\Phi(O \cap \mathfrak{g}) = \Phi(O) \cap G.$$

Given $g \in G$, we compose Φ with the translation L_g to obtain a diffeomorphism $L_g \circ \Phi : O \to L_g(W)$. Then

$$L_g \circ \Phi(O \cap \mathfrak{g}) = L_g \circ \Phi(O) \cap G.$$

The tangent space to G at an element $Id \in G$ is given by

$$D_0(\Phi)(\mathfrak{g}) = Id(\mathfrak{g}) = \mathfrak{g}.$$

□

4.6 homomorphisms of Lie groups and homomorphisms of Lie algebras

In this section we make explicit the relation between homomorphisms of a Lie algebra of a linear group and homomorphisms of the Lie group. But note that a homomorphism between Lie groups is not necessarily continuous (you may try to find one from $\mathbb{R}$ to $\mathbb{R}$). For the next proposition we will impose that the homomorphism is smooth in order to relate it to a morphism between their Lie algebras.

First we clarify the notion of smoothness. Let $G \subset \mathrm{Gl}_n(\mathbb{R})$ be a linear Lie subgroup and $\mathfrak{g} = \mathrm{Lie}(G)$ be its Lie algebra. We proved in theorem 4.28 that there exists a neighbourhood O of the origin in $\mathfrak{g}$ such that the restriction of the exponential map to its Lie algebra $\exp_{\mathfrak{g}} : \mathfrak{g} \to G$ is a local homeomorphism. We use this local homeomorphism to define smoothness of a homomorphism $\phi : G \to H$:

Definition 4.28. *Let $\phi : G \to H$ be a homomorphism between linear Lie groups. We say it is a smooth homomorphism if for each $g \in G$ there exits a neighbourhood of the origin $U \subset \mathfrak{g}$ such that the map*

$$\exp_{\mathfrak{h}}^{-1} \circ L_{(\phi(g))^{-1}} \circ \phi \circ L_g \circ \exp_{\mathfrak{g}} : U \to \mathfrak{h}$$

is smooth.

Remark 4.29. *At the identity ($g = e$), the map is $\exp_{\mathfrak{h}}^{-1} \circ \phi \circ \exp_{\mathfrak{g}} : U \to \mathfrak{h}$. The differential $D_e\phi$ is, by definition, the differential of that map at 0.*

Proposition 4.30. *Let $\phi : G \to H$ be a homomorphism (which we suppose smooth) between linear Lie groups. We define the homomorphism $D\phi_e : \mathfrak{g} \to \mathfrak{h}$ as the differential of the map at the identity. Then*

1. $\ker D\phi_e = \mathrm{Lie}(\ker \phi)$,
2. *$D\phi_e$ is injective if and only if $\ker \phi$ is discrete,*
3. *If $D\phi_e$ is surjective then $\phi(G)$ contains the connected component of the identity in H,*
4. *If G and H are connected then $D\phi_e$ is an isomorphism if and only if ϕ is a covering map.*

Proof. We verify that $\exp\left(tD\phi_e(X)\right)$ and $\phi(\exp tX)$ are two one parameter groups with the same derivative at $t = 0$. Indeed,

$$\frac{d}{dt} \exp\left(tD\phi_e(X)\right)_{|_{t=0}} = D\phi_e(X) = \frac{d}{dt}\phi(\exp tX)_{|_{t=0}}.$$

From theorem 4.6 we have, for all $X \in \mathfrak{g}$ and $t \in \mathbb{R}$,

$$\exp\left(tD\phi_e(X)\right) = \phi(\exp tX).$$

Therefore $X \in \ker D\phi_e$ if and only if $\exp tX \in \ker\phi$ for all t if and only if $X \in \mathrm{Lie}(\ker\phi)$. This proves the first assertion and the second follows immediately from the formula and the fact that the Lie algebra of a Lie group is $\{0\}$ if and only if the group is discrete.

If $D\phi_e$ is surjective then the image of ϕ contains the image, by the exponential map, of a neighbourhood of the origin in $\mathfrak{h}$. This implies that ϕ is surjective onto a neighbourhood of the identity in H. We conclude that $\phi(G)$ contains the connected component of the identity in H by the proposition .

For the last assertion: recall that a map $\phi : G \to H$ is a covering map if it is a local homeomorphism and for any $h \in H$ there exists an open neighbourhood $h \in U$ such that $\phi^{-1}(U)$ is a disjoint union of neighbourhoods which are homeomorphic to U via the map ϕ.

If ϕ is a covering map between Lie groups then, in particular, the Lie algebras are of the same dimension and $\dim(\ker\phi) = 0$ which implies that $D\phi_e$ is an isomorphism. On the other hand, suppose $D\phi_e$ is an isomorphism. We first verify the covering condition at the identity: we may choose an open neighbourhood $e \in V$ of the identity such that $V \cap \ker\phi = \{e\}$. Let U be a neighborhood of the identity in G such that $\phi(V) = U$. We may choose V to be symmetric (satisfying $V^{-1} = V$). Now we claim that

$$\phi^{-1}(U) = \bigcup_{n\in\ker\phi} nV$$

is a disjoint union of neighbourhoods homeomorphic to U via ϕ. Indeed, if $nv = n'v'$ with $v, v' \in V$ then $vv'^{-1} = n'n^{-1} \in \ker\phi$. That is $vv'^{-1} = e$ which implies $n = n'$. In general, if $h \in H$, choose $h' \in \phi^{-1}(h)$, then the neighbourhood hU satisfies

$$\phi^{-1}(hU) = \bigcup_{n\in\ker\phi} h'nV$$

which is a disjoint union of open sets homeomorphic to hU via ϕ and the proof is complete.

□

Example 4.31. *Consider* $\mathrm{SU}(2) \subset \mathrm{GL}_2(\mathbb{C})$ *the special unitary group. Its Lie algebra is*

$$\mathfrak{su}(2) = \{\ X \in \mathrm{M}_2(\mathbb{C}) \mid X + \overline{X}^T = 0\ \}.$$

A general element in $\mathfrak{su}(2)$ *is of the form*

$$X = \begin{pmatrix} iu & v + iw \\ -v + iw & -iu \end{pmatrix}.$$

Observe that the $\det X = u^2 + v^2 + w^2$ *is the standard euclidean quadratic form on* $\mathbb{R}^3$.
Define the adjoint map $\mathrm{Ad} : \mathrm{SU}(2) \to \mathrm{Aut}(\mathfrak{su}(2))$. *and observe that* $\det(\mathrm{Ad}(g)X) = \det(X)$.
This implies that, identifying $\mathfrak{su}(2)$ *to* $\mathbb{R}^3$ *with the euclidean metric as above,* $\mathrm{Ad}(g) \subset \mathrm{O}(3)$.
In fact the image is in $\mathrm{SO}(3)$ *because* $\mathrm{SU}(2)$ *is connected.*

Exercise: show that the map $\mathrm{Ad} : \mathrm{SU}(2) \to \mathrm{SO}(3)$ *is a covering map with* $\ker \mathrm{Ad} = \{\pm Id\}$.

5 Lie algebras

In this chapter we study the basic results on Lie algebras, including the theorems of Lie and Engels.

5.1 Derivations and automorphisms

The set of linear transformations defined on a vector spaces V is denoted $\mathrm{End}(V)$ and it is equipped with a Lie algebra structure. Indeed, for $T, S \in \mathrm{End}(V)$, define $[T,S] = T \circ S - S \circ T$.

Recall from definition that a derivation $D : \mathfrak{g} \to \mathfrak{g}$ of a Lie algebra $\mathfrak{g}$ is a linear map satisfying

$$D([X,Y]) = [D(X),Y] + [X,D(Y)].$$

Proposition 5.1. *Let $\mathfrak{g}$ be a Lie algebra.*

1. *The set of derivations* $\mathrm{Der}(\mathfrak{g}) \subset \mathrm{End}(\mathfrak{g})$ *is a Lie sub-algebra of* $\mathrm{End}(\mathfrak{g})$.
2. *The map* $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ *is a morphism of Lie algebras.*
3. $\mathrm{ad}(\mathfrak{g}) \subset \mathrm{Der}(\mathfrak{g})$.

Proof. 1. If A and B are two derivations, then for $X, Y \in \mathfrak{g}$,

$$[A,B]([X,Y]) = (AB-BA)[X,Y] = A([BX,Y]+[X,BY]) - B([AX,Y]+[X,AY])$$

$$= [ABX,Y]+[BX,AY]+[AX,BY]+[X,ABY]-([BAX,Y]+[AX,BY]+[BX,AY]+[X,BAY])$$

$$= [(AB-BA)X,Y]+[X,(AB-BA)Y].$$

2. Indeed,

$$\mathrm{ad}_{[X,Y]}(Z) = [[X,Y],Z] = [[X,Y],Z]$$

and, on the other hand

$$[\mathrm{ad}_X, \mathrm{ad}_Y](Z) = \mathrm{ad}_X \circ \mathrm{ad}_Y(Z) - \mathrm{ad}_Y \circ \mathrm{ad}_X(Z) = [X,[Y,Z]] + [Y,[X,Z]].$$

The equality of both sides follows then from Jacobi identity.

3. we leave to the reader the last inclusion.

□

Proposition 5.2. *Let* $\mathfrak{g}$ *be a finite dimensional Lie algebra. Then*

$$\mathrm{Der}(\mathfrak{g}) = \mathrm{Lie}\,(\mathrm{Aut}(\mathfrak{g}))\,.$$

Proof. If $\mathrm{D} \in \mathrm{Lie}\,(\mathrm{Aut}(\mathfrak{g}))$ then $\mathrm{Exp}(t\mathrm{D}) \in \mathrm{Aut}(\mathfrak{g})$ for all $t \in \mathbb{R}$. That is, $\mathrm{Exp}(t\mathrm{D})([X,Y]) = [\mathrm{Exp}(t\mathrm{D})X, \mathrm{Exp}(t\mathrm{D})Y])$ for all $X, Y \in \mathfrak{g}$. Differentiating at $t = 0$, we obtain

$$\mathrm{D}([X,Y]) = [\mathrm{D}\,X, Y] + [X, \mathrm{D}\,Y].$$

On the other direction, if D is a derivation, define two functions $F_i : \mathbb{R} \to \mathfrak{g}$, $1 \le i \le 2$, by

$$F_1(t) = \mathrm{Exp}(t\mathrm{D})[X,Y], \quad F_1(t) = [\mathrm{Exp}(t\mathrm{D})X, \mathrm{Exp}(t\mathrm{D})Y].$$

We show that they satisfy the same first differential equation with initial condition $F_1(0) = F_2(0) = [X,Y]$:

$$F_1'(t) = \mathrm{D}[X,Y] = [\mathrm{D}\,X, Y] + [X, \mathrm{D}\,Y] = F_2'(t).$$

We conclude $F_1(t) = F_2(t)$ for all t and therefore $\mathrm{Exp}(t\mathrm{D}) \in \mathrm{Aut}(\mathfrak{g})$.

□

Remark 5.3. *In view of the proposition, we may write the diagram (compare with remark 4.21)*

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\mathrm{ad}} & \mathrm{Der}(\mathfrak{g}) \\ \downarrow{\scriptstyle \exp} & & \downarrow{\scriptstyle \mathrm{Exp}} \\ G & \xrightarrow{\mathrm{Ad}} & \mathrm{Aut}(\mathfrak{g}) \end{array}$$

5.2 Ideals

Definition 5.4. *Let* $\mathfrak{g}$ *be a Lie algebra. A subalgebra* $\mathfrak{h} \subset \mathfrak{g}$ *is said to be an ideal if, for any* $X \in \mathfrak{g}$ *and* $Y \in \mathfrak{h}$, $[X,Y] \in \mathfrak{h}$.

Example 5.5. *Let* $\mathfrak{g}$ *be a Lie algebra. Then its center*

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

and its derived algebra

$$\mathscr{D}(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}]$$

are ideals.

Ideals are an infinitesimal version of normal subgroups as the following proposition shows.

Proposition 5.6. *Let G be a linear Lie group and Let $H \subset G$ be a normal closed subgroup. Then $\mathfrak{h} = \mathrm{Lie}(H)$ is an ideal of $\mathfrak{g} = \mathrm{Lie}(G)$.*

Proof. Let $Y \in \mathfrak{h}$ and $X \in \mathfrak{g}$. Then, for all $t, s \in \mathbb{R}$, $\exp(tX)\exp(sY)\exp(-tX) \in H$. Therefore, for all $t \in \mathbb{R}$,

$$\frac{d}{ds}_{|0} \exp(tX)\exp(sY)\exp(-tX) = \exp(tX)\,Y\exp(-tX) \in \mathfrak{h}.$$

Differentiating again, we obtain that

$$\frac{d}{dt}_{|0} \exp(tX)\,Y\exp(-tX) = \mathrm{ad}_X(Y) = [X, Y] \in \mathfrak{h}.$$

□

Definition 5.7. *A Lie algebra is simple if it is not abelian and the only ideals are $\{0\}$ and itself.*

Example 5.8. *The Lie algebra $\mathfrak{sl}(2,\mathbb{R})$ is simple. Indeed, suppose $\mathfrak{h} \subset \mathfrak{sl}(2,\mathbb{R})$ is an ideal. We recall the basis E, F, H of the Lie algebra satisfying $[E, F] = H$, $[H, E] = 2E$ and $[H, F] = -2F$. Writing a nonzero element of $\mathfrak{h}$ as $X = aE + bF + cH$ we obtain*

$$\mathrm{ad}_E \circ \mathrm{ad}_E(X) = \mathrm{ad}_E(bH - 2cE) = -2bE, \quad \mathrm{ad}_F \circ \mathrm{ad}_F(X) = \mathrm{ad}_F(-aH + 2cF) = -2aF.$$

Therefore, if a or b are nonzero we immediately obtain that $\mathfrak{h} = \mathfrak{sl}(2,\mathbb{R})$. On the other hand, if $a = b = 0$ then $c \neq 0$ and this implies again that $\mathfrak{h} = \mathfrak{sl}(2,\mathbb{R})$.

5.3 Representations: first definitions

Definition 5.9. *Let G be a topological group and V a finite dimensional vector space. A representation of G on V is a continuous homomorphism*

$$\pi : G \to \mathrm{GL}(V).$$

We say that a subspace $W \subset V$ is an invariant subspace if $\pi(g)W \subset W$ for all $g \in G$.

Definition 5.10. *Let $\mathfrak{g}$ be a Lie algebra and V a finite dimensional vector space. A representation of $\mathfrak{g}$ on V is a morphism of Lie algebras*

$$\rho : \mathfrak{g} \to \mathrm{End}(V).$$

We say that a subspace $W \subset V$ is an invariant subspace if $\rho(X)W \subset W$ for all $X \in \mathfrak{g}$.

Example 5.11. *The maps* $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$ *and* $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ *are representations of the Lie group G and its Lie algebra* $\mathfrak{g}$.

Proposition 5.12. *Let* $\pi : G \to \mathrm{GL}(V)$ *be a representation of a linear Lie group. Then the map*

$$\rho : \mathfrak{g} \to \mathrm{End}(V)$$

defined by

$$\rho(X) = \frac{d}{dt}_{|t=0} \pi\left(\exp(tX)\right)$$

is a representation of the Lie algebra $\mathfrak{g}$.

Proof. Observe that $\pi\left(\exp(tX)\right)$ is a one parameter group and therefore smooth. The map ρ is then well defined and we have $\exp(t\rho(X)) = \pi(\exp(tX))$. We will suppose, in order to simplify the proof, that π is differentiable. In this case $\rho = d\pi$, the differential of π. It is then clear that it is a linear map.

It remains to prove that it is a morphism of Lie algebras. Recall that $\exp(t\mathrm{Ad}_g Y) = g\exp(tY)g^{-1}$. Therefore $\pi\left(\exp(t\mathrm{Ad}_g Y)\right) = \pi(g)\pi(\exp(tY))\pi(g^{-1})$. Differentiating at $t = 0$ we obtain

$$\rho(\mathrm{Ad}_g Y) = \pi(g)\rho(Y)\pi(g^{-1})$$

Writing now $g = \exp(sX)$ and differentiating the linear map ρ, we obtain on the left side

$$\frac{d}{ds}\rho(\mathrm{Ad}_{\exp(sX)} Y)|_{s=0} = \rho(\frac{d}{ds}\mathrm{Ad}_{\exp(sX)} Y|_{s=0}) == \rho(\mathrm{ad}_X Y) = \rho([X,Y]).$$

On the right side, on the other hand,

$$\rho(X)\rho(Y) - \rho(Y)\rho(X).$$

□

Remark 5.13. *If we don't suppose smoothness of* π *we may still prove linearity of* ρ *considering*

$$\pi\left(\exp(t(X+Y))\right) = \pi\left(\lim_{k\to\infty}\left(\exp(\frac{t}{k}X)\exp(\frac{t}{k}Y)\right)^k\right)$$

$$= \lim_{k\to\infty}\left(\pi\left(\exp(\frac{t}{k}X)\right)\pi\left(\exp(\frac{t}{k}Y)\right)\right)^k = \lim_{k\to\infty}\left(\exp(\frac{t}{k}\rho(X))\exp(\frac{t}{k}\rho(Y))\right)^k = \exp\left(t(\rho(X)+\rho(Y))\right).$$

Although there exists Lie groups which are not linear Lie groups, Lie algebras always admit injective representations:

Theorem 5.14 (Ado's theorem)**.** *Every finite dimensional Lie algebra admits an injective representation on a finite dimensional vector space.*

5.4 Nilpotent and solvable Lie algebras

Let $\mathfrak{g}$ be a (finite dimensional) Lie algebra over $\mathbb{R}$ or $\mathbb{C}$. Define, for any subsets $A, B \in \mathfrak{g}$,

$$[A, B] \subset \mathfrak{g},$$

to be the vector subspace generated by the Lie brackets of A and B.

