Lie groups

E.P. van den Ban

Lecture Notes, Spring 2010

Contents

1 Groups

The purpose of this section is to collect some basic facts about groups. We leave it to the reader to prove the easy statements given in the text.

We recall that a *group* is a set G together with a map $\mu : G \times G \to G$, $(x, y) \mapsto xy$ and an element $e = e_G$, such that the following conditions are fulfilled

(a) $(xy)z = x(yz)$ for all $x, y, z \in G$;

(b) $xe = ex = x$ for all $x \in G$;

(c) for every $x \in G$ there exists an element $x^{-1} \in G$ such that $xx^{-1} = x^{-1}x = e$.

Remark 1.1 Property (a) is called *associativity* of the group operation. The element e is called the *neutral element* of the group.

The element x^{-1} is uniquely determined by the property (c); indeed, if $x \in G$ is given, and $y \in G$ an element with $xy = e$, then $x^{-1}(xy) = x^{-1}e = x^{-1}$, hence $x^{-1} = (x^{-1}x)y = ey = y$. The element x^{-1} is called the *inverse* of x.

Example 1.2 Let S be a set. Then $\mathrm{Sym}(S)$, the set of bijections $S \to S$, equipped with composition, is a group. The neutral element e equals I_S, the identity map $S \to S$, $x \mapsto x$.
If $S = \{1, \dots, n\}$, then $\mathrm{Sym}(S)$ equals S_n, the group of permutations of n elements.

A group G is said to be *commutative* or *abelian* if $xy = yx$ for all $x, y \in G$. We recall that a *subgroup* of G is a subset $H \subset G$ such that

(a) $e_G \in H$;

(b) $xy \in H$ for all $x \in H$ and $y \in H$;

(c) $x^{-1} \in H$ for every $x \in H$.

We note that a subgroup is a group of its own right. If G, H are groups, then a *homomorphism* from G to H is defined to be a map $\varphi : G \to H$ such that

(a) $\varphi(e_G) = e_H$;

(b) $\varphi(xy) = \varphi(x)\varphi(y)$ for all $x, y \in G$.

We note that the *image* $\mathrm{im}(\varphi) := \varphi(G)$ is a subgroup of H. The *kernel* of φ, defined by

$$\ker\varphi := \varphi^{-1}(\{e_H\}) = \{x \in G \mid \varphi(x) = e_H\}$$

is also readily seen to be a subgroup of G. A surjective group homomorphism is called an *epimorphism*. An injective group homomorphism is called a *monomorphism*. We recall that a group homomorphism $\varphi : G \to H$ is injective if and only if its kernel is trivial, i.e., $\ker\varphi = \{e_G\}$. A bijective group homomorphism is called an *isomorphism*. The inverse φ^{-1} of an isomorphism $\varphi : G \to H$ is a group homomorphism from H to G. Two groups G_1 and G_2 are called *isomorphic* if there exists an isomorphism from G_1 onto G_2.

If G is a group, then by an automorphism of G we mean an isomorphism of G onto itself. The collection of such automorphisms, denoted $\mathrm{Aut}(G)$, is a subgroup of $\mathrm{Sym}(G)$.

Example 1.3 If G is a group and $x \in G$, then the map $l_x : G \to G$, $y \mapsto xy$, is called *left translation* by x. We leave it to the reader to verify that $x \mapsto l_x$ is a group homomorphism from G to $\mathrm{Sym}(G)$.

Likewise, if $x \in G$, then $r_x : G \to G$, $y \mapsto yx$, is called *right translation* by x. We leave it to the reader to verify that $x \mapsto (r_x)^{-1}$ is a group homomorphism from G to $\mathrm{Sym}(G)$.

If $x \in G$, then $\mathcal{C}_x : G \to G$, $y \mapsto xyx^{-1}$ is called *conjugation* by x. We note that $\mathcal{C}_x$ is an automorphism of G, with inverse $\mathcal{C}_{x^{-1}}$. The map $\mathcal{C} : x \to \mathcal{C}_x$ is a group homomorphism from G into $\mathrm{Aut}(G)$. Its kernel is the subgroup of G consisting of the elements $x \in G$ with the property that $xyx^{-1} = y$ for all $y \in G$, or, equivalently, that $xy = yx$ for all $y \in G$. Thus, the kernel of $\mathcal{C}$ equals the *center* $Z(G)$ of G.