Definition 5.15. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{C}$ or $\mathbb{R}$.*

- *The derived series is defined by recurrence by $\mathscr{D}^1(\mathfrak{g}) = \mathscr{D}(\mathfrak{g})$ (the derived algebra) and then, for $n \geq 1$,*
$$\mathscr{D}^{n+1}(\mathfrak{g}) = \mathscr{D}\left(\mathscr{D}^n(\mathfrak{g})\right).$$
- *central series is defined by recurrence by $\mathscr{C}^1(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}]$ and then, for $n \geq 1$,*
$$\mathscr{C}^{n+1}(\mathfrak{g}) = \left[\mathscr{C}^n(\mathfrak{g}), \mathfrak{g}\right].$$

Remark 5.16.

1. *An alternative notation is $\mathfrak{g}^{(n)} = \mathscr{D}^n(\mathfrak{g})$ and $\mathfrak{g}^n = \mathscr{C}^n(\mathfrak{g})$.*
2. *Each $\mathscr{D}^n(\mathfrak{g})$ and each $\mathscr{C}^n(\mathfrak{g})$ is an ideal.*
3. *for $n \geq 1$, $\mathscr{D}^n(\mathfrak{g}) \subset \mathscr{C}^n(\mathfrak{g})$.*
4. *The series are decreasing ($\mathscr{C}^{n+1}(\mathfrak{g}) \subset \mathscr{C}^n(\mathfrak{g})$, $\mathscr{D}^{n+1}(\mathfrak{g}) \subset \mathscr{D}^n(\mathfrak{g})$) and therefore they are constant after a certain rank.*

Definition 5.17. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{C}$ or $\mathbb{R}$.*

- *The Lie algebra is said to be solvable if there exists n such that $\mathscr{D}^n(\mathfrak{g}) = \{0\}$.*
- *The Lie algebra is said to be nilpotent if there exists n such that $\mathscr{C}^n(\mathfrak{g}) = \{0\}$.*

Remark 5.18. *Note that if $\mathfrak{g}$ is nilpotent then it is also solvable.*

Example 5.19. *The set $\mathfrak{t}_0(n)$ of strictly upper triangular matrices of size n is a nilpotent Lie algebra. On the other hand, the set of upper triangular matrices $\mathfrak{t}(n)$ is a solvable Lie algebra.*

Proposition 5.20. *Let $\mathfrak{h} \subset \mathfrak{g}$ be an ideal of the Lie algebra $\mathfrak{g}$.*

1. *If $\mathfrak{h}$ and $\mathfrak{g}/\mathfrak{h}$ are solvable then $\mathfrak{g}$ is solvable.*
2. *If $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are two solvable ideals of $\mathfrak{g}$ then $\mathfrak{h}_1 + \mathfrak{h}_2$ is a solvable ideal.*

Proof. 1. There exists n such that $\mathfrak{D}^n(\mathfrak{g}/\mathfrak{h}) = \{0\}$. That is, means that $\mathfrak{D}^n(\mathfrak{g}) \subset \mathfrak{h}$. As there exists m such that $\mathfrak{D}^m(\mathfrak{h}) = \{0\}$, we obtain that

$$\mathfrak{D}^{n+m}(\mathfrak{g}) \subset \mathfrak{D}^m(\mathfrak{h}) = \{0\}.$$

2. We apply the first result to the quotient $(\mathfrak{h}_1 + \mathfrak{h}_2)/\mathfrak{h}_2$ which is solvable by the isomorphism

$$(\mathfrak{h}_1 + \mathfrak{h}_2)/\mathfrak{h}_2 \simeq \mathfrak{h}_1/\mathfrak{h}_1 \cap \mathfrak{h}_2.$$

□

The proposition allows us to define the maximal solvable ideal:

Definition 5.21. *Let* $\mathfrak{g}$ *be a Lie algebra. Define* $\mathrm{rad}(\mathfrak{g})$ *to be the maximal solvable ideal of* $\mathfrak{g}$.

Recall that a nilpotent matrix $N \in \mathrm{M}_n(\mathbb{R})$ (or a nilpotent transformation $T : V \to V$ defined on a vector space) satisfies $N^n = 0$ (or $T^n = 0$) for a $n \in \mathbf{N}$.

Lemma 5.22. *If an element* $X \in \mathrm{End}(V)$ *is nilpotent then* $\mathrm{ad}_X \in \mathrm{End}(\mathrm{End}(V))$ *is nilpotent.*

Proof. If X is nilpotent then $X^N = 0$ for a $N \in \mathbf{N}$. We compute, for all $A \in \mathrm{End}(V)$,

$$(\mathrm{ad}_X)^n(A) = (L_X - R_X)^n(A) = \sum_k^n \binom{n}{k} L_X^k R_X^{n-k} A,$$

where $L_X A = XA$ and $R_X A = AX$. Therefore, as $X^N = 0$ we obtain that each monomial is null if, for all $0 \le i \le n$, $i \ge N$ or $n - k \ge N$, that is, $n \ge N + k$. We conclude that $(\mathrm{ad}_X)^n$ for $n \ge 2N - 1$.

□

5.4.1 Nilpotent algebras: Engel's theorem

A Lie algebra $\mathfrak{g}$ is nilpotent if there exists N such that for any $Y \in \mathfrak{g}$ and $X_1, \cdots, X_N \in \mathfrak{g}$ we have

$$\mathrm{ad}_{X_1} \circ \cdots \circ \mathrm{ad}_{X_N} Y = 0.$$

In particular, for all $X \in \mathfrak{g}$, $(\mathrm{ad}_X)^N = 0$. Engel's theorem is the converse statement:

Theorem 5.23. *Let* $\mathfrak{g}$ *be a Lie algebra such that, for all* $X \in \mathfrak{g}$, (ad_X) *is nilpotent, that is, for any* $X \in \mathfrak{g}$ *there exists N such* $(\mathrm{ad}_X)^N = 0$. *Then* $\mathfrak{g}$ *is nilpotent.*

The proof is based on the following theorem:

Theorem 5.24. *Let $\rho : \mathfrak{g} \to \mathrm{End}(V)$ be a representation such that each element in $\rho(\mathfrak{g})$ is nilpotent. Then there exists $v \in V \setminus \{0\}$ such that $\rho(X)v = 0$ for all $X \in \mathfrak{g}$.*

Proof. It is sufficient to prove the theorem for an injective representation. We suppose then that $\mathfrak{g} \subset \mathrm{End}(V)$ and $\rho = Id$. The proof is an induction on the dimension of $\mathfrak{g}$. If $\dim(\mathfrak{g}) = 0$ the result is obvious. If $\dim(\mathfrak{g}) = 1$ then $\mathfrak{g} = \langle X \rangle$ and, as X is nilpotent, consider the minimum N such that $X^N = 0$. Then there exists $w \in V \setminus \{0\}$ such that $v = X^{N-1}w \neq 0$ and it follows that $Xv = 0$.

Suppose the proposition true for $\dim(\mathfrak{g}) = n-1$. Consider a maximal subalgebra $\mathfrak{h} \subsetneq \mathfrak{g}$. We claim it is an ideal of dimension $n-1$. Indeed, define the representation $\mathfrak{h} \to \mathrm{End}(\mathfrak{g}/\mathfrak{h})$ by $X \to (Y + \mathfrak{h} \to [X, Y] + \mathfrak{h})$. By lemma 5.22 the representation is nilpotent and therefore by the induction hypothesis, there exist $Y_0 \in \mathfrak{g}/\mathfrak{h} \setminus \{0\}$ such that, for all $X \in \mathfrak{h}$, we have $[X, Y_0] = 0$ in $\mathfrak{g}/\mathfrak{h}$. Let $X_0 \in \mathfrak{g}$ be a lift to $\mathfrak{g}$ of Y_0. Then $[X, X_0] \in \mathfrak{h}$ and therefore $\mathfrak{h} + \langle X_0 \rangle$ is a subalgebra. As $\mathfrak{h}$ was maximal, $\mathfrak{g} = \mathfrak{h} + \langle X_0 \rangle$. We conclude that $\dim(\mathfrak{h}) = n-1$ and that $\mathfrak{h}$ is an ideal.

We apply the induction hypotheses to $\mathfrak{h}$: Let V_0 be the set of all vectors $v \in V$ such that $\rho(\mathfrak{h})v = 0$ (V_0 is non-empty because of the induction hypothesis). V_0 is stable by X_0: if $v \in V_0$ then, for all $X \in \mathfrak{h}$, $XX_0v = X_0Xv + [X, X_0]v = 0$. We apply now the induction hypothesis to $\langle X_0 \rangle$ and obtain that there exists $v \in V_0$ such that $X_0v = 0$ and this concludes the proof. □

We may now give the proof of Engel's theorem:

Proof. We are supposing that $\mathfrak{g}$ is such that $\mathrm{ad}_X \subset \mathrm{End}(\mathfrak{g})$ is nilpotent for all $X \in \mathfrak{g}$. From theorem 5.24 applied to the representation $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$, there exists a nonzero element $X \in \mathfrak{g}$ such that $\mathrm{ad}_{\mathfrak{g}}(X) = 0$. That is, the center $\mathrm{Z}(\mathfrak{g})$ of $\mathfrak{g}$ is non-trivial. We prove the theorem by induction on the dimension of $\mathfrak{g}$. Indeed, the theorem is clear if the dimension of $\mathfrak{g}$ is null. Assume the theorem true for the Lie algebra $\mathfrak{g}/\mathrm{Z}(\mathfrak{g})$. From the nilpotency of $\mathrm{Z}(\mathfrak{g})$ and $\mathfrak{g}/\mathrm{Z}(\mathfrak{g})$ we conclude that $\mathfrak{g}$ is nilpotent (if there exists n such that $\mathscr{D}^n(\mathfrak{g}/\mathrm{Z}(\mathfrak{g})) = \{0\}$, that is $\mathscr{D}^n(\mathfrak{g}) \subset \mathrm{Z}(\mathfrak{g})$ then $\mathscr{D}^{n+1}(\mathfrak{g}) = \{0\}$).

□

The following corollary makes it explicit the fact that a nilpotent algebra has a representation on the space of strictly upper triangular matrices.

Corollary 5.25. *Let $\rho : \mathfrak{g} \to \mathrm{End}(V)$ be a representation such that each element in $\rho(\mathfrak{g})$ is nilpotent. Then there exists a basis of V such that the matrix of $\rho(X)$ is upper triangular with vanishing diagonal coefficients for all $X \in \mathfrak{g}$.*

Proof. Let $v_1 \in V$ be the an element given by the previous theorem such that $\rho(X)(v_1) = 0$ for all $X \in \mathfrak{g}$. Let V_1 be the vector space generated by v_1 and consider now the representation induced by ρ on V/V_1. It is also nilpotent and therefore one can find a vector $v_2 \in V$ such that $\rho(X)v_2 \in V_1$ for all X. By induction we construct a basis (v_i) such that the matrix of ρ with respect to it is upper triangular with zero diagonal coefficients. □

5.4.2 Solvable algebras: Lie's theorem

We will prove that solvable algebras always admit a representation with values in upper triangular matrices. This is the analogue of the corollary following Engel's theorem in the case of nilpotent algebras. But one should be careful because the theorem is not valid in general for Lie algebras defined over $\mathbb{R}$.

Theorem 5.26 (Lie's theorem)**.** *Let $\mathfrak{g}$ be a solvable Lie algebra over $\mathbb{C}$. Let $\rho : \mathfrak{g} \to \mathrm{End}(V)$ be a representation on a finite dimensional complex vector space. Then there exists $v \in V \setminus \{0\}$ and a function $\lambda : \mathfrak{g} \to \mathbb{C}$ such that $\rho(X)v = \lambda(X)v$ for all $X \in \mathfrak{g}$.*

Proof. By induction on the dimension of $\mathfrak{g}$. The zero dimensional case is obvious. Suppose now that the result is true for dimensions less than or equal to $n-1$. Let $\mathfrak{g}$ be a solvable Lie algebra with $\dim(\mathfrak{g}) = n$. We divide the proof in several steps:

1. Find an ideal of codimension one $\mathfrak{h} \subset \mathfrak{g}$. Consider any codimension one subspace containing $[\mathfrak{g},\mathfrak{g}] \subsetneq \mathfrak{g}$ (it is a proper inclusion because the algebra is solvable). $\mathfrak{h}$ is clearly an ideal because $[\mathfrak{h},\mathfrak{g}] \subset [\mathfrak{g},\mathfrak{g}]$.

2. Apply the induction hypothesis to $\mathfrak{h}$: there exists $w_0 \in V$ and $\lambda : \mathfrak{h} \to \mathbb{C}$ such that $\rho(Y)w_0 = \lambda(Y)w_0$ for all $Y \in \mathfrak{h}$.

3. Let $X_0 \in \mathfrak{g} \setminus \mathfrak{h}$. Define the sequence $w_0, \cdots, w_k$ where $w_{i+1} = \rho(X)w_i$ and k is the largest index so that $(w_0, \cdots, w_k)$ is a linearly independent family. The subspaces $W_i = \langle w_0, \cdots, w_i \rangle$ clearly satisfy, for $0 \le i \le k-1$,

$$\rho(X_0)W_i \subset W_{i+1},$$

   moreover, $\rho(X_0)W_k \subset W_k$.

4. We show now that $\rho(Y)(X) = \lambda(Y)X$ for all $X \in W_k$ and $Y \in \mathfrak{h}$.

- Prove first that $\rho(Y)w_j = \lambda(Y)w_j \mod W_{j-1}$ for $j \le k$ and all $Y \in \mathfrak{h}$. The case $j = 0$ is obvious defining $W_{-1} = \{0\}$. Compute, for $Y \in \mathfrak{h}$, using the induction hypothesis,

$$\rho(Y)w_{j+1} = \rho(Y)\rho(X_0)w_j = \rho(X_0)\rho(Y)w_j + \rho([Y,X_0])w_j$$

$$= \rho(X_0)\lambda(Y)w_j + \lambda([Y,X_0])w_j \mod (W_{j-1}) = \lambda(Y)w_{j+1} \mod (W_j)$$

- If $Z = [Y, X_0]$ for $Y \in \mathfrak{h}$ then $\lambda(Z) = 0$. Indeed, compute (from the previous item)

$$\operatorname{tr}(\rho(Z)_{|W_k}) = (k+1)\lambda(Z),$$

and observe that the trace of any commutator is null.

- We finish the proof that $\rho(Y)w_j = \lambda(Y)w_j$ for all j and $Y \in \mathfrak{h}$ by induction. By definition, $\rho(Y)w_0 = \lambda(Y)w_0$ for all $Y \in \mathfrak{h}$. Suppose the formula correct for j. We compute

$$\rho(Y)w_{j+1} = \rho(Y)\rho(X_0)w_j = \rho(X_0)\rho(Y)w_j + \rho([Y,X_0])w_j$$

$$= \rho(X_0)\lambda(Y)w_j + 0 = \lambda(Y)w_{j+1}.$$

5. We choose finally an eigenvector v with eigenvalue μ for $\rho(X_0)$ in W_k and define the extension of the map $\lambda : \mathfrak{h} \to \mathbb{C}$ to $\mathfrak{g}$ by $\lambda(X_0) = \mu$.

□

Analogously to the case of nilpotent algebras, there exists a representation of a solvable algebra in the space of upper triangular matrices. We leave the proof to the reader.

Corollary 5.27. *Let $\mathfrak{g}$ be a solvable Lie algebra over $\mathbb{C}$. Let $\rho : \mathfrak{g} \to \operatorname{End}(V)$ be a representation on a finite dimensional complex vector space. Then there exists a basis of V such that the matrix of $\rho(X)$ is upper triangular for all $X \in \mathfrak{g}$.*

5.5 Semi-simple Lie algebras and the Killing form

Recall that a Lie algebra is simple if it is not abelian and it admits no non-trivial ideal.

Remark 5.28. *The Lie algebra $\mathfrak{g}$ is simple if and only if the representation $\operatorname{ad} : \mathfrak{g} \to \operatorname{End}(\mathfrak{g})$ is irreducible.*

Definition 5.29. *A Lie algebra $\mathfrak{g}$ is semi-simple if its only abelian ideal is $\{0\}$.*

Remark 5.30. *If* $\mathfrak{g}$ *is semi-simple then its center is trivial. Equivalently,* $\ker(\mathrm{ad}) = \{0\}$.

Example 5.31. *A direct sum of simple algebras is a semi-simple algebra and we prove later that any semi-simple algebra is of this form.*

We state a theorem which decomposes a Lie algebra into a solvable ideal and a subalgebra which is semisimple:

Theorem 5.32 (Levi-Malcev)**.** *Any Lie algebra* $\mathfrak{g}$ *has a decomposition as a direct sum of vector spaces*

$$\mathfrak{g} = \mathfrak{s} + \mathrm{rad}(\mathfrak{g}),$$

where $\mathfrak{s}$ *is a semi-simple algebra.*

We now define a bilinear form which has an important role in characterizing semi-simplicity and also solvability.

Definition 5.33. *Let* $\mathfrak{g}$ *be a Lie algebra and* $\rho : \mathfrak{g} \to \mathrm{End}(V)$ *be a representation (*V *is finite dimensional). The Killing form of the representation is the bilinear form*

1. $\mathrm{B}_\rho(X, Y) = \mathrm{tr}(\rho(X)\rho(Y))$,
2. *In the case* $\rho = \mathrm{ad}$ *we write simply* $\mathrm{B}(X, Y)$ *or* $\mathrm{B}_{\mathfrak{g}}(X, Y) = \mathrm{tr}(\mathrm{ad}(X)\,\mathrm{ad}(Y))$ *and say that the form is the Killing form of the Lie algebra.*

Clearly, the Killing form of a representation is symmetric.

Lemma 5.34. *For any* $X \in \mathfrak{g}$ *the linear map* $\mathrm{ad}_X \in \mathrm{End}(\mathfrak{g})$ *is skew-symmetric with respect to* B_ρ*. That is,*

$$B_\rho(\mathrm{ad}_X Y, Z) + B_\rho(Y, \mathrm{ad}_X Z) = 0$$

for all $Y, Z \in \mathfrak{g}$.

Proof. We compute

$$B_\rho(\mathrm{ad}_X Y, Z) = \mathrm{tr}(\rho([X,Y]), \rho(Z)) = \mathrm{tr}([\rho(X), \rho(Y)]\rho(Z)) = \mathrm{tr}\left(\left(\rho(X)\rho(Y) - \rho(Y)\rho(X)\right)\rho(Z)\right)$$

$$= \mathrm{tr}\left(\rho(X)\rho(Y)\rho(Z) - \rho(Y)\rho(X)\rho(Z)\right) = \mathrm{tr}\left(\rho(Y)\rho(Z)\rho(X) - \rho(Y)\rho(X)\rho(Z)\right)$$

$$= \mathrm{tr}\left(\rho(Y)\left(\rho(Z)\rho(X) - \rho(X)\rho(Z)\right)\right) = \mathrm{tr}\left(\rho(Y)\left(\rho(Z)\rho(X) - \rho(X)\rho(Z)\right)\right) = -B_\rho(Y, \mathrm{ad}_X Z).$$

□

Example 5.35. *Let* $\mathfrak{g} = \mathrm{M}_n(\mathbb{R})$. *Then*

$$\mathrm{B}(X,Y) = 2n\,\mathrm{tr}(XY) - 2\,\mathrm{tr}(X)\,\mathrm{tr}(Y).$$

We verify the equality on the basis of $\mathfrak{g}$ *given by the elementary matrices* E_{ij} *whose only non-zero coefficient is* 1 *at the* (i,j) *entry.*

$$B(E_{ij},E_{kl}) = \mathrm{tr}(\mathrm{ad}_{E_{ij}}\,\mathrm{ad}_{E_{kl}}) = \sum_{r,s}\langle \mathrm{ad}_{E_{ij}}\,\mathrm{ad}_{E_{kl}}\,E_{rs},E_{rs}\rangle,$$

where $\langle\cdot,\cdot\rangle$ *is defined so that* $\langle E_{ij},E_{kl}\rangle = \delta_{ij}\delta_{kl}$. *We obtain*

$$\begin{aligned}\mathrm{ad}_{E_{ij}}\,\mathrm{ad}_{E_{kl}}\,E_{rs} &= [E_{ij},[E_{kl},E_{rs}]] = [E_{ij},\delta_{lr}E_{ks} - \delta_{sk}E_{rl}]\\ &= \delta_{lr}(\delta_{jk}E_{is} - \delta_{si}E_{kj}) - \delta_{sk}(\delta_{jr}E_{il} - \delta_{li}E_{rj}).\end{aligned}$$

Now,

$$\langle \mathrm{ad}_{E_{ij}}\,\mathrm{ad}_{E_{kl}}\,E_{rs},E_{rs}\rangle = \delta_{lr}(\delta_{jk}\delta_{ir} - \delta_{si}\delta_{kr}\delta_{sj}) - \delta_{sk}(\delta_{jr}\delta_{ir}\delta_{ls} - \delta_{li}\delta_{rs})$$

and then

$$\sum_{r,s}\langle \mathrm{ad}_{E_{ij}}\,\mathrm{ad}_{E_{kl}}\,E_{rs},E_{rs}\rangle = 2n\delta_{jk}\delta_{li} - 2\delta_{ij}\delta_{kl}.$$

On the other hand we have the easy computation

$$2n\,\mathrm{tr}(E_{ij}E_{kl}) - 2\,\mathrm{tr}(E_{ij})\,\mathrm{tr}(E_{kl}) = 2n\delta_{jk}\delta_{li} - 2\delta_{ij}\delta_{kl},$$

which shows that the two sides are equal.

The Killing form of an ideal of a Lie algebra is the restriction of the Killing form of the Lie algebra:

Proposition 5.36. *Let* $\mathfrak{h} \subset \mathfrak{g}$ *be an ideal. Then*

$$B_{\mathfrak{h}} = \mathrm{B}|_{\mathfrak{h}},$$

wher $B_{\mathfrak{h}}$ *is the Killing form of* $\mathfrak{h}$ *and* $\mathrm{B}|_{\mathfrak{h}}$ *is the Killing form of* $\mathfrak{g}$ *restricted to* $\mathfrak{h}$.

Proof. We consider a basis of $\mathfrak{h}$ and complete it to a basis of $\mathfrak{g}$. In that basis, for all $X \in \mathfrak{h}$, the matrix of $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ is of the form

$$[\mathrm{ad}_X] = \begin{pmatrix} A_X & \star \\ 0 & 0 \end{pmatrix}.$$

Note that the matrix of $\mathrm{ad}_X : \mathfrak{h} \to \mathfrak{h}$ is simply A_X. Therefore, for all $X, Y \in \mathfrak{h}$,

$$B(X,Y) = \mathrm{tr}([\mathrm{ad}_X][\mathrm{ad}_Y]) = B_{\mathfrak{h}}(X,Y).$$

□

Example 5.37. $\mathfrak{sl}(n,\mathbb{R}) \subset \mathfrak{gl}(n,\mathbb{R})$ *is an ideal. Therefore for* $X, Y \in \mathfrak{sl}(n,\mathbb{R})$,

$$B_{\mathfrak{sl}(n,\mathbb{R})}(X,Y) = 2n\,\mathrm{tr}(XY).$$

5.6 Cartan's criteria

5.6.1 Linear algebra recollection: The Jordan-Chevalley theorem

Here we assume that V is a vector space over $\mathbb{C}$.

Definition 5.38. *A linear map $T : V \to V$ is semi-simple if it is diagonalizable.*

We will assume the following theorem of linear algebra:

Theorem 5.39 (Jordan-Chevalley decomposition)**.** *Let $T : V \to V$ be a linear map. Then there exists a unique decomposition*

$$T = T_s + T_n$$

where T_s is semi-simple and T_n is nilpotent with $[T_s, T_n] = 0$. Moreover, T_s and T_n are polynomials in T (with no constant term) and therefore commute also with T.

The decomposition of ad_T is inherited from the decomposition of T:

Proposition 5.40. *If $T = T_s + T_n$ is the Jordan-Chevalley decomposition of a linear map $T : V \to V$ then*

$$\mathrm{ad}_T = \mathrm{ad}_{T_s} + \mathrm{ad}_{T_n}$$

is the Jordan-Chevalley decomposition of ad_T.

Proof. The proof that ad_{T_n} is nilpotent was already given in lemma 5.22. In the next lemma we prove that ad_{T_s} is semi-simple. It remains to show that $[\mathrm{ad}_{T_n}, \mathrm{ad}_{T_s}] = 0$. This follows from the fact that ad is a morphism of Lie algebras.

□

Lemma 5.41. *If a linear map $T : V \to V$ is semi-simple then the linear map*

$$\mathrm{ad}_T : \mathrm{End}(V) \to \mathrm{End}(V)$$

is semi-simple.

Proof. Write the matrix of T in a basis where it is a diagonal matrix D with coefficients λ_i, $1 \le i \le n$. In the same basis a basis of $\mathrm{End}(V)$ is given by E_{ij} where the coefficient at (i, j) is 1 and 0 elsewhere. Compute

$$[D, E_{ij}] = DE_{ij} - E_{ij}D = \sum_{k=1}^{n} \lambda_k (E_{kk}E_{ij} - E_{ij}E_{kk}) = \sum_{k=1}^{n} \lambda_k (\delta_{ki}E_{kj} - \delta_{jk}E_{ik}) = (\lambda_i - \lambda_j)E_{ij}.$$

This shows that the vectors E_{ij} are eigenvectors of ad_D. □

5.6.2 Cartan's criterium for a solvable Lie algebra

We state in this section the following Cartan's criterium:

Theorem 5.42. *A Lie algebra* $\mathfrak{g}$ *is solvable if and only if* $B_{\mathfrak{g}}(X,Y)=0$ *for all* $X\in\mathfrak{g}$ *and* $Y\in[\mathfrak{g},\mathfrak{g}]$.

Proof. We suppose that $\mathfrak{g}$ is a solvable Lie algebra defined over $\mathbb{C}$ and use Lie's theorem so that ad_X is a triangular matrix for all $X\in\mathfrak{g}$. Then $\mathrm{ad}_{[X_1,X_2]}$ is strictly triangular for all $X_1,X_2\in\mathfrak{g}$ and so is $\mathrm{ad}_X\, ad_{[X_1,X_2]}$. This shows that $B_{\mathfrak{g}}(X,Y)=0$ for all $X\in\mathfrak{g}$ and $Y\in[\mathfrak{g},\mathfrak{g}]$.

Suppose now $B_{\mathfrak{g}}(X,Y)=0$ for all $X\in\mathfrak{g}$ and $Y\in[\mathfrak{g},\mathfrak{g}]$. We claim that $[\mathfrak{g},\mathfrak{g}]$ is nilpotent which clearly implies that $\mathfrak{g}$ is solvable. It suffices, by Engel's theorem, to show that each element in $[\mathfrak{g},\mathfrak{g}]$ is nilpotent.

The proof is by contradiction: suppose ad_Y is not nilpotent and write the Jordan-Chevalley decomposition $\mathrm{ad}_Y=S+N$ where S is semi-simple and N nilpotent with $NS=SN$. Let μ_i, $1\le i\le n$ be the eigenvalues and v_i be the corresponding eigenvectors of ad_Y. Define the conjugate $\bar{S}$ with eigenvalues $\bar{\mu_i}$ so that $\mathrm{tr}\,(S\bar{S})=\sum|\mu_i|^2>0$. Note that $N\bar{S}=\bar{S}N$ which implies $(\bar{S}N)^m=\bar{S}^mN^m$ and therefore $\bar{S}N$ is nilpotent. We obtain that

$$\mathrm{tr}\,(\bar{S}\,\mathrm{ad}_Y)=\mathrm{tr}\,(\bar{S}(S+N))=\mathrm{tr}\,\bar{S}S>0.$$

On the other hand, writing $Y=[X_1,X_2]$ (or a sum of commutators) we obtain

$$\mathrm{tr}\,(\bar{S}\,\mathrm{ad}_Y)=\mathrm{tr}\,(\bar{S}ad_{[X_1,X_2]})=\mathrm{tr}\,(\bar{S}[ad_{X_1},ad_{X_2}])=\mathrm{tr}\,([\bar{S},ad_{X_1}]\,\mathrm{ad}_{X_2})=\mathrm{tr}\,((\mathrm{ad}_{\bar{S}}(\mathrm{ad}_{X_1}))\,\mathrm{ad}_{X_2}).$$

Note that, by proposition 5.40, ad_S is semisimple and $\mathrm{ad}_S=q(\mathrm{ad}_{\mathrm{ad}_Y})$ for a polynomial q (with $q(0)=0$). There exists also a polynomial r (with $r(0)=0$) such that $\mathrm{ad}_{\bar{S}}=r(\mathrm{ad}_{\mathrm{ad}_Y})$. Indeed ad_S is diagonal with eigenvalues $\mu_i-\mu_j$. It suffices to find a polynomial p such that $p(\mu_i-\mu_j)=\bar{\mu}_i-\bar{\mu}_j$ and $p(0)=0$ and then make $r=p\circ q$.

We have to compute then $\mathrm{tr}\,((r(\mathrm{ad}_{\mathrm{ad}_Y})(\mathrm{ad}_{X_1}))\,\mathrm{ad}_{X_2})$ where r is a polyomial with $r(0)=0$. In particular

$$\begin{aligned}&\mathrm{tr}\,((\mathrm{ad}_{\mathrm{ad}_Y}(\mathrm{ad}_{X_1}))\,\mathrm{ad}_{X_2})\\ &=\mathrm{tr}\left((\mathrm{ad}_Y\,\mathrm{ad}_{X_1}-\mathrm{ad}_{X_1}\,\mathrm{ad}_Y)\,\mathrm{ad}_{X_2}\right)\\ &=\mathrm{tr}\left((\mathrm{ad}_{[Y,X_1]})\,\mathrm{ad}_{X_2}\right)\end{aligned}$$

which is null by hypothesis. Analogously,

$$\mathrm{tr}\,((\mathrm{ad}^2_{\mathrm{ad}_Y}(\mathrm{ad}_{X_1}))\,\mathrm{ad}_{X_2})=\mathrm{tr}\left((\mathrm{ad}_{\mathrm{ad}_Y}\,\mathrm{ad}_{\mathrm{ad}_Y}(\mathrm{ad}_{X_1}))\,\mathrm{ad}_{X_2}\right)=\mathrm{tr}\left((\mathrm{ad}_{\mathrm{ad}_Y}(\mathrm{ad}_{[Y,X_1]}))\,\mathrm{ad}_{X_2}\right)$$

$$=\mathrm{tr}\left((\mathrm{ad}_{[Y,[Y,X_1]]})\,\mathrm{ad}_{X_2}\right)=0$$

and, also for each monomial,

$$\operatorname{tr}((\operatorname{ad}^n_{\operatorname{ad}_Y}(\operatorname{ad}_{X_1}))\operatorname{ad}_{X_2}) == \operatorname{tr}\left((\operatorname{ad}_{[Y,[Y,[\dots X_1]\dots]]})\operatorname{ad}_{X_2}\right) = 0.$$

The contradiction with $\operatorname{tr}(\bar{S}\operatorname{ad}_Y) > 0$ implies that $S = 0$ and therefore any $Y \in [\mathfrak{g},\mathfrak{g}]$ is nilpotent and we conclude that $\mathfrak{g}$ is solvable.

□

Remark 5.43. *In particular, if the Killing form is null, the Lie algebra is solvable.*

5.6.3 Cartan's criterium for a semi-simple Lie algebra

We give a characterization of semi-simple Lie algebras using the Killing form. Its proof is based on the Cartan's criterium for a solvable algebra.

Proposition 5.44. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{C}$ or $\mathbb{R}$. The following assertions are equivalent.*

1. $\mathfrak{g}$ *is semi-simple.*
2. $\operatorname{rad}(\mathfrak{g}) = \{0\}$.
3. *The Killing form is non-degenerate.*

Proof.

1. We prove 1. implies 2. by contraction: Suppose $\operatorname{rad}(\mathfrak{g}) \neq \{0\}$. As it is solvable, there exists a minimal n such that $\mathfrak{D}^n(\operatorname{rad}(\mathfrak{g})) = \{0\}$. But this implies that $\mathfrak{D}^{n-1}(\operatorname{rad}(\mathfrak{g})) \neq \{0\}$ is an abelian ideal. A contradiction with the hypothesis that $\mathfrak{g}$ is semi-simple.

2. Assume the Killing form is degenerate. We will prove that $\operatorname{rad}(\mathfrak{g}) \neq \{0\}$ establishing then that 2. implies 3. : Let
$$I = \mathfrak{g}^{\perp} = \{\, X \in \mathfrak{g} \mid B(X,Y) = 0 \text{ for all } Y \in \mathfrak{g}\,\}.$$
We show that I is an ideal. Indeed, for a $X \in I$ and $Y, Z \in \mathfrak{g}$ $B([X,Y],Z) = B(X,[Y,Z]) = 0$. The fact that it is abelian follows from Cartan's criterium because for all $X \in I$, $B(X,X) = 0$ and lemma 5.36.