We end this preparatory section with the isomorphism theorem for groups. To start with we recall that a *relation* on a set S is a subset R of the Cartesian product $S \times S$. We agree to also write xRy in stead of $(x, y) \in R$. A relation $\sim$ on S is called an *equivalence relation* if the following conditions are fulfilled, for all $x, y, z \in S$,

(a) $x \sim x$ (reflexivity);

(b) $x \sim y \Rightarrow y \sim x$ (symmetry);

(c) $x \sim y \wedge y \sim z \Rightarrow x \sim z$ (transitivity).

If $x \in S$, then the collection $[x] := \{y \in S \mid y \sim x\}$ is called the *equivalence class* of x. The collection of all equivalence classes is denoted by $S/\sim$.

A *partition* of a set S is a collection $\mathcal{P}$ of non-empty subsets of S with the following properties

(a) if $A, B \in \mathcal{P}$, then $A \cap B = \emptyset$ or $A = B$;

(b) $\cup_{A\in\mathcal{P}} A = S$.

If $\sim$ is an equivalence relation on S then $S/\sim$ is a partition of S. Conversely, if $\mathcal{P}$ is a partition of S, we may define a relation $\sim_{\mathcal{P}}$ as follows: $x \sim_{\mathcal{P}} y$ if and only if there exists a set $A \in \mathcal{P}$ such that x and y both belong to A. One readily verifies that $\sim_{\mathcal{P}}$ is an equivalence relation; moreover, $S/\sim_{\mathcal{P}} = \mathcal{P}$.

Equivalence relations naturally occur in the context of maps. If $f : S \to T$ is a map between sets, then the relation $\sim$ on S defined by $x \sim y \iff f(x) = f(y)$ is an equivalence relation. If $x \in S$ and $f(x) = c$, then the class $[x]$ equals the *fiber*

$$f^{-1}(c) := f^{-1}(\{c\}) = \{y \in S \mid f(y) = c\}.$$

Let π denote the natural map $x \mapsto [x]$ from S onto $S/\sim$. Then there exists a unique map $\bar{f} : S/\sim \to T$ such that the following diagram commutes

$$\begin{array}{ccc} S & \xrightarrow{f} & T \\ \pi \downarrow & \nearrow \bar{f} & \\ S/\sim & & \end{array}$$

We say that f *factors* through a map $\bar{f} : S/\sim \to T$. Note that $\bar{f}([x]) = f(x)$ for all $x \in S$. The map $\bar{f}$ is injective, and has image equal to $f(S)$. Thus, if f is surjective, then $\bar{f}$ is a bijection from $S/\sim$ onto T.

Partitions, hence equivalence relations, naturally occur in the context of subgroups. If K is a subgroup of a group G, then for every $x \in G$ we define the *right coset* of x by $xK := l_x(K)$. The collection of these cosets, called the right *coset space*, is a partition of G and denoted by G/K. The associated equivalence relation is given by $x \sim y \iff xK = yK$, for all $x, y \in G$.

The subgroup K is called a *normal subgroup* if $xKx^{-1} = K$, for every $x \in G$. If K is a normal subgroup then G/K carries a unique group structure for which the natural map $\pi : G \to G/K$, $x \mapsto xK$ is a homomorphism. Accordingly, $xK \cdot yK = \pi(x)\pi(y) = \pi(xy) = xyK$.

Lemma 1.4 (The isomorphism theorem) *Let $f : G \to H$ be an epimorphism of groups. Then $K := \ker f$ is a normal subgroup of G. There exists a unique map $\bar{f} : G/K \to H$, such that $\bar{f} \circ \pi = f$. The factor map $\bar{f}$ is an isomorphism of groups.*

Proof: Let $x \in G$ and $k \in K$. Then $f(xkx^{-1}) = f(x)f(k)f(x)^{-1} = f(x)e_H f(x)^{-1} = e_H$, hence $xkx^{-1} \in \ker f = K$. It follows that $xKx^{-1} \subset K$. Similarly it follows that $x^{-1}Kx \subset K$, hence $K \subset xKx^{-1}$ and we see that $xKx^{-1} = K$. It follows that K is normal.

Let $x \in G$ and write $f(x) = h$. Then, for every $y \in G$, we have $yK = xK \iff f(y) = f(x) \iff y \in f^{-1}(h)$. Hence G/K consists of the fibers of f. In the above we saw that there exists a unique map $\bar{f} : G/K \to H$, such that $\bar{f} \circ \pi = f$. The factor map is bijective, since f is surjective. It remains to be checked that $\bar{f}$ is a homomorphism. Now $\bar{f}(eK) = f(e_G) = e_H$, since f is a homomorphism. Moreover, if $x, y \in G$, then $\bar{f}(xKyK) = \bar{f}(xyK) = f(xy) = f(x)f(y)$. This completes the proof. $\square$

2 Lie groups, definition and examples

Definition 2.1 (Lie group) A Lie group is a smooth (i.e., C^∞) manifold G equipped with a group structure so that the maps $\mu : (x, y) \mapsto xy,\ G \times G \to G$ and $\iota : x \mapsto x^{-1},\ G \to G$ are smooth.

Remark 2.2 For a Lie group, the group operation is usually denoted multiplicatively as above. The neutral element is denoted by $e = e_G$. Sometimes, if the group is commutative, i.e., $\mu(x, y) = \mu(y, x)$ for all $x, y \in G$, the group operation is denoted additively, $(x, y) \mapsto x + y$; in this case the neutral element is denoted by 0.

Example 2.3 We begin with a few easy examples of Lie groups.

(a) $\mathbb{R}^n$ together with addition $+$ and the neutral element 0 is a Lie group.

(b) $\mathbb{C}^n \simeq \mathbb{R}^{2n}$ together with addition $+$ and the neutral element 0 is a Lie group.

(c) $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$ is an open subset of $\mathbb{R}$, hence a smooth manifold. Equipped with the ordinary scalar multiplication and the neutral element 1, $\mathbb{R}^*$ is a Lie group. Similarly, $\mathbb{R}^+ :=]\,0, \infty\,[$ together with scalar multiplication and 1 is a Lie group.

(d) $\mathbb{C}^* := \mathbb{C} \setminus \{0\}$ is an open subset of $\mathbb{C} \simeq \mathbb{R}^2$, hence a smooth manifold. Together with complex scalar multiplication and 1, $\mathbb{C}^*$ is a Lie group.

If G_1 and G_2 are Lie groups, we may equip the product manifold $G = G_1 \times G_2$ with the product group structure, i.e., $(x_1, x_2)(y_1, y_2) := (x_1 y_1, x_2 y_2)$, and $e_G = (e_{G_1}, e_{G_2})$.

Lemma 2.4 *Let G_1, G_2 be Lie groups. Then $G := G_1 \times G_2$, equipped with the above manifold and group structure, is a Lie group.*

Proof: The multiplication map $\mu : G \times G \to G$ is given by $\mu((x_1, x_2), (y_1, y_2)) = [\mu_1 \times \mu_2]((x_1, y_1), (x_2, y_2)$. Hence, $\mu = (\mu_1 \times \mu_2) \circ (I_{G_1} \times S \times I_{G_2})$, where $S : G_2 \times G_1 \to G_1 \times G_2$ is the 'switch' map given by $S(x_2, y_1) = (y_1, x_2)$. It follows that μ is the composition of smooth maps, hence smooth.

The inversion map ι of G is given by $\iota = (\iota_1, \iota_2)$, hence smooth. $\square$

Lemma 2.5 *Let G be a Lie group, and let $H \subset G$ be both a subgroup and a smooth submanifold. Then H is a Lie group.*

Proof: Let $\mu = \mu_G : G \times G \to G$ be the multiplication map of G. Then the multiplication map μ_H of H is given by $\mu_H = \mu|_{H\times H}$. Since μ is smooth and $H \times H$ a smooth submanifold of $G \times G$, the map $\mu_H : H \times H \to G$ is smooth. Since H is a subgroup, μ_H maps into the smooth submanifold H, hence is smooth as a map $H \times H \to H$. Likewise, $\iota_H = \iota_G|_H$ is smooth as a map $H \to H$. $\square$

Example 2.6

(a) The unit circle $\mathbb{T} := \{z \in \mathbb{C} \mid |z| = 1\}$ is a smooth submanifold as well as a subgroup of the Lie group $\mathbb{C}^*$. Therefore it is a Lie group.

(b) The q-dimensional torus $\mathbb{T}^q$ is a Lie group.

So far, all of our examples of Lie groups were commutative. We shall formulate a result that asserts that interesting connected Lie groups are not to be found among the commutative ones. For this we need the concept of isomorphic Lie groups.

Definition 2.7 Let G and H be Lie groups.

(a) A *Lie group homomorphism* from G to H is a smooth map $\varphi : G \to H$ that is a homomorphism of groups.