3. We prove that 3. implies 1. again by contradiction: Suppose I is an abelian ideal. Then for a $X \in I$ and all $Y, Z \in \mathfrak{g}$
$$\operatorname{ad}_X \circ \operatorname{ad}_Y(Z) = [X,[Y,Z]] \in I.$$
Therefore $(\operatorname{ad}_X \circ \operatorname{ad}_Y)^2(Z) = 0$ for all Z, which implies that $\operatorname{ad}_X \circ \operatorname{ad}_Y$ is nilpotent. We conclude that $\operatorname{tr}(\operatorname{ad}_X \circ \operatorname{ad}_Y) = 0$ and B is degenerate.

□

A final proposition gives the structure of a semi-simple Lie algebra.

Proposition 5.45. *A semi-simple Lie algebra* $\mathfrak{g}$ *is a direct sum of simple algebras.*

Proof. Let $I \subset \mathfrak{g}$ be an ideal. Consider the set

$$I^{\perp} = \{ X \in \mathfrak{g} \mid B(X, Y) = 0 \text{ for all } Y \in I \}.$$

One shows easily that $I^{\perp}$ is an ideal and $\mathfrak{g} = I \oplus I^{\perp}$ is a direct sum of algebras. The proposition is proved by repeating this decomposition until we find simple ideals

□

6 Examples

In this chapter we list low dimensional Lie algebras and Lie groups and state the classification of semi-simple complex Lie algebras.

6.1 Low dimensional Lie algebras

In general, once a Lie algebra is fixed, there exists a unique simply connected Lie group with that Lie algebra up to isomorphisms. But this simply connected Lie group is not necessarily linear.

Clearly, there is only one Lie algebra of dimension one over $\mathbb{R}$ or $\mathbb{C}$. There exists two non-isomorphic Lie groups corresponding to the one dimensional Lie algebra over $\mathbb{R}$. Indeed, the abelian Lie group $\mathbb{R}$ is the simply connected one but there exists also a compact quotient, the circle $\mathrm{U}(1)$. The simply connected Lie group corresponding to a one dimensional Lie algebra over $\mathbb{C}$ is the abelian group $\mathbb{C}$.

6.1.1 Two dimensional Lie algebras

Let $\mathfrak{g}$ be a two dimensional Lie algebra over $\mathbb{R}$ or $\mathbb{C}$ which is not abelian. Then $[\mathfrak{g},\mathfrak{g}]$ is a one dimensional ideal. If one defines X as a generator of this ideal one obtains that for any $Y \in \mathfrak{g}$ which is linearly independent to X, $[X,Y] = \lambda X$ where $\lambda \neq 0$. Therefore one may chose a basis (X,Y) such that $[X,Y] = X$. One recognizes then the Lie algebra $\mathfrak{aff}(1)$ of the Lie group Aff(1) of affine transformations. Observe that the algebra $\mathfrak{aff}(1)$ is solvable but not nilpotent.

6.1.2 Three dimensional Lie algebras over $\mathbb{R}$

In this section we classify Lie algebras $\mathfrak{g}$ of dimension three. We will divide in cases according to the dimension of $[\mathfrak{g},\mathfrak{g}]$. We have already met the abelian algebra $\mathbb{R}^3$, the nilpotent Lie algebra of the Heisenberg group, $\mathfrak{h}(3,\mathbb{R})$, and the simple Lie algebras $\mathfrak{sl}(2,\mathbb{R})$ and $\mathfrak{su}(2)$. Observe that one may also form the direct sum $\mathfrak{aff}(1,\mathbb{R}) \oplus \mathbb{R}$ of the affine algebra with the one dimensional Lie algebra to obtain a solvable three dimensional example.

Clearly, there exists only one abelian Lie algebra and in this case $\dim([\mathfrak{g},\mathfrak{g}]) = 0$. We enumerate now the other cases.

1. Suppose $\dim([\mathfrak{g},\mathfrak{g}]) = 1$. Let Z be a generator of $[\mathfrak{g},\mathfrak{g}]$. There are two possibilities:

    - The algebra is nilpotent, that is, $\mathscr{C}^2(\mathfrak{g}) = [\mathfrak{g},[\mathfrak{g},\mathfrak{g}]] = \{0\}$. Therefore Z is in the center and we recognize the Lie algebra of the Heisenberg group.

- $\dim(\mathscr{C}^2(\mathfrak{g})) = 1$, that is $\mathscr{C}^2(\mathfrak{g}) = [\mathfrak{g},[\mathfrak{g},\mathfrak{g}]] = [\mathfrak{g},\mathfrak{g}]$ (the algebra is solvable but not nilpotent). In this case, there exists $X \in \mathfrak{g}$ such that $[X,Z] = Z$. We claim now that we may chose Y in the center of the algebra so that (X,Y,Z) is a basis of the algebra (and we conclude that $\mathfrak{g} = \mathfrak{aff}(1) \oplus \mathbb{R}$): write $Y = aX + bZ + cY'$ for a basis (X,Y',Z). We impose that $[X,Y] = [Y,Z] = 0$ to obtain $bZ + c[X,Y'] = 0$ and $aZ + c[Y',Z] = 0$ which has a solution with an appropriate choice of a and b.

2. Suppose $\dim([\mathfrak{g},\mathfrak{g}]) = 2$ and let (X,Y) be a basis of $[\mathfrak{g},\mathfrak{g}]$. Let $Z \in \mathfrak{g}$ so that (X,Y,Z) is a basis of $\mathfrak{g}$. There are two possibilities for the ideal $[\mathfrak{g},\mathfrak{g}]$.

   - $[\mathfrak{g},\mathfrak{g}]$ is a solvable but not abelian ideal. Than it is the algebra $\mathfrak{aff}(1)$ and we may assume $[X,Y] = Y$. Then $\mathrm{ad}_X X = 0$, $\mathrm{ad}_X Y = Y$ and $\mathrm{ad}_X Z \in [\mathfrak{g},\mathfrak{g}]$ which implies $\mathrm{tr}(\mathrm{ad}_X) = 1$ which contradicts the fact that X is a commutator (and therefore $\mathrm{tr}(\mathrm{ad}_X) = 0$). This case is then impossible.
   - $[\mathfrak{g},\mathfrak{g}]$ is an abelian ideal. Consider the linear map $\mathrm{ad}_Z|_{[\mathfrak{g},\mathfrak{g}]} : [\mathfrak{g},\mathfrak{g}] \to [\mathfrak{g},\mathfrak{g}]$. From the fact that $[\mathfrak{g},\mathfrak{g}] = \langle X,Y\rangle$ is two dimensional and abelian we conclude that $\mathrm{ad}_Z|_{[\mathfrak{g},\mathfrak{g}]} \in \mathrm{GL}(\mathfrak{g})$ (indeed, $[\mathfrak{g},\mathfrak{g}] = \langle [Z,X],[Z,Y]\rangle$). We have, on the basis (X,Y) for $[\mathfrak{g},\mathfrak{g}]$,
   $$[\mathrm{ad}_Z] = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$
   The idea now is to find a basis which puts the matrix in a canonical form. Two algebras which have different canonical forms are not isomorphic. Changing the basis of $[\mathfrak{g},\mathfrak{g}]$ through a matrix M and Z by $\lambda Z \mod [\mathfrak{g},\mathfrak{g}]$ changes the matrix above by
   $$\lambda M^{-1}\begin{pmatrix} a & b \\ c & d \end{pmatrix} M.$$
   Therefore, the algebra $\mathfrak{g}$ is determined by the classes of similarity of GL_2 up to a scalar multiple. Note that conversely any matrix in GL_2 determines a Lie algebra because the Jacobi identity is satisfied.

   We may divide then in cases according to the similarity classes:

     - ad_Z is diagonalizable: We may choose X, Y and Z so that the matrix of ad_Z is
     $$[\mathrm{ad}_Z|_{[\mathfrak{g},\mathfrak{g}]}] = \begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix}$$
     with $|d| \geq 1$.

– ad_Z is not diagonalizable. In this case one may find a basis such that

$$[\mathrm{ad}_Z|_{[\mathfrak{g},\mathfrak{g}]}] = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

or, for $b > 0$ (if $b < 0$ we permute X and Y),

$$[\mathrm{ad}_Z|_{[\mathfrak{g},\mathfrak{g}]}] = \begin{pmatrix} 1 & -b \\ b & 1 \end{pmatrix}$$

or, finally,

$$[\mathrm{ad}_Z|_{[\mathfrak{g},\mathfrak{g}]}] = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Each of these solvable Lie algebras are non-isomorphic because no further reduction is possible by a change of basis.

3. Suppose $\dim([\mathfrak{g},\mathfrak{g}]) = 3$. We claim the algebra is simple and it is either $\mathfrak{sl}(2,\mathbb{R})$ or $\mathfrak{su}(2)$. First note that if (X_1, X_2, X_3) is any basis of $\mathfrak{g}$ then $f_3 = [X_1, X_2], f_1 = [X_2, X_3], f_2 = [X_3, X_1]$ generate $[\mathfrak{g},\mathfrak{g}]$ and therefore are linearly independent. Define the matrix $A = (a_{ij})$ by $f_i = \sum_{j=1}^3 a_{ij} X_j$. We claim that A is symmetric. Indeed, the Jacobi identity (the only non-trivial one in a three dimensional Lie algebra)

$$[X_1,[X_2,X_3]] + [X_3,[X_1,X_2]] + [X_2,[X_3,X_1]] = 0$$

imposes that

$$[X_1, f_1] + [X_3, f_3] + [X_2, f_2] = \sum_{j=1}^3 a_{1j}[X_1, X_j] + \sum_{j=1}^3 a_{3j}[X_3, X_j] + \sum_{j=1}^3 a_{2j}[X_2, X_j]$$

$$= (a_{12} - a_{21}) f_3 - (a_{13} - a_{31}) f_2 + (a_{32} - a_{23}) f_2 = 0.$$

Changing the basis $X'_i = \sum_{j=1}^3 m_{ij} X_j$ (where $M = (m_{ij}) \in \mathrm{GL}_3$) we compute a change of basis $f'_i = [X'_j, X'_k] = \sum_{j=1}^3 n_{ij} f_j$ given by

$$N = (n_{ij}) = \det(M) {M^{-1}}^T.$$

Indeed,

$$f'_i = [X'_j, X'_k] = \sum_{r,s} m_{jr} m_{ks} [X_r, X_s]$$

$$= (m_{j2} m_{k3} - m_{j3} m_{k2}) f_1 - (m_{j1} m_{k3} - m_{j3} m_{k1}) f_2 + (m_{j1} m_{k2} - m_{j2} m_{k1}) f_3$$

and we recognize the adjoint matrix, that is, $\det(M){M^{-1}}^T$. Therefore the change of basis by the matrix M implies the following change of A

$$A \to \det(M){M^{-1}}^T A M^{-1}.$$

As A is symmetric one can normalize to a diagonal matrix with eigenvalues $1, \alpha, \beta$ (use first an orthogonal matrix to diagonalize A and then use a diagonal matrix to make one of the eigenvalues equal to one). We obtain then

$$[X_1, X_2] = X_3, [X_2, X_3] = \alpha X_1, [X_3, X_1] = \beta X_2.$$

Changing the basis to $X_1' = X_1, X_2' = \lambda X_2, X_3' = \lambda X_3$ changes the constants to $\alpha' = \lambda^2 \alpha$ and $\beta' = \beta$, therefore one may reduce to bases with $\alpha = \pm 1$. Analogously, by changing to $X_1' = \lambda X_1, X_2' = X_2, X_3' = \lambda X_3$ one may reduce to bases with $\beta = \pm 1$. Note now that by permuting the basis, and eventually multiplying all the basis element by -1, one can arrange so that we have a basis with $\beta = 1$ and $\alpha = 1$ or $\alpha = -1$ and we finally obtain the two simple Lie algebras, respectively, $\mathfrak{su}(2)$ or $\mathfrak{sl}(2, \mathbb{R})$.

Remark 6.1. *In order to obtain the classification of Lie algebras over $\mathbb{C}$ we follow the same path but there are simplifications:*

1. *The only matrices of ad_Z appearing in the classification of the algebras with $\dim([\mathfrak{g}, \mathfrak{g}]) = 1$ are*

$$[\mathrm{ad}_Z] = \begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix}$$

with $|d| \geq 1$ and

$$[\mathrm{ad}_Z] = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Indeed the matrices

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

and

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

are diagonalizable over $\mathbb{C}$.

2. *There exists only one simple Lie algebra (over $\mathbb{C}$) of dimension three, up to isomorphism. Indeed, the last two algebras are isomorphic (considered as algebras over $\mathbb{C}$).*

Remark 6.2. *The solvable Lie algebras such that* $\dim([\mathfrak{g},\mathfrak{g}]) = 2$ *can be realized in* M_3 *as subalgebras of the affine algebra*

$$\mathrm{aff}(2,\mathbb{R}) = \left\{ \begin{pmatrix} \mathfrak{gl}(2,\mathbb{R}) & v \\ 0 & 0 \end{pmatrix} \mid v \in \mathbb{R}^2 \right\}.$$

We describe them via the adjoint representation. Note that $\mathrm{ad} : \mathfrak{g} \to \mathrm{M}_3$ *is faithful: We write in the basis* (X, Y, Z) *the first case where*

$$[\mathrm{ad}_Z |_{[\mathfrak{g},\mathfrak{g}]}] = \begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix}.$$

Let $W = -xX - yY + tZ \in \mathfrak{g}$ *then the matrix of* ad_W *is*

$$[\mathrm{ad}_W] = \begin{pmatrix} t & 0 & x \\ 0 & td & yd \\ 0 & 0 & 0 \end{pmatrix}$$

and we obtain an isomorphic algebra in M_3 *as the image of the map* ad*:*

$$\mathrm{ad}(\mathfrak{g}) \simeq \left\{ \begin{pmatrix} t & 0 & x \\ 0 & td & y \\ 0 & 0 & 0 \end{pmatrix} \mid x, y, t \in \mathbb{R} \right\}.$$

The other examples are obtained analogously, the matrix $[\mathrm{ad}_Z |_{[\mathfrak{g},\mathfrak{g}]}]$ *appearing as an upward left block matrix.*

$$\left\{ \begin{pmatrix} t[\mathrm{ad}_Z |_{[\mathfrak{g},\mathfrak{g}]}] & v \\ 0 & 0 \end{pmatrix} \mid v \in \mathbb{R}^2 \right\}.$$

The results of the classification are put together in the following table which uses the description Bianchi of the Lie algebras enumerating them into types from I to IX.

Three dimensional Lie algebras over $\mathbb{R}$		
Bianchi classification	Lie brackets	$\dim((\mathfrak{g},\mathfrak{g}])$
I Abelian algebra	$[X,Y]=[X,Z]=[X,Z]=0$	0
II Heisenberg algebra	$[X,Y]=Z, [X,Z]=[X,Z]=0$	1
III $\mathfrak{aff}(2,\mathbb{R})\times\mathbb{R}$	$[X,Y]=Y, [X,Z]=[Y,Z]=0$	1
IV	$[X,Y]=0$, $[\mathrm{ad}_Z\,\vert_{\mathscr{D}}]=\begin{pmatrix}1&1\\0&1\end{pmatrix}$	2
V Similarity algebra	$[X,Y]=0$, $[\mathrm{ad}_Z\,\vert_{\mathscr{D}}]=\begin{pmatrix}1&0\\0&1\end{pmatrix}$	2
VI	$[X,Y]=0$, $[\mathrm{ad}_Z\,\vert_{\mathscr{D}}]=\begin{pmatrix}1&0\\0&d\end{pmatrix}$, $\vert d\vert>1$	2
VI_0 Poincaré algebra	$[X,Y]=0$, $[\mathrm{ad}_Z\,\vert_{\mathscr{D}}]=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$	2
VII	$[X,Y]=0$, $[\mathrm{ad}_Z\,\vert_{\mathscr{D}}]=\begin{pmatrix}1&-b\\b&1\end{pmatrix}$, $b>0$	2
VII_0 Euclidean algebra	$[X,Y]=0$, $[\mathrm{ad}_Z\,\vert_{\mathscr{D}}]=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$	2
VIII $\mathfrak{sl}(2,\mathbb{R})$	$[X_1,X_2]=X_3, [X_2,X_3]=-X_1, [X_3,X_1]=X_2$	3
IX $\mathfrak{su}(2,\mathbb{R})$	$[X_1,X_2]=X_3, [X_2,X_3]=X_1, [X_3,X_1]=X_2$	3

6.2 Low dimensional Lie groups

Classifying Lie groups with the above Lie algebras can be achieved considering first the simply connected Lie group with a fixed Lie algebra and then classify possible quotients by discrete subgroups. For instance, in the case of the two dimensional abelian Lie algebra, the simply connected Lie group is isomorphic to $\mathbb{R}^2$. But we will not deal here with the classification of the continuous families of quotient groups. Sometimes it is easier to give a Lie group which is not simply connected but a quotient which is easier to describe as a linear Lie group.

In two dimensions we have already seen the Lie group $\mathrm{Aff}^+(1,\mathbb{R})\subset$ whose Lie algebra is $\mathfrak{aff}(1,\mathbb{R})$.

In three dimensions, the Lie groups $\mathbb{R}^3$, the Heisenberg group $\mathrm{Heis}(3,\mathbb{R})$ and $\mathrm{Aff}^+(1,\mathbb{R})\times\mathbb{R}$ are simply connected and correspond to the Lie algebras $\mathbb{R}^3$, the Heisenberg algebra $\mathfrak{h}(3,\mathbb{R})$ and $\mathfrak{aff}(1,\mathbb{R})\oplus\mathbb{R}$.

Also the two simple real Lie algebras correspond to the Lie groups $\mathrm{SU}(2)$ and $\mathrm{SL}(2,\mathbb{R})$. Note that $\mathrm{SU}(2)$ is diffeomorphic to S^3 and therefore simply connected but $\mathrm{SL}(2,\mathbb{R})$ is not simply connected.

It remains to find Lie groups corresponding to the family of solvable Lie algebras found in the previous section. Using 6.2 we can compute Lie groups of each of the solvable Lie algebras by exponentiation:

1. In the case
$$[\mathrm{ad}_Z|_{[\mathfrak{g},\mathfrak{g}]}] = \begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix}$$
we obtain
$$\exp([\mathrm{ad}(\mathfrak{g})]) = \left\{ \begin{pmatrix} e^t & 0 & u \\ 0 & e^{dt} & v \\ 0 & 0 & 1 \end{pmatrix} \mid t,u,v \in \mathbb{R} \right\}.$$
Note that if $d = 1$ this is the similarity group of $\mathbb{R}^2$.

2. In the case
$$[\mathrm{ad}_Z|_{[\mathfrak{g},\mathfrak{g}]}] = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$
we obtain
$$\exp([\mathrm{ad}(\mathfrak{g})]) = \left\{ \begin{pmatrix} e^t & te^t & u \\ 0 & e^t & v \\ 0 & 0 & 1 \end{pmatrix} \mid t,u,v \in \mathbb{R} \right\}.$$

3. In the case
$$[\mathrm{ad}_Z|_{[\mathfrak{g},\mathfrak{g}]}] = \begin{pmatrix} 1 & -b \\ b & 1 \end{pmatrix}$$
we obtain
$$\exp([\mathrm{ad}(\mathfrak{g})]) = \left\{ \begin{pmatrix} e^t \cos bt & -e^t \sin bt & u \\ e^t \sin bt & e^t \cos bt & v \\ 0 & 0 & 1 \end{pmatrix} \mid t,u,v \in \mathbb{R} \right\}.$$

4. In the case
$$[\mathrm{ad}_Z|_{[\mathfrak{g},\mathfrak{g}]}] = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$
we obtain
$$\exp([\mathrm{ad}(\mathfrak{g})]) = \left\{ \begin{pmatrix} \cos t & \sin t & u \\ \sin t & \cos t & v \\ 0 & 0 & 1 \end{pmatrix} \mid t,u,v \in \mathbb{R} \right\},$$
which is the Euclidean group of isometries of $\mathbb{R}^2$.

Note that they are all subgroups of the affine linear group $\mathrm{Aff}(2,\mathbb{R})$.

6.3 The classical simple Lie algebras over $\mathbb{C}$

The simple Lie algebras over $\mathbb{C}$ were classified by Killing and Lie. Here we only describe the classical Lie algebras which are gathered into families called series A, B, C and D:

A_n $(n \geq 1)$	$\mathfrak{sl}(n+1, \mathbb{C})$	$\{X \in M_{n+1}(\mathbb{C}) \mid \mathrm{tr}(X) = 0\}$	$n^2 + 2n$
B_n $(n \geq 2)$	$\mathfrak{so}(2n+1, \mathbb{C})$	$\{X \in M_{2n+1}(\mathbb{C}) \mid X + X^T = 0\}$	$2n^2 + n$
C_n $(n \geq 3)$	$\mathfrak{sp}(2n, \mathbb{C})$	$\{X \in M_{2n}(\mathbb{C}) \mid X^T J + JX = 0\}$, where $J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$	$2n^2 + n$
D_n $(n \geq 4)$	$\mathfrak{so}(2n, \mathbb{C})$	$\{X \in M_{2n}(\mathbb{C}) \mid X + X^T = 0\}$	$2n^2 - n$

Remark 6.3.

- *The restrictions on n ($n \geq 1$ for A_n, $n \geq 2$ for B_n, $n \geq 3$ for C_n, $n \geq 4$ for D_n) avoid isomorphisms in low dimensions:*

$$A_1 \cong B_1 \cong C_1 \qquad (\mathfrak{sl}(2) \cong \mathfrak{so}(3) \cong \mathfrak{sp}(2))$$
$$B_2 \cong C_2 \qquad (\mathfrak{so}(5) \cong \mathfrak{sp}(4))$$
$$D_2 \cong A_1 \oplus A_1 \qquad (\mathfrak{so}(4) \cong \mathfrak{sl}(2) \oplus \mathfrak{sl}(2))$$
$$D_3 \cong A_3 \qquad (\mathfrak{so}(6) \cong \mathfrak{sl}(4))$$

- *All algebras are simple (except D_2, which is semisimple).*

Remark 6.4. *Besides the classical Lie algebras there exist five exceptional (simple) Lie algebras E_6, E_7, E_8, F_4 and G_2.*

7 The Haar measure

7.1 Haar measure

The goal of this chapter is to introduce invariant measures on Lie groups which will be important in order to understand the theory of representations. There exists a general construction of an invariant measure on any locally compact topological group.

Consider the Borel σ-algebra on a locally compact Hausdorff topological space X. A Radon measure μ is a positive measure which is finite on compact subsets and which is inner and outer regular (for any open set U, $\mu(U) = \sup_K \mu(K)$ where the supremum is taken over all subsets $K \subset U$ and for any Borel set B, $\mu(B) = \inf_U \mu(U)$ where the infimum is taken over all subsets $U \supset B$).

Remark 7.1. Let X be a locally compact Hausdorff topological space. Given a Radon measure there exits then a continuous linear map (integration with respect to the measure) from the space of compactly supported continuous functions to $\mathbb{R}$,

$$I : C_c(X) \to \mathbb{R},$$

which is positive on positive functions. Continuity meaning that, for every compact $K \subset X$, there exists a constant C_K such that, for all continuous f with support contained in K,

$$|I(f)| \leq C_K \sup_{x\in X} |f(x)|.$$

Reciprocally, by the Riez-Markov-Kakutani theorem, if there exists such a positive continuous linear map, there exists a Radon measure defining it.

Definition 7.2. *A Radon measure μ on a locally compact topological group G is said to be a left invariant (right invariant) measure if, for any Borel subset B and any g ∈ G,*

$$\mu(gB) = \mu(B) \ \ \big(\mu(Bg) = \mu(B)\big).$$

This is equivalent to imposing that, for any $f \in C_c(G)$ and $h \in G$,

$$\int_X f(hg)\mu(dg) = \int_X f(g)\mu(dg), \ \ \left(\int_X f(gh)\mu(dg) = \int_X f(g)\mu(dg)\right).$$

Example 7.3. The Lebesgue measure on $\mathbb{R}^n$ is a left and right invariant measure. On ${\mathbb{R}^*}_+$ with the multiplication group structure, $\mu = \frac{dt}{t}$ is a left and right invariant measure. Indeed,

$$\mu(sB) = \int_{sB} \frac{dt}{t} = \int_B \frac{d(s^{-1}t)}{s^{-1}t} = \int_B \frac{dt}{t} = \mu(B).$$

The left and right invariant measure with total mass one on the circle is given by $d\theta$ in the usual parametrization $\theta \to \mathrm{e}^{2\pi i\theta}$.

Example 7.4. *Let G be the affine group in dimension 1. That is the group given by affine transformations $x \to ax + b$. It has a representation in $GL(\mathbb{R}^2)$ given by matrices of the form (with $a > 0$)*

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}.$$

The group is homeomorphic to $\mathbb{R}_+^* \times \mathbb{R}$ and the measure $\mu_L = \frac{1}{a^2} da db$ is left invariant ($da db$ is the usual Lebesgue measure on $\mathbb{R}_+^* \times \mathbb{R}$). One needs to check what happens when translating by a fixed group element. One computes

$$\begin{pmatrix} h_1 & h_2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} h_1 a & h_1 b + h_2 \\ 0 & 1 \end{pmatrix}$$

and therefore

$$\frac{1}{(h_1 a)^2} d(h_1 a) d(h_1 b + h_2) = \frac{1}{a^2} da db.$$

The right invariant measure is different and is given by $\mu_R = \frac{1}{a} da db$.

The following theorem will not be proven in these notes. We will, on the other hand, give a proof of the existence of an invariant measure for certain matrix groups.

Theorem 7.5 (Existence of the Haar measure)**.** *Let G be a locally compact group. There exits a unique, up to a positive factor, left invariant measure which is non-trivial.*

Remark 7.6. *A non-trivial left invariant measure has the property that any open set has strictly positive measure. Indeed, there exists a compact subset $K \subset G$ such that $\mu(K) > 0$ and if $U \subset G$ is an open subset, consider the cover $\bigcup_{x \in K} xU$ of K. One can extract a finite cover $\bigcup_{i=1}^N x_i U$ and then*

$$0 < \mu(K) \leq \mu(\bigcup_{i=1}^N x_i U) \leq \sum_{i=1}^N \mu(x_i U) = N\mu(U),$$

where the last equality is due to the left invariance property.

If the group is compact we can pick a unique measure by imposing that the measure is a probability measure.

7.1.1 Unimodular groups

Groups such that there exists a measure which is right and left invariant are called unimodular. We will prove here that compact groups are unimodular.

Consider a locally compact topological group with a left invariant measure μ. Let $\phi: G \to G$ be a continuous automorphism of G. Then the measure $\phi_*(\mu)$ (that is, for B a Borel subset in G, $\phi_*(\mu)(B) = \mu(\phi^{-1}(B)))$ is also left invariant. Indeed, by left invariance of μ,

$$\phi_*(\mu)(gB) = \mu(\phi^{-1}(gB)) = \mu(\phi^{-1}(g)\phi^{-1}(B)) = \mu(\phi^{-1}(gB))) = \phi_*(\mu)(B).$$

In particular for an inner automorphism $Ad_{h^{-1}}$, we obtain that $Ad_{h^{-1}*}\mu$ is also left invariant. That is the measure $B \to \mu(hBh^{-1})$ is left invariant. But $\mu(hBh^{-1}) = \mu(Bh^{-1})$ by left invariance of μ. Therefore the measure $B \to \mu(Bh^{-1})$ is left invariant. By unicity of the left invariant measure we must have

$$\mu(Bh^{-1}) = \Delta(h)\mu(B)$$

for a function $\Delta: G \to \mathbb{R}_+^*$.

Definition 7.7. *We call $\Delta: G \to \mathbb{R}_+^*$ the modular function of G.*

Proposition 7.8. *The modular function $\Delta: G \to \mathbb{R}_+^*$ is a continuous homomorphism.*

Proof. Δ is an homomorphism: By definition

$$\mu(B(h_1h_2)^{-1}) = \Delta(h_1h_2)\mu(B).$$

On the other hand

$$\mu(B(h_1h_2)^{-1}) = \mu((Bh_2^{-1})h_1^{-1}) = \Delta(h_1)\mu(Bh_2^{-1}) = \Delta(h_1)\Delta(h_2)\mu(B).$$

It remains to show continuity. It is enough to show continuity at the identity, that is, for every $\varepsilon > 0$ there exists a neighborhood V of the identity such that for all $v \in V$,

$$-\varepsilon < \Delta(v) - 1 < \varepsilon.$$

This is equivalent to show, for a fixed $K \subset G$ and all $v \in V$,

$$1 - \varepsilon < \frac{\mu(Kv^{-1})}{\mu(K)} < 1 + \varepsilon.$$

Note that $\frac{\mu(Kv^{-1})}{\mu(K)} = \frac{\mu(Uv^{-1})}{\mu(U)}$ for any two sets K, U with non zero measure. The inequalities follow from the regularity of the measure. Let $K \subset G$ be a compact set with non-empty interior (so that $\mu(K) > 0$ by remark 7.6) and $U \supset K$ an open set such that $\mu(U) < \mu(K)(1+\varepsilon)$. Let V be a symmetric neighborhood ($V^{-1} = V$) of the identity element such that $KV^{-1} \subset U$ (this is possible by the continuity of the group multiplication). Then, for $v \in V$,

$$1-\varepsilon < \frac{1}{1+\varepsilon} < \frac{\mu(K)}{\mu(U)} \le \frac{\mu(Uv^{-1})}{\mu(U)} = \frac{\mu(Kv^{-1})}{\mu(K)} \le \frac{\mu(KV^{-1})}{\mu(K)} \le \frac{\mu(U)}{\mu(K)} < 1+\varepsilon.$$

□

A corollary to the above proposition is the following result because the image of a compact group by the modular function should be a compact subgroup of $\mathbb{R}_+^*$ which is $\{1\}$. Alternatively, one could have noticed that, for a compact group, $\mu(G) < \infty$, so $\mu(G) = \mu(Gg) = \Delta(g)\mu(G)$ implies $\Delta(g) = 1$ for all g.

Proposition 7.9. *Any compact group is unimodular.*

7.2 Haar measure on matrix groups

7.2.1 Groups which are homeomorphic to open sets in euclidean space

We suppose now that $G \subset \mathbb{R}^N$ is an open subset and the group operation is smooth. We will find a measure with a continuous density $h(x)$ with respect to the Lebesgue measure dl. We want, for any Borel set B and $g \in G$,

$$\mu(gB) = \int_{gB} h(x)dl(x) = \int_B h(x)dl(x) = \mu(B).$$

Recall the change of variable formula:

$$\int_{gB} h(x)dl(x) = \int_B h(gx)|J(g\cdot)|dl(x),$$

Where $J(g\cdot)$ is the Jacobian of the map $x \to gx$. Therefore we need to satisfy the equation

$$h(gx)|J(g\cdot)| = h(x),$$

for all $x, g \in G$. One may fix $x = e$, the identity of the group, and then we have the expression

$$h(g) = C\frac{1}{|J(g\cdot)|_e|}.$$

Example 7.10. Revisit the affine group with the formula above. We need to compute the Jacobian of the map, for fixed h_1, h_2,

$$(a,b) \to (h_1 a, h_1 b + h_2).$$

That is $|J((h_1,h_2)\cdot)| = h_1^2$. We obtain again the invariant measure

$$\frac{1}{h_1^2} dh_1 dh_2.$$

7.2.2 Matrix groups

This section uses basic facts of differential geometry. If you are not familiar with this language you may skip this section. The matrix groups are submanifolds of an euclidean space and one can describe their Haar measure using differential forms. Let m be the dimension of a submanifold G in $\mathbb{R}^n$ and suppose ω is a form of degree m on G. One defines a positive measure through ω:

Definition 7.11. *Let ω be an m-form on G written in local coordinates $(x_1, \cdots, x_m)$ as*

$$a(x)dx^1 \wedge \cdots \wedge dx^m.$$

The measure associated to ω, denoted $|\omega|$, is defined on the same local coordinates by a density with respect to the Lebesgue measure:

$$|\omega|(dx) = |a(x)|dx^1 \cdots dx^m.$$

We need to show that the measure is well defined, that is, the measure does not depend on the local coordinates. Suppose $\phi : V_y \to V_x$ is a diffeomorphism. Then

$$\phi^*(\omega) = a(\phi(y))J(\phi, y)dy^1 \wedge \cdots \wedge dy^m$$

and therefore

$$|\phi^*(\omega)|(dy) = |a(\phi(y))J(\phi, y)|dy^1 \cdots dy^m.$$

Now, let $f : V_x \to \mathbb{C}$ be a continuous function with compact support. Then the formula of change of variables implies

$$\int_{V_x} f(x)|\omega|(dx) = \int_{V_x} f(x)|a(x)|dx^1 \cdots dx^m$$

$$= \int_{V_y} f(\phi(y))|a(\phi(y))J(\phi, y)||dy^1 \cdots dy^m = \int_{V_y} f(\phi(y))|\omega|(dy).$$

This shows that the measure is locally well defined. In order to define the measure on the whole group we use a partition of unity.

Recall that a tangent vector X at a point g of a submanifold G is given by $X = \gamma'(0)$ where γ is a smooth curve defined on a neighbourhood of the origin with values in the submanifold G. We may then identify the Lie algebra $\mathrm{Lie}(G)$ of a linear group with its tangent space at the identity. If $G \subset \mathrm{GL}_n$ then one identifies the tangent space at $g \in G$ (denoted $T_g G$) to a linear subspace of M_n which itself is the tangent space of GL_n at that point.