(b) An *Lie group isomorphism* from G onto H is a bijective Lie group homomorphism $\varphi : G \to H$ whose inverse is also a Lie group homomorphism.

(c) An *automorphism* of G is an isomorphism of G onto itself.

Remark 2.8 (a) If $\varphi : G \to H$ is a Lie group isomorphism, then φ is smooth and bijective and its inverse is smooth as well. Hence, φ is a diffeomorphism.

(b) The collection of Lie group automorphisms of G, equipped with composition, forms a group, denoted $\mathrm{Aut}(G)$.

We recall that a topological space X is said to be *connected* if $\emptyset$ and X are the only subsets of X that are both open and closed. The space X is said to be *arcwise connected* if for each pair of points $a, b \in X$ there exists a continous curve $c : [0,1] \to X$ with initial point a and end point b, i.e., $c(0) = a$ and $c(1) = b$. If X is a manifold then X is connected if and only if X is arcwise connected.

We can now formulate the promised results about connected commutative Lie groups.

Theorem 2.9 *Let G be a connected commutative Lie group. Then there exist integers $p, q \geq 0$ such that G is isomorphic to $\mathbb{T}^p \times \mathbb{R}^q$.*

The proof of this theorem will be given at a later stage, when we have developed enough technology. See Theorem 6.1.

A more interesting example is the following. In the sequel we will often discuss new general concepts in the context of this important particular example.

Example 2.10 Let n be a positive integer, and let $\mathrm{M}(n, \mathbb{R})$ be the set of real $n \times n$ matrices. Equipped with entry wise addition and scalar multiplication, $\mathrm{M}(n, \mathbb{R})$ is a linear space, which in an obvious way may be identified with $\mathbb{R}^{n^2}$. For $A \in \mathrm{M}(n, \mathbb{R})$ we denote by A_{ij} the entry of A in the i-th row and the j-th column. The maps $\xi_{ij} : A \mapsto A_{ij}$ may be viewed as a system of (linear) coordinate functions on $\mathrm{M}(n, \mathbb{R})$.

In terms of these coordinate functions, the determinant function $\det : \mathrm{M}(n, \mathbb{R}) \to \mathbb{R}$ is given by

$$\det = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \xi_{1\sigma(1)} \cdots \xi_{n\sigma(n)},$$

where S_n denotes the group of permutations of $\{1, \ldots, n\}$, and where sgn denotes the sign of a permutation. It follows from this formula that det is smooth.

The set $\mathrm{GL}(n, \mathbb{R})$ of invertible matrices in $\mathrm{M}(n, \mathbb{R})$, equipped with the multiplication of matrices, is a group. As a set it is given by

$$\mathrm{GL}(n, \mathbb{R}) = \{A \in \mathrm{M}(n, \mathbb{R}) \mid \det A \neq 0\}.$$

Thus, $\mathrm{GL}(n, \mathbb{R})$ is the pre-image of the open subset $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ of $\mathbb{R}$ under det. As the latter function is continuous, it follows that $\mathrm{GL}(n, \mathbb{R})$ is an open subset of $\mathrm{M}(n, \mathbb{R})$. As such, it may be viewed as a smooth manifold of dimension n^2. In terms of the coordinate functions ξ_{ij}, the multiplication map $\mu : \mathrm{GL}(n, \mathbb{R}) \times \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$ is given by

$$\xi_{kl}(\mu(A, B)) = \sum_{i=1}^{n} \xi_{ki}(A) \xi_{il}(B).$$

It follows that μ is smooth. Given $A \in \mathrm{M}(n, \mathbb{R})$ we denote by A^{T} the transpose of A. Moreover, for $1 \leq i, j \leq n$ we denote by $M_{ij}(A)$ the matrix obtained from A by deleting the i-th row and j-th column. The co-matrix of A is defined by $A^{co}_{ij} = (-1)^{i+j} \det M_{ij}(A^{\mathrm{T}})$. Clearly, the map $A \mapsto A^{co}$ is a polynomial, hence smooth map from $\mathrm{M}(n, \mathbb{R})$ to itself. By Cramer's rule the inversion $\iota : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$, $A \to A^{-1}$ is given by

$$\iota(A) = (\det A)^{-1} A^{co}.$$

It follows that ι is smooth, and we see that $\mathrm{GL}(n, \mathbb{R})$ is a Lie group.

Example 2.11 Let V be a real linear space of finite dimension n. Let $\mathbf{v} = (v_1, \ldots, v_n)$ be an ordered basis of V. Then there is a unique linear isomorphism $e_{\mathbf{v}}$ from $\mathbb{R}^n$ onto V, mapping the j-th standard basis vector e_j onto v_j. If $\mathbf{w}$ is a second basis, then $L := e_{\mathbf{v}}^{-1} e_{\mathbf{w}}$ is a linear isomorphism of $\mathbb{R}^n$ onto itself, hence a diffeomorphism. It follows that V has a unique structure of smooth manifold such that the map $e_{\mathbf{v}}$ is a diffeomorphism, for any choice of basis $\mathbf{v}$.

We denote by $\mathrm{End}(V)$ the set of linear endomorphisms of V, i.e., linear maps of V into itself. Equipped with pointwise addition and scalar multiplication, $\mathrm{End}(V)$ is a linear space. Let $\mathbf{v} = (v_1, \ldots, v_n)$ be an ordered basis of V. Given $A \in \mathrm{End}(V)$, we write $\mathrm{mat}(A) = \mathrm{mat}_{\mathbf{v}}(A)$ for the matrix of A with respect to $\mathbf{v}$. The entries A_{ij} of this matrix are determined by $Av_j = \sum_{i=1}^n A_{ij} v_i$, for all $1 \leq j \leq n$. As in Example 2.10 we denote by $\mathrm{M}(n, \mathbb{R})$ the set of all real $n \times n$ matrices. Equipped with entry wise addition and scalar multiplication, $\mathrm{M}(n, \mathbb{R})$ is a linear space. Accordingly, mat is a linear isomorphism from $\mathrm{End}(V)$ onto $\mathrm{M}(n, \mathbb{R})$. Via this map, composition in $\mathrm{End}(V)$ corresponds with matrix multiplication in $\mathrm{M}(n, \mathbb{R})$. More precisely,

$$\mathrm{mat}(A \circ B) = \mathrm{mat}(A)\ \mathrm{mat}(B)$$

for all $A, B \in \mathrm{End}(V)$.

We note that the matrix $\mathrm{mat}_{\mathbf{v}}(A)$ equals the matrix of $e_{\mathbf{v}}^{-1} \circ A \circ e_{\mathbf{v}}$ with respect to the standard basis of $\mathbb{R}^n$. Let now $\mathbf{w} = (w_1, \ldots, w_n)$ be a second ordered basis of V and let S be the matrix of the linear endomorphism $L = e_{\mathbf{v}}^{-1} \circ e_{\mathbf{w}} \in \mathrm{End}(\mathbb{R}^n)$ with respect to the standard basis. Then from $e_{\mathbf{v}}^{-1} A e_{\mathbf{v}} = L \circ e_{\mathbf{w}}^{-1} A e_{\mathbf{w}} \circ L^{-1}$ it follows that

$$\mathrm{mat}_{\mathbf{v}}(A) = S\, \mathrm{mat}_{\mathbf{w}}(A)\, S^{-1}.$$

By conjugation invariance of determinant and trace, we find that

$$\mathrm{det mat}_{\mathbf{v}}(A) = \mathrm{det mat}_{\mathbf{w}} A \quad \text{and} \quad \mathrm{tr mat}_{\mathbf{v}}(A) = \mathrm{tr mat}_{\mathbf{w}} A$$

for all $A \in \mathrm{End}(V)$. It follows that determinant and trace are independent of the choice of basis. Hence, there exist unique maps $\det, \mathrm{tr} : \mathrm{End}(V) \to \mathbb{R}$ such that $\det A = \mathrm{det mat} A$ and $\mathrm{tr} A = \mathrm{tr mat} A$ for any choice of basis.

We denote by $\mathrm{GL}(V)$, or also $\mathrm{Aut}(V)$, the set of invertible elements of $\mathrm{End}(V)$. Then $\mathrm{GL}(V)$ is a group. Moreover, fix a basis of V, then the associated matrix map $\mathrm{mat} : \mathrm{End}(V) \to \mathrm{M}(n, \mathbb{R})$ is a diffeomorphism, mapping $\mathrm{GL}(V)$ onto $\mathrm{GL}(n, \mathbb{R})$. It follows that $\mathrm{GL}(V)$ is an open subset, hence a submanifold of $\mathrm{End}(V)$. Moreover, as mat restricts to a group isomorphism from $\mathrm{GL}(V)$ onto $\mathrm{GL}(n, \mathbb{R})$, it follows from the discussion in the previous example that $\mathrm{GL}(V)$ is a Lie group and that mat is an isomorphism of Lie groups from $\mathrm{GL}(V)$ onto $\mathrm{GL}(n, \mathbb{R})$.