Given a smooth map $\phi : G \to G$ we define the differential $\phi_* : T_g G \to T_{\phi(g)} G$ by

$$\phi_*(X) = \frac{d}{dt}\phi(\gamma(t))_{|_{t=0}}$$

Given an element $X \in \mathrm{Lie}(G)$ one defines a left invariant vector field $X^*(g) \in T_g G$ as follows:

$$X^*(g) = \frac{d}{dt} g.\mathrm{e}^{tX}{}_{|t=0}.$$

Identifying X as an element in $T_e G$ we may write $X^*(g) = L_{g_*}(X)$. It is invariant in the sense that $L_{h*}(X^*(g)) = X^*(hg)$:

$$X^*(hg) = \frac{d}{dt} hg.\mathrm{e}^{tX}{}_{|t=0} = h\frac{d}{dt} g.\mathrm{e}^{tX}{}_{|t=0} = L_{h*}(X^*(g)).$$

Differential forms which are left invariant are described similarly. Recall that the dual space of the vector space $T_g G$ (denoted $T_g^* G$) is the space of 1-forms at g. Given a smooth map $\phi : G \to G$, we define the map $\phi^* : T^*_{\phi(g)} G \to T_g^* G$ by, for $v \in T_g G$,

$$(\phi^*(\omega))(v) = \omega(\phi_*(v)).$$

As for left invariant vector fields, one can define invariant 1-forms by translating along the group a 1-form defined at $T_e^* G$:

$$L_g^*(\omega(g)) = \omega(e).$$

One extends the definitions to k-forms. If the dimension of G is m then invariant m-forms are unique up to a constant factor. They give rise to the Haar measure on G.

Proposition 7.12. *Let ω be a left invariant n-form on a linear group of dimension n. Then $|\omega|$ is a left Haar measure.*

Proof. We explained before that $|\omega|$ defines a measure. Now ω is left invariant so that $L_g^*(\omega) = \omega$. To show that it is an invariant measure we compute, for a continuous function of compact support $f \in C_c(G)$,

$$\int_G f(x)|\omega|(dx) = \int_{gG} f(x)|\omega|(dx) = \int_G f(gx)|L_g^*\omega|(dx) = \int_G f(gx)|\omega|(dx),$$

where we used the formula of change of variables in the second equality. □

Example 7.13. Consider in $\mathbb{R}^n$ the $(n-1)$-form ω defined as, where $\widehat{dx^i}$ means that this term does not appear in the sum,

$$\omega = \sum_{i=1}^{n} (-1)^{n-1} x^i dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n.$$

This form is invariant under the action of $SL(n,\mathbb{R})$. Indeed, let $v_1, \cdots, v_{n-1}$ be $n-1$ vectors in $\mathbb{R}^n$. Then at $x \in \mathbb{R}^n$,

$$\omega_x(v_1, \cdots, v_{n-1}) = \det(x, v_1, \cdots, v_{n-1}).$$

We need to show that $\omega_{gx}(gv_1, \cdots, gv_{n-1}) = \omega_x(v_1, \cdots, v_{n-1})$ which follows from the expression as a determinant.

Let $\iota : S^{n-1} \to \mathbb{R}^n$ be the embedding of the $n-1$-dimensional sphere in euclidean space and define $\omega_S = \iota^*\omega$. The measure $|\omega_S|$ is invariant under the action of the orthogonal group $SO(n)$.

8 Representations of compact groups

In this section we will deal with the general theory of representations of compact groups. The goal is to prove that all irreducible representations are finite dimensional and Peter-Weyl theorem which constructs a Hilbert space basis of $L^2(G)$ out of the matrix coefficients of irreducible representations.

8.1 Representations of topological groups

Definition 8.1. *Let G be a topological group and V a Banach space (over $\mathbb{R}$ or $\mathbb{C}$). A representation of G in V is a continuous homomorphism*

$$\rho : G \to \mathrm{GL}(V)$$

into the continuous operators with continuous inverse of V. The continuity of the map ρ is defined by, fixing any $v \in V$, the induced map $g \to \rho(g)v$ is continuous. If $\dim V = d$ we say that the representation is of degree d.

The space $\mathrm{GL}(V)$ was seen in example 1.4 to be a topological group.

An important role is played by representations in $\mathbb{C}$:

Definition 8.2. *A character is a representation of dimension one. That is a continuous homomorphism*

$$\chi : G \to \mathbb{C}^*.$$

Example 8.3. *The set of characters on $U(1) = \{ z \in \mathbb{C} \mid |z| = 1 \}$ are of the form*

$$\chi_m(z) = z^m,$$

for $m \in \mathbb{Z}$.

Definition 8.4. *Given a representation $\rho : G \to \mathrm{GL}(V)$, a subspace $W \subset V$ is said to be stable (or invariant) if $\rho(g)W = W$ for any $g \in G$. The representation $\rho^W : G \to \mathrm{GL}(W)$, defined by restricting each $\rho(g)$ to W, is said to be a subrepresentation of ρ. The representations without stable subspaces (other than the subspaces $\{0\}$ and V) are said to be irreducible.*

Definition 8.5. *A map between two representations of G, $\rho_1 : G \to \mathrm{GL}(V_1)$ and $\rho_2 : G \to \mathrm{GL}(V_2)$ is a continuous linear map $\phi : V_1 \to V_2$ (called interlacing homomorphism) satisfying, for all $g \in G$ and $v \in V_1$,*

$$\phi(\rho_1(g)v) = \rho_2(g)\phi(v).$$

The two representations are said to be similar or equivalent if the map ϕ is an isomorphism.

Example 8.6. *Representations of finite groups are the source of the theory of representations of groups and in particular the theory of representations of compact groups.*

1. *Consider an action of G on a finite set X with d elements. Define a representation of degree d as follows. Let V_X be the vector space with basis $(e_x)_{x\in X}$ and define $\tau(g)e_x = e_{gx}$. The matrices of that representation are permutation matrices.*

2. *A special case of the above representation is the regular representation of a group G. We let $X = G$ so the basis is described by $(e_g)_{g\in X}$.*

3. *Consider the space of functions $V = \{ f : G \to \mathbb{C} \}$. Define the representation $\rho : G \to$ $\mathrm{GL}(V)$, defined by*
$$(\rho(h)f)(g) = f(h^{-1}g).$$
We have indeed
$$(\rho(h_1h_2)f)(g) = f((h_1h_2)^{-1}g) = (\rho(h_2)f)({h_2}^{-1}g) = (\rho(h_1)\rho(h_2)f)(g).$$
It is equivalent to the regular representation: Let $(\delta_g)_{g\in G}$ be the basis of V consisting of characteristic functions of the singletons of G. Observe that $\rho(h)\delta_g = \delta_{hg}$. The homomorphism ϕ defined by the map on the basis elements $\phi(\delta_g) = e_g$ satisfies
$$\tau(h)\phi(\delta_g) = \tau(h)e_g = e_{hg} = \phi(\delta_{hg}) = \phi(\rho(h)\delta_g).$$

4. *A representation of degree one of a group of order d have its image contained in the set of d-roots of unity.*

Example 8.7. $\mathrm{U}(1) \simeq \mathrm{SO}(2)$ *has irreducible real representations of dimension two. Indeed, the the representation of* $\mathrm{SO}(2) \simeq [0,2\pi]/0 \sim 2\pi$ *defined by*
$$\rho_m(\theta) = \begin{pmatrix} \cos m\theta & \sin m\theta \\ -\sin m\theta & \cos m\theta \end{pmatrix}$$
is irreducible.

8.2 Duals, direct sums and tensor products

One can construct representations from given ones using direct sums, tensor products and by duality. Suppose the vector space V is finite dimensional.

Definition 8.8. *Given a representation $\rho : G \to \mathrm{GL}(V)$, one defines the dual representation $\rho^* : G \to \mathrm{GL}(V^*)$ by, for all $\phi \in V^*$ and $v \in V$,*
$$(\rho^*(g)\phi)(v) = \phi(\rho(g^{-1})v).$$

The dual representation ρ^* associated to a representation ρ is defined so that $(\rho^*(g)\phi)(\rho(g)(v)) = \phi(v)$. Note that in the infinite dimensional case one has to be careful with the definition of the topology in the dual space.

Consider two vector spaces V_1 and V_2 and their direct sum $V_1 \oplus V_2$. In the case the vector spaces spaces V_1 and V_2 are finite dimensional, $V_1 \oplus V_2$ has dimension equal to $\dim V_1 \oplus V_2 = \dim V_1 + \dim V_2$. Given basis (e_{1_i}) and (e_{2_j}) of V_1 and V_2, a basis of $V_1 \oplus V_2$ is given by $(e_{1_i} \oplus 0)$ together with $(0 \oplus e_{2_j})$.

Definition 8.9. *Let ρ_1 and ρ_2 be two representations of G in V_1 and V_2 respectively. The direct sum representation $\rho_1 \oplus \rho_2$ is the representation of G in $V_1 \oplus V_2$ defined by*

$$(\rho_1 \oplus \rho_2)(g)(v_1 \oplus v_2) = \rho_1(g)v_1 \oplus \rho_2(g)(v_2).$$

One defines, given n representations $\rho_1, \cdots, \rho_n$ in $V_1, \cdots, V_n$, a representatiion $\rho_1 \oplus \cdots \oplus \rho_n$ in the direct sum $V_1 \oplus \cdots \oplus V_n$. One may easily prove that any permutation in the components of the direct sum defines an equivalent representation.

Consider now two vector spaces V_1 and V_2 and their tensor product $V_1 \otimes V_2$. In the case the vector spaces spaces V_1 and V_2 are finite dimensional, $V_1 \otimes V_2$ has dimension equal to $\dim V_1 \oplus V_2 = \dim V_1 . \dim V_2$. Given basis (e_{1_i}) and (e_{2_j}) of V_1 and V_2, a basis of $V_1 \otimes V_2$ is given by $(e_{1_i} \otimes e_{2_j})$.

Definition 8.10. *Let ρ_1 and ρ_2 be two representations of G in V_1 and V_2 respectively. The tensor product representation $\rho_1 \otimes \rho_2$ is the representation of G in $V_1 \otimes V_2$ defined by*

$$(\rho_1 \otimes \rho_2)(g)(v_1 \otimes v_2) = \rho_1(g)v_1 \otimes \rho_2(g)(v_2).$$

If $r_{1_{ik}}(g)$ and $r_{2_{jl}}(g)$ are the matrices of ρ_1 and ρ_2 in the basis (e_{1_i}) and (e_{2_j}) of V_1 and V_2, the matrix of $(\rho_1 \otimes \rho_2)(g)$ in the basis $(e_{1_i} \oplus e_{2_j})$ is given by $r_{1_{ik}}(g) r_{2_{jl}}(g)$. That is,

$$(\rho_1 \otimes \rho_2)(g)(e_{1_k} \otimes e_{2_l}) = \sum_{i,j} r_{1_{ik}}(g) r_{2_{jl}}(g) e_{1_i} \otimes e_{2_j}.$$

As for direct sums, one defines, given n representations $\rho_1, \cdots, \rho_n$ in $V_1, \cdots, V_n$, a representatiion $\rho_1 \otimes \cdots \otimes \rho_n$ in the tensor product $V_1 \otimes \cdots \otimes V_n$. Again, any permutation in the components of the tensor product defines an equivalent representation.

Example 8.11. *Given a representation $\rho : G \to \mathrm{GL}(V)$ one may construct the representation $\rho \otimes \rho$. There are always two supplementary stable subspaces in $V \otimes V$:*

$$V \otimes V = Sym^2(V) \oplus Alt^2(V).$$

They are the eigenspaces (corresponding to eigenvalues $+1$ and -1) of the involution $\theta : V \otimes V \to V \otimes V$ defined by extending the formula $\theta(v \otimes w) = w \otimes v$ by linearity.

8.3 Invariant hermitian product

Definition 8.12. *Let $(V, \langle\cdot,\cdot\rangle)$ be a Hilbert space with a hermitian product $\langle\cdot,\cdot\rangle$. A representation $\rho : G \to \mathrm{GL}\, V$ is said to be unitary if, for all $g \in G$ and $v \in V$,*

$$||\rho(g)v|| = ||v||.$$

We also say in that case that the hermitian metric $\langle\cdot,\cdot\rangle$ is invariant.

Proposition 8.13. *Let $\rho : G \to \mathrm{GL}(V)$ be a finite dimensional representation of a compact group. Then there exits an invariant inner product on V.*

Proof. Start with any inner product $(\cdot,\cdot)$ defined on V and average with respect to the left invariant Haar measure μ. That is, define

$$\langle v, w\rangle = \int_G (\rho(g)v, \rho(g)w)\mu(dg).$$

The group being compact, the integral of the left side is finite and it is clearly an inner product. By the definition of μ, it is invariant. □

Proposition 8.14. *Given a representation $\rho : G \to \mathrm{GL}(V)$ of a compact group and a stable subspace $W \subset V$, there exists a supplementary stable subspace W'.*

Proof. Let $\langle\cdot,\cdot\rangle$ be an invariant hermitian metric on V. Then it is clear that the orthogonal space to W is stable. Indeed, note that $\rho(g)W = W$ so if $\langle w, w'\rangle = 0$ for all $w \in W$ we obtain that $\langle \rho(g)w, \rho(g)w'\rangle = 0$ which implies that $\langle w, \rho(g)w'\rangle = 0$ for all w. □

Theorem 8.15. *Any representation of a compact group in a finite dimensional vector space is a direct sum of irreducible representations.*

Proof. By induction. A degree one representation is always irreducible. Suppose the assertion is true for representations of degree strictly less than d and let $\rho : G \to \mathrm{GL}(V)$ be a representation of degree d. If it is not irreducible, there exists a stable space $W \subset V$. But then there exists a stable subspace $W' \subset V$ with $V = W \oplus W'$ and therefore $\rho = \rho^W \oplus \rho^{W'}$. The result follows because each of the subrepresentations has degree strictly less than d. □

8.4 Schur's lemma

In order to understand of the uniqueness in the above decomposition theorem we first prove Schur's lemma. It is an important tool in the theory.

Proposition 8.16 (Schur's lemma)**.** *Suppose $\phi : V_1 \to V_2$ is an interlacing homomorphism, that is, a linear map satisfying, for all $g \in G$,*

$$\phi \circ \rho_1(g) = \rho_2(g) \circ \phi$$

where $\rho_1 : G \to \mathrm{GL}(V_1)$ and $\rho_2 : G \to \mathrm{GL}(V_2)$ are two irreducible finite dimensional representations. Then either $\phi = 0$ or ϕ is an isomorphism. If $\rho_1 = \rho_2$ and the vector space V is defined over $\mathbb{C}$ then ϕ is a multiple of the identity.

Proof. From the formula satisfied by ϕ we deduce that $\mathrm{Ker}\,\phi$ and $\mathrm{Im}\,\phi$ are stable. This implies the first assertion.

If $\rho_1 = \rho_2$, consider an eigenvalue λ of $\phi : V_1 \to V_1 = V_2$ (it exits because we assume $V_1 = V_2$ is defined over $\mathbb{C}$). We obtain that $\mathrm{Ker}(\phi - \lambda Id)$ is stable. Indeed, we also have the formula $\rho_1 \circ (\phi - \lambda Id) = (\phi - \lambda Id) \circ \rho_1$. We conclude that $\phi - \lambda Id = 0$ as the kernel is not zero dimensional. □

Remark 8.17. *Define the space $\mathrm{Hom}_G(V_1, V_2)$ to be the space of linear maps from V_1 to V_2 which are equivariant with respect to representations ρ_1 and ρ_2. That is $\phi \in \mathrm{Hom}_G(V_1, V_2)$ when $\phi \circ \rho_1(g) = \rho_2(g) \circ \phi$. Schur's lemma is the assertion that the elements of $\mathrm{Hom}_G(V_1, V_2)$, if ρ_1 and ρ_2 are irreducible, are either trivial homomorphisms or isomorphims.*

Remark 8.18. *Let $\rho = \rho_1 \oplus \rho_2$ be a decomposition of a representation with $V = V_1 \oplus V_2$. Then the morphism $\iota_i : V_i \to V$ belongs to $\mathrm{Hom}_G(V_i, V)$ and the projections $\pi_i : V \to V_i$ belong to $\mathrm{Hom}_G(V, V_i)$.*

For the statement of the unicity of the decomposition it is convenient to write $V^{\oplus a}$ for the direct sum of a copies of V.

Theorem 8.19. *Let $\rho : G \to V$ be a representation. Then, up to a permutation, there exists a unique direct sum decomposition*

$$V = V_1^{\oplus a_1} \oplus \cdots \oplus V_k^{\oplus a_k}$$

corresponding to a direct sum $\rho = \rho_1^{\oplus a_1} \oplus \cdots \oplus \rho_k^{\oplus a_k}$ where the family of distinct irreducible representations $(\rho_i)_{0 \le i \le k}$ is unique up to isomorphisms.