Remark 2.12 In the above example we have distinguished between linear maps and their matrices with respect to a basis. In the particular situation that $V = \mathbb{R}^n$, we shall often use the map $\mathrm{mat} = \mathrm{mat}_{\mathbf{e}}$, defined relative to the standard basis $\mathbf{e}$ of $\mathbb{R}^n$ to identify the linear space $\mathrm{End}(\mathbb{R}^n)$ with $\mathrm{M}(n, \mathbb{R})$ and to identify the Lie group $\mathrm{GL}(\mathbb{R}^n)$ with $\mathrm{GL}(n, \mathbb{R})$.

We shall now discuss an important criterion for a subgroup of a Lie group G to be a Lie group. In particular this criterion will have useful applications for $G = \mathrm{GL}(V)$. We start with a result that illustrates the idea of homogeneity.

Let G be a Lie group. If $x \in G$, then the left translation $l_x : G \to G$, see Example 1.3, is given by $y \mapsto \mu(x, y)$, hence smooth. The map l_x is bijective with inverse $l_{x^{-1}}$, which is also smooth. Therefore, l_x is a diffeomorphism from G onto itself. Likewise, the right multiplication map $r_x : y \mapsto yx$ is a diffeomorphism from G onto itself. Thus, for every pair of points $a, b \in G$ both $l_{ba^{-1}}$ and $r_{a^{-1}b}$ are diffeomorphisms of G mapping a onto b. This allows us to compare structures on G at different points. As a first application of this idea we have the following.

Lemma 2.13 *Let G be a Lie group and H a subgroup. Let $h \in H$ be a given point (in the applications $h = e$ will be most important). Then the following assertions are equivalent.*

(a) *H is a submanifold of G at the point h;*

(b) *H is a submanifold of G.*

Proof: Obviously, (b) implies (a). Assume (a). Let n be the dimension of G and let m be the dimension of H at h. Then $m \leq n$. Moreover, there exists an open neighborhood U of h in G and a diffeomorphism χ of U onto an open subset of $\mathbb{R}^n$ such that $\chi(h) = 0$ and such that $\chi(U \cap H) = \chi(U) \cap (\mathbb{R}^m \times \{0\})$. Let $k \in H$. Put $a = kh^{-1}$. Then l_a is a diffeomorphism of G onto itself, mapping h onto k. We shall use this to show that H is a submanifold of dimension m at the point k. Since $a \in H$, the map l_a maps the subset H bijectively onto itself. The set $U_k := l_a(U)$ is an open neighborhood of k in G. Moreover, $\chi_k = \chi \circ l_a^{-1}$ is a diffeomorphism of U_k onto the open subset $\chi(U)$ of $\mathbb{R}^n$. Finally,

$$\chi_k(U_k \cap H) = \chi_k(l_a U \cap l_a H) = \chi_k \circ l_a(U \cap H) = \chi(U \cap H) = \chi(U) \cap (\mathbb{R}^m \times \{0\}).$$

This shows that H is a submanifold of dimension m at the point k. Since k was an arbitrary point of H, assertion (b) follows. $\square$

Example 2.14 Let V be a finite dimensional real linear space. We define the *special linear group*

$$\mathrm{SL}(V) := \{A \in \mathrm{GL}(V) \mid \det A = 1\}.$$

Note that det is a group homomorphism from $\mathrm{GL}(V)$ to $\mathbb{R}^*$. Moreover, $\mathrm{SL}(V)$ is the kernel of det. In particular, $\mathrm{SL}(V)$ is a subgroup of $\mathrm{GL}(V)$. We will show that $\mathrm{SL}(V)$ is a submanifold of $\mathrm{GL}(V)$ of codimension 1. By Lemma 2.13 it suffices to do this at the element $I = I_V$.

Since $G := \mathrm{GL}(V)$ is an open subset of the linear space $\mathrm{End}(V)$ its tangent space $T_I G$ may be identified with $\mathrm{End}(V)$. The determinant function is smooth from G to $\mathbb{R}$ hence its tangent map is a linear map from $\mathrm{End}(V)$ to $\mathbb{R}$. In Lemma 2.15 below we show that this tangent map is the trace $\mathrm{tr} : \mathrm{End}(V) \to \mathbb{R}$, $A \mapsto \mathrm{tr}(A)$. Clearly tr is a surjective linear map. This implies that det is submersive at I. By the *submersion theorem,* it follows that $\mathrm{SL}(V)$ is a smooth codimension 1 submanifold at I.

Lemma 2.15 *The function* $\det : \mathrm{GL}(V) \to \mathbb{R}^*$ *has tangent map at* I *given by* $T_I \det = \mathrm{tr} : \mathrm{End}(V) \to \mathbb{R}$, $A \mapsto \mathrm{tr} A$.

Proof: Put $G = \mathrm{GL}(V)$. In the discussion in Example 2.14 we saw that $T_I G = \mathrm{End}(V)$ and, similarly, $T_1\mathbb{R}^* = \mathbb{R}$. Thus $T_I \det$ is a linear map $\mathrm{End}(V) \to \mathbb{R}$. Let $H \in \mathrm{End}(V)$. Then by the chain rule,

$$T_I(\det)(H) = \left.\frac{d}{dt}\right|_{t=0} \det(I + tH).$$

Fix a basis $v_1, \dots, v_n$ of V. We denote the matrix coefficients of a map $A \in \mathrm{End}(V)$ with respect to this basis by A_{ij}, for $1 \le i, j \le n$. Using the definition of the determinant, we obtain

$$\det(I + tH) = 1 + t(H_{11} + \cdots + H_{nn}) + t^2 R(t, H),$$

where R is polynomial in t and the matrix coefficients H_{ij}. Differentiating this expression with respect to t and substituting $t = 0$ we obtain

$$T_I(\det)(H) = H_{11} + \cdots + H_{nn} = \mathrm{tr} H.$$

□

We shall now formulate a result that allows us to give many examples of Lie groups. The complete proof of this result will be given at a later stage. Of course we will make sure not to use the result in the development of the theory until then.

Theorem 2.16 *Let* G *be a Lie group and let* H *be a subgroup of* G. *Then the following assertions are equivalent.*

(a) H *is closed in the sense of topology.*

(b) H *is a submanifold.*

Proof: For the moment we will only prove that (b) implies (a). Assume (b). Then there exists an open neighborhood U of e in G such that $U \cap \bar{H} = U \cap H$. Let $y \in \bar{H}$. Since l_y is a diffeomorphism from G onto itself, yU is an open neighborhood of y in G, hence $yU \cap H \neq \emptyset$. Select $h \in yU \cap H$. Then $y^{-1}h \in U$. On the other hand, from $y \in \bar{H}$, $h \in H$ it follows that $y^{-1}h \in \bar{H}$. Hence, $y^{-1}h \in U \cap \bar{H} = U \cap H$, and we see that $y \in H$. We conclude that $\bar{H} \subset H$. Therefore, H is closed. □

By a *closed subgroup* of a Lie group G we mean a subgroup that is closed in the sense of topology.

Corollary 2.17 *Let G be a Lie group. Then every closed subgroup of G is a Lie group.*

Proof: Let H be a closed subgroup of G. Then H is a smooth submanifold of G, by Theorem 2.16. By Lemma 2.5 it follows that H is a Lie group. $\square$

Corollary 2.18 *Let $\varphi : G \to H$ be a homomorphism of Lie groups. Then the kernel of φ is a closed subgroup of G. In particular,* $\ker \varphi$ *is a Lie group.*

Proof: Put $K = \ker\varphi$. Then K is a subgroup of G. Now φ is continuous and $\{e_H\}$ is a closed subset of H. Hence, $K = \varphi^{-1}(\{e_H\})$ is a closed subset of G. Now apply Corollary 2.17. $\square$

Remark 2.19 We may apply the above corollary in Example 2.14 as follows. The map $\det : \mathrm{GL}(V) \to \mathbb{R}^*$ is a Lie group homomorphism. Therefore, its kernel $\mathrm{SL}(V)$ is a Lie group.

Example 2.20 Let now V be a complex linear space of finite complex dimension n. Then by $\mathrm{End}(V)$ we denote the complex linear space of complex linear maps from V to itself, and by $\mathrm{GL}(V)$ the group of invertible maps. The discussion of Examples 2.10 and 2.11 goes through with everywhere $\mathbb{R}$ replaced by $\mathbb{C}$. In particular, the determinant det is a complex polynomial map $\mathrm{End}(V) \to \mathbb{C}$, hence continuous. Since $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ is open in $\mathbb{C}$, the set $\mathrm{GL}(V) = \det^{-1}(\mathbb{C}^*)$ is open in $\mathrm{End}(V)$. As in Example 2.11 we now see that $\mathrm{GL}(V)$ is a Lie group.