Proof. Consider another decomposition $V = W_1^{\oplus b_1} \oplus \cdots \oplus W_l^{\oplus b_l}$ and the map $f_{j_\beta i_\alpha}$, the composition of an injection and a projection $V_{i_\alpha} \to V \to W_{j_\beta}$ where $V_{i_\alpha}, W_{j_\beta}$ are irreducible factors. It satisfies, for $v \in V_{i_\alpha}$, $f_{j_\beta i_\alpha} \circ \rho(g) v = \rho(g) \circ f_{j_\beta i_\alpha} v$. By Schur's lemma $f_{j_\beta i_\alpha}$ is either null or an isomorphism. It follows that the number of different irreducible components is the same in both decompositions, that is $k = l$ and that the multiplicity of each irreducible representation is the same. Up to a permutation we suppose that $V_i \simeq W_i$. Now, the identity map restricted to $V_i^{\oplus a_i}$ composed with the projection on $W_i^{\oplus a_i}$ should be an isomorphism. On the other hand, the identity restricted to $V_i^{\oplus a_i}$ composed with the projection on each $W_j^{\oplus a_j}$ with $j \neq i$ is trivial. This implies that $V_i^{\oplus a_i} = W_i^{\oplus a_i}$. □

Remark 8.20. *The subspaces V_i are not uniquely determined. For instance, the trivial representation (that is, $\rho(g) = Id$ for all g) of degree d has only one irreducible (which is of degree one) of multiplicity d but it can be chosen arbitrarily.*

Remark 8.21. *The components $V_i^{\oplus a_i}$ are called isotypic components of the representation and are unique.*

We fix a notation for the set of fixed vectors by a fixed representation. The dimension of this subspace is the multiplicity of the trivial representation in the decomposition by irreducible representations.

Definition 8.22. *Given a representation $\rho : G \to GL(V)$, define*

$$V^G = \{\, v \in V \mid \rho(g)v = v \text{ for all } g \in G\,\}.$$

Exercise 8.23. *All finite dimensional representations of an abelian group are of degree one.*

8.5 Characters

Definition 8.24. *Let $\rho : G \to \mathrm{GL}(V)$ be a representation in a finite dimensional space V. The character of ρ is the map $\chi : G \to \mathbb{C}^*$ given by*

$$\chi(g) = \mathrm{tr}(\rho(g)).$$

Remark that the trace of an element $H \in \mathrm{GL}(V)$ can be computed through the matrix of H in a fixed basis. In particular, the character of a representation ρ of degree d evaluated at the identity element is d. From the properties of the trace function, we have, for all $g, h \in G$,

$$\chi(ghg^{-1}) = \chi(h)$$

and

$$\chi(gh) = \chi(hg).$$

If G is a compact group one can assume that the representation is unitary. This means that there exists a basis such that the matrix representation is formed by unitary matrices. It follows that

$$\chi(g^{-1}) = \overline{\chi(g)}.$$

Proposition 8.25. *Let ρ_1 and ρ_2 be two representations and χ_1 and χ_2 be their associated characters. Then*

1. *The character of $\rho_1 \oplus \rho_2$ is $\chi_1 + \chi_2$,*
2. *The character of $\rho_1 \otimes \rho_2$ is $\chi_1 \chi_2$.*

Proof. For the character of $\rho_1 \oplus \rho_2$, just organize the matrix representation after a choice of basis as a block diagonal matrix.

Writing the tensor product representation in matricial form after a choice of basis, we compute its character as

$$\chi(g) = \sum_{i,j} r_{1_{ii}}(g) r_{2_{jj}}(g) = \chi_1(g)\chi_2(g).$$

□

8.6 Averages

In this section we start with averages on groups. We will first deal mainly with finite groups and later repeat the definitions in the case of compact groups. The arguments are very similar with a sum over element of the group having the role of an integral over the group equipped with the Haar measure.

For finite groups (or compact groups) one can always construct invariant objects by averaging over the group. Let $G \times A \to A$ be an action on a vector space A. For $a \in A$ define the average of the orbit of a as

$$\bar{a} = \frac{1}{|G|} \sum_{g \in G} g.a.$$

Then, clearly, $\bar{a}$ is fixed by the whole group G. If the group is compact we use the Haar measure to define

$$\bar{a} = \frac{1}{\mu(G)} \int_G g.a\, \mu(dg).$$

An important case of the average process is to let $A = \mathrm{Hom}(V_1, V_2)$. Given two representations $\rho_i : G \to V_i$ $(i = 1, 2)$, and any homomorphism

$$\phi : V_1 \to V_2$$

we may form the average

$$\bar{\phi} = \frac{1}{|G|}\sum_{g\in G} \rho_2(g)\circ\phi\circ\rho_1(g^{-1}).$$

Another way to describe this average is to consider ϕ as an element of $Hom(V_1, V_2) = V_1^* \otimes V_2$ with the action of the group G given by $\rho_1^*\otimes\rho_2$. The average $\bar{\phi}$ is invariant under the action $\rho_1^*\otimes\rho_2$. Schur's lemma leads to the following:

Proposition 8.26. *Suppose $\rho_i : G \to \mathrm{GL}(V_i)$ ($i = 1, 2$) are irreducible representations and $\phi : V_1 \to V_2$ is a homomorphism. Then*

1. *If V_1 and V_2 are not isomorphic then $\bar{\phi} = 0$.*
2. *If $V_1 = V_2$ (with $\dim V_1 = n$) and $\rho_1 = \rho_2$ then $\bar{\phi} = \frac{1}{n}\operatorname{tr}\phi Id$.*

Proof. As $\bar{\phi}$ is an equivariant homomorphism we may apply Schur's lemma. The first assertion is clear. For the second assertion, from Schur's lemma, we have that $\bar{\phi}$ is an homothety. To obtain the constant of homothety we compute

$$\operatorname{tr}\bar{\phi} = \frac{1}{|G|}\operatorname{tr}(\sum_{g\in G}\rho_1(g)\circ\phi\circ\rho_1(g^{-1})) = \frac{1}{|G|}\sum_{g\in G}\operatorname{tr}(\phi) = \operatorname{tr}(\phi).$$

□

Choosing basis vectors of V_1 and V_2 one obtains matrix representations $[\rho_1(g)] = (r_{i_1 j_1}(g))$ and $[\rho_2(g)] = (r_{i_2 j_2}(g))$.

Corollary 8.27. 1. *If V_1 is not isomorphic to V_2 then, for any fixed $i_1 j_1, i_2, j_2$,*

$$\frac{1}{|G|}\sum_{g\in G} r_{i_2 j_2}(g) r_{i_1 j_1}(g^{-1}) = 0$$

2. *If $V_1 = V_2$ and $\rho_1 = \rho_2$ then*

$$\frac{1}{|G|}\sum_{g\in G} r_{i_2 j_2}(g) r_{i_1 j_1}(g^{-1}) = \frac{1}{n}\delta_{j_2 i_1}\delta_{i_2 j_1}.$$

Proof. We prove the second assertion and leave the first one to the reader. If we write any homomorphism ϕ as a matrix, $\phi = (x_{ij})$, we obtain

$$\bar{\phi} = \frac{1}{|G|}\sum_{g\in G, j_2, i_1} r_{i_2 j_2}(g) x_{j_2 i_1} r_{i_1 j_1}(g^{-1}) = \frac{1}{n}\operatorname{tr}(\phi)\delta_{i_2 j_1} = \frac{1}{n}\sum_{j_2, i_1}\delta_{j_2 i_1} x_{j_2 i_1}\delta_{i_2 j_1}.$$

This formula is valid for all (x_{ij}) which implies that the coefficients are equal. Therefore

$$\frac{1}{|G|}\sum_{g\in G} r_{i_2 j_2}(g) r_{i_1 j_1}(g^{-1}) = \frac{1}{n}\delta_{j_2 i_1}\delta_{i_2 j_1}.$$

□

Remark 8.28. *The relations above may be interpreted as orthogonality relations between coefficients of the matrices of representations in a fixed basis (see the next section). More precisely, suppose the representations involved are unitary so that, in a hermitian basis,* $[\rho_1(g^{-1})] = \overline{[\rho_1(g)]}^T$*. Then one may write in 8.27*

$$\frac{1}{|G|}\sum_{g\in G} r_{i_2 j_2}(g)\overline{r_{j_1 i_1}}(g) = 0$$

and

$$\frac{1}{|G|}\sum_{g\in G} r_{i_2 j_2}(g)\overline{r_{j_1 i_1}}(g) = \frac{1}{n}\delta_{j_2 i_1}\delta_{i_2 j_1}.$$

8.6.1 The projection formula

A particular case of the average arises choosing $A = \mathrm{Hom}(V,V)$, $\rho_1(g) = Id$ the trivial representation and $\rho = \rho_2 : G \to \mathrm{GL}(V)$. Let $Id \in \mathrm{Hom}(V,V)$ be the identity map. The average of Id will be

$$\phi = \frac{1}{|G|}\sum_{g\in G}\rho(g).$$

We claim that ϕ is the projection of V into the set of fixed vectors $V^G = \{\, v \mid \rho(g)v = v \,\}$ defined in 8.22. Indeed,

$$\rho(h)\phi(w) = \frac{1}{|G|}\sum_{g\in G}\rho(hg)(w) = \frac{1}{|G|}\sum_{g\in G}\rho(g)(w) = \phi(w).$$

Moreover, it is clear that $\phi(v) = v$ for all $v \in V^G$ proving the claim. Observing that the dimension of V^G is precisely $\mathrm{tr}(\phi)$ we obtain the following:

Corollary 8.29. *Given a representation* $\rho : G \to \mathrm{GL}(V)$*, the multiplicity of the trivial representation in the irreducible decomposition is*

$$\frac{1}{|G|}\sum_{g\in G}\chi(g).$$

8.7 Orthogonality

It is convenient to introduce a Hermitian product on the space of complex functions on G.

Definition 8.30. *For φ_1, φ_2 complex functions on G we define the Hermitian product*

$$(\varphi_1, \varphi_2) = \frac{1}{|G|} \sum_{g \in G} \varphi_1(g) \overline{\varphi_2(g)}.$$

Remark 8.31. *We proved before that, for any representation of a finite group, one may find a basis such that its matrix representation on that basis is unitary. That is $[\rho(g^{-1})] = [\rho(g)]^*$. It follows that a character χ satisfies the property*

$$\chi(g^{-1}) = \overline{\chi(g)}.$$

Remark 8.32. *The proposition in the previous section, applied when the representations are unitary, imply that the matrix entries of non-isomorphic irreducible representations are orthogonal. Also, different entries of an irreducible representation are orthogonal.*

The proof of the following proposition follows from corollary 8.27 and will be left to the reader.

Proposition 8.33. *If χ and χ' are characters of two non-isomorphic irreducible representations then $(\chi, \chi') = 0$ and $(\chi, \chi) = 0$.*

Proposition 8.34. *Let χ be the character of a representation ρ and χ_0 be the character of an irreducible representation ρ_0. Then (χ, χ_0) is the number of times the representation ρ_0 appears in the decomposion by irreducibles of ρ.*

Proof. The character of a representation is the sum of the characters of a decomposition in direct sums. The last remark then implies the result. □

As a representation is determined up to isomorphism by its irreducible direct summands, the last proposition implies the following corollary.

Corollary 8.35. *Two representations with the same character are isomorphic.*

Remark 8.36. *Given a representation $\rho = \rho_1^{\oplus a_1} \oplus \cdots \oplus \rho_k^{\oplus a_k}$ where each ρ_i $(1 \leq i \leq k)$ is irreducible, we obtain from the orthogonality relations*

$$(\chi, \chi) = \sum_{1 \leq i \leq k} {a_i}^2.$$

Therefore a representation is irreducible if and only if $(\chi, \chi) = 1$.

8.8 Example: the regular representation

Recall that the regular representation of a finite group is given by defining a basis e_g and then the action $\rho_R(h)e_g = e_{hg}$. The corresponding matrices $[\rho_R(h)]$ on that basis are permutation matrices with null diagonal elements if $h \neq e$. This observation proves the following description of the character of the regular representation.

Proposition 8.37. *Let $r : G \to \mathbb{C}$ be the character of the regular representation. Then $r(e) = |G|$ and $r(h) = 0$ if $h \neq e$.*

Proposition 8.38. *Let G be a finite group.*

1. *There exists a finite familly $(\rho_i)_{1\leq i\leq k}$ of non-isomorphic irreducible representations of G. We denote by n_i their degree.*

2. *Any irreducible representation ρ_i is a direct summand in the regular representation with multiplicity equal to its degree n_i.*

3. $$\sum_{i=1}^{k} n_i^2 = |G|.$$

4. *For $h \neq e$,*
$$\sum_{i=1}^{k} n_i \chi_i(h) = 0.$$

Proof. Let ρ_i be an irreducible representation with character χ_i. The multiplicity of ρ_i in the decomposition by irreducibles of ρ_R is given by

$$(\chi_i, r) = \frac{1}{|G|}\sum_{g\in G}\chi_i(g)r(g) = \chi_i(e) = \deg(\rho_i).$$

This proves the first and second assertions. The other two then follow from the formula

$$r(h) = \sum_{i=1}^{k} n_i \chi_i(h)$$

evaluated at e and $h \neq e$. □

8.9 Characters as basis vectors of class functions

A class function $f : G \to \mathbb{C}$ is a function satisfying, for all $g, h \in G$,

$$f(hgh^{-1}) = f(g).$$

That is, it is a function which factor through the set of the conjugacy classes of G. Let H be the set of class functions. The character of a representation is an example of a class function and we prove in this section that the set of characters of irreducible representations of G is a basis of the space of class functions H. This implies that the number of non-isomorphic irreducible representations is equal to the number of conjugacy classes in the group.

8.9.1 Averages with weight a class function

Observe that if $\rho : G \to \mathrm{GL}(V)$ is a representation and $f : G \to \mathbb{C}$ is a class function then

$$\bar{\rho} = \frac{1}{|G|} \sum_{g \in G} f(g)\rho(g)$$

is equivariant with respect to ρ. Indeed, for any $v \in V$,

$$\bar{\rho}(\rho(h)v) = \frac{1}{|G|} \sum_{g \in G} f(g)\rho(g)\rho(h)v = \frac{1}{|G|} \sum_{g \in G} f(g)\rho(gh)v = \frac{1}{|G|} \sum_{g \in G} f(g)\rho(h)\rho(h^{-1}gh)v$$

$$= \frac{1}{|G|} \sum_{g \in G} f(g)\rho(h)\rho(h^{-1}gh)v = \rho(h)\frac{1}{|G|} \sum_{g \in G} f(g)\rho(h^{-1}gh)v = \rho(h)\frac{1}{|G|} \sum_{g' \in G} f(g')\rho(g')v$$

where we made a change $g' = h^{-1}gh$ and used the fact that f is a class function.

Proposition 8.39. *If ρ is irreducible with character χ and degree n then*

$$\bar{\rho} = \frac{1}{|G|} \sum_{g \in G} f(g)\rho(g)$$

is a homothety with homothety constant equal to $\frac{1}{\dim V}(f, \overline{\chi})$.

Proof. As $\bar{\rho}$ is an equivariant homomorphism we may apply Schur's lemma and conclude that it is a homothety. The constant of homothety is given by $\frac{1}{\dim V} \operatorname{tr} \bar{\rho}$:

$$\operatorname{tr} \bar{\rho} = \frac{1}{|G|} \operatorname{tr}(\sum_{g \in G} f(g)\rho(g)) = \frac{1}{|G|} \sum_{g \in G} f(g)\chi(g) = (f, \overline{\chi}).$$

□

Theorem 8.40. *The characters of irreducible representations of a finite group G form a basis of the class functions space H.*

Proof. Characters form a linearly independent family by the orthogonality relations. It remains to show that it generates H. Let $f \in H$ be orthogonal to all characters. The previous proposition implies that, for each irreducible representation $\rho_i : G \to \mathrm{GL}(V_i)$,

$$\bar{\rho}_i = \frac{1}{\dim V_i}(f, \overline{\chi_i}) Id = 0.$$

By decomposing into irreducible representations, we obtain for any representation ρ that $\bar{\rho} = 0$. Choose now $\rho = \rho_R$ the regular representation. Then

$$\bar{\rho}(e_{g_i}) = 0 = \frac{1}{|G|}\sum_{g\in G} f(g)\rho(g)e_{g_i} = \frac{1}{|G|}\sum_{g\in G} f(g)e_{gg_i}.$$

This implies $f = 0$. □

Corollary 8.41. *The number of irreducible representations of a finite group G is equal the number of conjugacy classes in G.*

Proof. Observe that the space of class functions H has dimension precisely the number of conjugacy classes in G. □

8.10 Examples and character tables

Character tables list the irreducible representations. Each row corresponds to an irreducible representation and each column corresponds to a conjugacy class. It is useful to include other informations as the number of elements in each conjugacy class as well an element in it.

8.10.1 S_3

Consider the symmetric group S_3 (of order $3! = 6$). There are three conjugacy classes whose representatives are $Id = 1, (12), (123)$ with respectively $1, 3, 2$ elements each.

Thre are two obvious irreducible representations of degree 1: the trivial representation $\mathbf{1}$ and the sign representation $\boldsymbol{\varepsilon}$. The character of the trivial one evaluated on the three conjugacy classes is the row $(1, 1, 1)$. The character of the sign representation corresponds to the row $(1, -1, 1)$.

The standard representation ρ_{st} of the permutation group in n letters is obtained by restricting the permutation matrix representation ρ_P to the subspace $W \subset \mathbb{C}^n$ defined by the equation $z_1 + \cdots + z_n = 0$. It is a degree $n-1$ representation such that $\rho_P = \rho_{st} \oplus \mathbf{1}$.

We show that ρ_{st} of S_3 is irreducible. Compute the character of the representation by chosing the basis $(1,-1,0),(0,1,-1)$ of W. Then $\chi(e)=2$, $\chi((12))=0$ and $\chi((123))=-1$. Now, $(\chi,\chi)=\frac{1}{6}(1.2^2+3.0+2.(-1)^2)=1$ and therefore the representation is irreducible.

S_3	1	(12)	(123)
	1	3	2
trivial: **1**	1	1	1
alternate: $\boldsymbol{\varepsilon}$	1	-1	1
standard: $\rho_{\mathbf{st}}$	2	0	-1

8.10.2 Abelian groups

Proposition 8.42. *The irreducible representations of an abelian group are one dimensional. If the group is finite and all irreducible representations are one dimensional then it is abelian.*

Proof. Let ρ be an irreducible representation of an abelian group G. For a fixed $g \in G$, $\rho(g): V \to V$ is equivariant with respect to G. Indeed, for all $h \in G$, $\rho(h)\rho(g)=\rho(g)\rho(h)$. By Schur's lemma, $\rho(g)$ is an homothety. It follows that, for all g, $\rho(g)$ is an homothety and therefore, as the representation is irreducible, its degree is one.

Suppose now that all irreducible representations of a finite group G are of degree one. From the formula $|G|=\sum n_i^2$ we obtain that the number of irreducible representations (which equals the number of conjugacy classes of G) is equal to the order of G. This implies that the group is abelian. □

8.10.3 The cyclic group

Let C_n be the cyclic group $\mathbb{Z}/n\mathbb{Z}$. There are n irreducible representations all of degree one. We may organize the list of representations defining ρ_0 to be the trivial representation and, for $1 \le k \le n-1$,

$$\rho_k(1)=\mathrm{e}^{\frac{2\pi i}{k}}.$$

For instance, for $n=3$ we obtain the following character table (where $w=\mathrm{e}^{\frac{2\pi i}{3}}$).

C_3	0	1	2
	1	1	1
ρ_0	1	1	1
ρ_1	1	w	w^2
ρ_2	1	w^2	w

8.10.4 Abelian subgroups

If a group has a large abelian subgroup the irreducible representations are more easily described.

Proposition 8.43. *If $A \subset G$ is an abelian subgroup of index n then the degree of the irreducible representations of G are at most n.*

Proof. Let $\rho : G \to GL(V)$ be an irreducible representation and ρ_A be its restriction to A. Let W be an invariant subspace under ρ_A of dimension one (it exist by the previous section). Consider now the space V' generated by all the images $\rho(g)W$ for $g \in G$. Clearly, because W is A invariant,

$$V' = \sum_{r \in G/A} \rho(r) W.$$

Therefore, $\dim V' \leq |G/A|$. But, V' is invariant by G. So $V' = V$ by irreducibility of G. □

8.10.5 The dihedral group

The dihedral group D_n (of order $2n$) is the isometry group of a regular polygon of n vertices. It contains a cyclic group C_n of rotations and n reflections. A presentation of the group is given choosing the generator of the rotation group r to be the rotation by $\frac{2\pi}{n}$ and one reflection s:

$$\langle r, s \mid r^n = 1,\ s^2 = 1,\ srs = r^{-1} \rangle.$$

The relation $srs = r^{-1} = r^{n-1}$ implies that $sr^k s = r^{-k} = r^{n-k}$ and therefore there are at most $n/2 + 1$ (from 0 to $[n/2]$) classes in C_n.

The cyclic group $C_n \subset D_n$ is an index two subgroup. Therefore all irreducible representations of D_n are of degree at most two. The elements of D_n are listed as r^k and sr^k, $0 \leq k \leq n-1$.

We may identify two irreducible representations of degree one: The trivial representation, $\psi_1(r^k) = 1,\ \rho_1(sr^k) = 1$, and $\psi_2(r^k) = 1,\ \rho_2(sr^k) = -1$.

The description of the irreducible representations depends on the parity of n. The number of conjugacy classes are described in each case as follows:

- If n is even there are $n/2 + 1$ classes in C_n and two classes of reflections,
- if n is odd there are $(n+1)/2$ classes in C_n and only one class of reflections.

Indeed, suppose first that n is even. There are precisely $n/2+1$ classes in C_n as one can easily verify. We claim that there are only two classes of reflections: Observe first that $r^k(sr^h)r^{-k} = sr^{-k}r^{h-k} = sr^{h-2k}$. Therefore there are at most two classes corresponding to sr^h with h even or odd. On the other hand $sr^k(sr^h)r^{-k}s = sr^k(sr^h)sr^k = sr^k r^{-h} r^k = sr^{2k-h} = sr^{h+n-2k}$ which proves that there are exactly two classes of reflections corresponding to sr^h with h odd of even. The result for n odd is proven similarly.

For n even, we may identify two other one dimensional representations: $\psi_3(r^k) = (-1)^k$, $\psi_3(sr^k) = (-1)^k$, and $\psi_4(r^k) = (-1)^k$, $\psi_4(sr^k) = (-1)^{k+1}$. They don't appear for odd n as one should have $\rho(r^n) = 1$.

Degree two representations ρ_h, $0 \le h \le n/2$, may be defined as follows:

$$\rho_h(r^k) = \begin{pmatrix} w^{hk} & 0 \\ 0 & w^{-hk} \end{pmatrix}, \quad \rho_h(sr^k) = \begin{pmatrix} 0 & w^{-hk} \\ w^{hk} & 0 \end{pmatrix},$$

where $w = \mathrm{e}^{\frac{2\pi i}{n}}$. Observe that for the values $h = 0, n/2$ the representations are reducible. But for all $0 < h < n/2$, the representations are irreducible because the matrix $\rho_h(r^k)$ have only two invariant linear spaces which, on the other hand, are not invariant under $\rho_h(sr^k)$.

We obtained all the irreducible representations as $4.1^2 + (n/2-1).2^2 = 2n = |D_n|$. The character table is:

D_n	r^k	sr^k
ψ_1	1	1
ψ_2	1	-1
ψ_3	$(-1)^k$	$(-1)^k$
ψ_4	$(-1)^k$	$(-1)^{k+1}$
ρ_h	$2\cos\frac{2\pi i}{n}$	0

We leave to the reader to find the irreducible representations and the character table in the case n is odd.

9 Peter-Weyl theorem

In this section we will deal with the general theory of representations of compact groups. The goal is to prove that all irreducible representations are finite dimensional and Peter-Weyl theorem which constructs a Hilbert space basis of $L^2(G)$ out of the matrix coefficients of irreducible representations.

9.1 Averages on compact groups

On compact groups, as for finite groups, one can always construct invariant objects by averaging over the group. We consider functions with values on a finite dimensional vector space V equipped with a hermitian metric.

Definition 9.1. *Let G be a compact Lie group and μ its normalized Haar measure. Let $f : G \to V$ be a continuous function where V is finite dimensional. Define*

$$\int_G f(g)\mu(dg)$$

as the unique vector in V associated to the continuous form

$$w \to \int_G \langle f(g), w\rangle \mu(dg).$$

Remark 9.2. .

1. One may use a basis and write the function $f : G \to V$ as a column of complex valued functions. The integral is the column of formed by integrals of each component.

2. Because the one form is invariant, the integral just defined is clearly invariant, that is, for all $h \in G$,
$$\int_G f(hg)\mu(dg) = \int_G f(g)\mu(dg).$$

3. Let $\phi : \mathscr{H} \to \mathscr{H}$ be a continuous linear map. Then
$$\phi\left(\int_G f(g)\mu(dg)\right) = \int_G \phi\big(f(g)\big)\,\mu(dg).$$

4. One can also extend the integral to Hilbert space valued functions.

Example 9.3. *Let $\rho : G \to \mathrm{GL}(V)$ be a representation. For $v \in V$ define the average of the orbit of v as*

$$\bar{v} = \int_G \rho(g)\, v\, \mu(dg).$$

Then, clearly, $\rho(g)v$ is a continuous vector valued function and using the definition above we obtain that $\bar{v}$ is fixed by the whole group G.

Example 9.4. *Given two finite dimensional representations $\rho_i : G \to \mathrm{GL}(V_i)$ ($i = 1,2$), and any homomorphism*

$$\phi : V_1 \to V_2$$

we may form the average

$$\bar{\phi} = \int_{g\in G} \rho_2(g) \circ \phi \circ \rho_1(g^{-1}) \mu(dg).$$

Another way to describe this average is to understand ϕ as an element of $Hom(V_1, V_2) = V_1^* \otimes V_2$ with the action of the group G given by $\rho_1^* \otimes \rho_2$. The average $\bar{\phi}$ is invariant under the action $\rho_1^* \otimes \rho_2$ and this is equivalent to the condition that $\bar{\phi}$ be equivariant.

Combining the discussion above with Schur's lemma we obtain as in the case of finite groups:

Proposition 9.5. *Suppose $\rho_i : G \to \mathrm{GL}(V_i)$ ($i = 1,2$) are irreducible representations and $\phi : V_1 \to V_2$ is a homomorphism. Then*

1. *If V_1 and V_2 are not isomorphic then $\bar{\phi} = 0$.*
2. *If $V_1 = V_2$ (with $\dim V_1 = n$) and $\rho_1 = \rho_2$ then $\bar{\phi} = \frac{1}{n} \operatorname{tr} \phi Id$.*

The proof of the following fundamental theorem is postponed. It is the core of Peter-Weyl theorem.

Theorem 9.6. *Irreducible representations of a compact group are finite dimensional.*

9.2 Orthogonality relations

We consider a Hermitian product on the space of complex functions on G as in the case of finite groups.

Definition 9.7. *For φ_1, φ_2 complex functions on G we define the Hermitian product*

$$(\varphi_1, \varphi_2) = \int_{g\in G} \varphi_1(g)\overline{\varphi_2(g)}\mu(dg).$$

Remark 9.8. *We let $L^2(G)$ be the space square integrable functions with respect to the left invariant Haar measure.*

Now we will study the subspace of $L^2(G)$ generated by the coefficients of a matrix representation. That is, functions of the form $g \to (\rho(g)v, w)$ where ρ is a representation of G.

Definition 9.9. *Given a unitary representation $\rho : G \to \mathrm{GL}(V)$, where V is a finite dimensional Hermitian space we define $\mathcal{M}_\rho \subset L^2(G)$ to be the subset generated by the functions*

$$g \to \langle \rho(g)v, w\rangle,$$

where $v, w \in V$ are any two vectors. They are called representative functions.

Remark 9.10. *If $(e_i)_{1\le i\le d}$ is a basis of the vector space (of dimension d) then $\mathcal{M}_\rho$ is generated by $\big(\langle \rho(e_i), e_j\rangle\big)_{1\le i,j\le d}$.*

Proposition 9.11.

1. *Let $\rho_1 : G \to \mathrm{GL}(V_1)$ and $\rho_2 : G \to \mathrm{GL}(V_2)$ be two non-isomorphic irreducible unitary representations. Then for any $u, v \in V_1$ and $u', v' \in V_2$,*

$$\int_G \langle \rho_1(g)u, v\rangle \overline{\langle \rho_2(g)u', v'\rangle} \mu(dg) = 0.$$

2. *Let $\rho : G \to \mathrm{GL}(V)$ be an irreducible unitary representation. Then*

$$\int_G \langle \rho(g)u, v\rangle \overline{\langle \rho(g)u', v'\rangle} \mu(dg) = \frac{1}{\dim V}\langle u, u'\rangle\langle v, v'\rangle.$$

Proof. Use proposition 9.5 to average the map

$$\phi_{v,v'}(u) = \langle u, v\rangle v'.$$

That is

$$\bar{\phi}_{v,v'}(u) = \int_G \langle \rho_1(g^{-1})u, v\rangle \rho_2(g)v' \ \mu(dg).$$

In the first item, $v \in V_1$ and $v' \in V_2$. We obtain an equivariant map $\bar{\phi}_{v,v'}$ (satisfying $\bar{\phi}_{v,v'} \circ \rho_1(g) = \rho_2(g) \circ \bar{\phi}_{v,v'}$), which by Schur's lemma, as the representations are not isomorphic, is null.

In the second item $v, v', u' \in V$ to obtain

$$\langle \bar{\phi}_{v,v'}(u), u'\rangle = \int_G \langle \rho(g^{-1})u, v\rangle \langle \rho(g)v', u'\rangle \ \mu(dg)$$

and using the unitary property of ρ we get

$$= \int_G \langle \rho(g^{-1})u, v\rangle \langle v', \rho(g^{-1})u'\rangle\, \mu(dg) = \int_G \langle \rho(g^{-1})u, v\rangle \overline{\langle \rho(g^{-1})u', v'\rangle}\, \mu(dg).$$

The equivariant map $\bar{\phi}_{v,v'}$ is a multiple of the identity by Schur's lemma. That is, for all $u \in V$,

$$\bar{\phi}_{v,v'} u = \lambda(v, v')u.$$

We get

$$\int_G \langle \rho(g^{-1})u, v\rangle \overline{\langle \rho(g^{-1})u', v'\rangle}\, \mu(dg) = \lambda(v, v')\langle u, u'\rangle$$

where λ is the trace of the map $\bar{\phi}_{v,v'}$ divided by $\dim V$. Now,

$$\operatorname{tr}\left(\int_G \rho(g)\circ\phi_{v,v'}\circ\rho(g^{-1})\,\mu(dg)\right) = \int_G \operatorname{tr}\left(\rho(g)\circ\phi_{v,v'}\circ\rho(g^{-1})\right)\,\mu(dg)$$

$$= \int_G \operatorname{tr}\phi_{v,v'}\,\mu(dg) = \operatorname{tr}\phi_{v,v'}\int_G \mu(dg) = \operatorname{tr}\phi_{v,v'}.$$

But, choosing a basis of V,

$$\operatorname{tr}\phi_{v,v'} = \sum_{i=1}^{\dim V} \langle \phi_{v,v'}(e_i), e_i\rangle = \sum_{i=1}^{\dim V} \langle e_i, v\rangle\langle v', e_i\rangle\,\mu(dg) = \langle v', v\rangle.$$

□

9.3 Peter-Weyl theorem

Peter-Weyl theorem states that the coefficients of all irreducible representations of a compact group generate the space $L^2(G)$. The classical case being for the group $U(1) \simeq S^1$ the fact that the irreducible representations (given by $z \to z^n$, where $|z| = 1$) generate $L^2(S^1)$.

We first consider the set of all equivalence classes of irreducible representations.

Definition 9.12. *We denote by $\hat{G}$ the set of equivalence classes of irreducible unitary representations of a compact group G.*

Recall definition 9.9 for a finite dimensional unitary representation $\rho : G \to \mathrm{GL}(V)$, the set $\mathscr{M}_\rho \subset L^2(G)$ generated by representative functions

$$g \to \langle \rho(g)e_i, e_j\rangle,$$

where (e_i) is an orthogonal basis of V.

Remark 9.13. *One can show that if two unitary representations are equivalent, they are isomorphic through a unitary isomorphism.*

Remark 9.14. *One verifies easily that equivalent representations define the same subspace and therefore we will denote $\mathscr{M}_\rho$ by $\mathscr{M}_\lambda$ where the equivalent class $[\rho]$ defined by ρ is equal to $\lambda \in \hat{G}$.*

We may state now the main theorem.

Theorem 9.15 (Peter-Weyl)**.** *Let G be a compact group. Then*

$$L^2(G) = \widehat{\bigoplus}_{\lambda \in \hat{G}} M_\lambda,$$

where $\widehat{\oplus}$ denotes the closure of the direct sum in $L^2(G)$.

Proof. Suppose, by contradiction, that the orthogonal complement of $\mathscr{H} = \widehat{\oplus}_{\lambda \in \hat{G}} M_\lambda$ (call the complement $\mathscr{H}_0$) is not empty. Note that $\mathscr{H}_0$ is closed. It is also invariant by the right regular representation: if $f \in \mathscr{H}_0$ and $u \in \mathscr{H}$, then

$$\int_G (\rho(h)v)(g)\overline{u(g)}\mu(dg) = \int_G v(gh)\overline{u(g)}\mu(dg) = \int_G v(g)\overline{u(gh^{-1})}\mu(dg).$$

But if $u(g) = \langle \rho_\lambda(g)a, b\rangle$ is a representative function then

$$u(gh^{-1}) = \langle \rho_\lambda(gh^{-1})a, b\rangle = \langle \rho_\lambda(g)(\rho(h^{-1})a), b\rangle$$

is also a representative function and therefore the integral above is zero.

The key argument is to use theorem 9.6. The regular representation restricted to $\mathscr{H}_0$ has an invariant finite dimensional subspace W such that the restricted representation is irreducible. We show now that an element $f \in W$ is itself a representative function. Indeed, choose a basis (e_i) of W and write

$$(\rho_R(h)f)(g) = \sum_{i=1}^{\dim W} c_i(h)e_i(g).$$

We have then

$$f(gh) = \sum_{i=1}^{\dim W} c_i(h)e_i(g)$$

and, making $g = 1$ we obtain

$$f(h) = \sum_{i=1}^{\dim W} c_i(h)e_i(1),$$

which is a linear combination of the representative functions $c_i(h) = \langle (\rho_R(h)f)(g), e_i\rangle$. We conclude that f itself is a representative function, a contradiction, unless $f = 0$.

□