The map $\det : \mathrm{GL}(V) \to \mathbb{C}^*$ is a Lie group homomorphism. Hence, by Corollary 2.18 its kernel, $\mathrm{SL}(V) := \{A \in \mathrm{GL}(V) \mid \det A = 1\}$, is a Lie group.

Finally, let $\mathbf{v} = (v_1, \ldots, v_n)$ be a basis of V (over $\mathbb{C}$). Then the associated matrix map $\mathrm{mat} = \mathrm{mat}_{\mathbf{v}}$ is a complex linear isomorphism from $\mathrm{End}(V)$ onto the space $\mathrm{M}(n, \mathbb{C})$ of complex $n \times n$ matrices. It restricts to a Lie group isomorphism $\mathrm{GL}(V) \simeq \mathrm{GL}(n, \mathbb{C})$ and to a Lie group isomorphism $\mathrm{SL}(V) \simeq \mathrm{SL}(n, \mathbb{C})$.

Another very useful application of Corollary 2.17 is the following. Let V be a finite dimensional real linear space, and let $\beta : V \times V \to W$ be a bilinear map into a finite dimensional real linear space W. For $g \in \mathrm{GL}(V)$ we define the bilinear map $g \cdot \beta : V \times V \to W$ by $g \cdot \beta(u, v) = \beta(g^{-1}u, g^{-1}v)$. From $g_1 \cdot (g_2 \cdot \beta) = (g_1 g_2) \cdot \beta$ one readily deduces that the stabilizer of β in $\mathrm{GL}(V)$,

$$\mathrm{GL}(V)_\beta = \{g \in \mathrm{GL}(V) \mid g \cdot \beta = \beta\}$$

is a subgroup of $\mathrm{GL}(V)$. Similarly $\mathrm{SL}(V)_\beta := \mathrm{SL}(V) \cap \mathrm{GL}(V)_\beta$ is a subgroup.

Lemma 2.21 *The groups* $\mathrm{GL}(V)_\beta$ *and* $\mathrm{SL}(V)_\beta$ *are closed subgroups of* $\mathrm{GL}(V)$. *In particular, they are Lie groups.*

Proof: Define $C_{u,v} = \{g \in \mathrm{GL}(V) \mid \beta(g^{-1}u, g^{-1}v) = \beta(u,v)\}$, for $u, v \in V$. Then $\mathrm{GL}(V)_\beta$ is the intersection of the sets $C_{u,v}$, for all $u, v \in V$. Thus, to establish closedness of this group, it suffices to show that each of the sets $C_{u,v}$ is closed in $\mathrm{GL}(V)$. For this, we consider the function $f : \mathrm{GL}(V) \to W$ given by $f(g) = \beta(g^{-1}u, g^{-1}v)$. Then $f = \beta \circ (\iota, \iota)$, hence f is continuous. Since $\{\beta(u,v)\}$ is a closed subset of W, it follows that $C_{u,v} = f^{-1}(\{\beta(u,v)\})$ is closed in $\mathrm{GL}(V)$. This establishes that $\mathrm{GL}(V)_\beta$ is a closed subgroup of $\mathrm{GL}(V)$. By application of Corollary 2.17 it follows that $\mathrm{GL}(V)_\beta$ is a Lie group.

Since $\mathrm{SL}(V)$ is a closed subgroup of $\mathrm{GL}(V)$ as well, it follows that $\mathrm{SL}(V)_\beta = \mathrm{SL}(V) \cap \mathrm{GL}(V)_\beta$ is a closed subgroup, hence a Lie group. □

By application of the above to particular bilinear forms, we obtain interesting Lie groups.

Example 2.22 (a) Take $V = \mathbb{R}^n$ and β the standard inner product on $\mathbb{R}^n$. Then $\mathrm{GL}(V)_\beta = \mathrm{O}(n)$, the *orthogonal group*. Moreover, $\mathrm{SL}(V)_\beta = \mathrm{SO}(n)$, the *special orthogonal group*.

Example 2.23 Let $n = p + q$, with p, q positive integers and put $V = \mathbb{R}^n$. Let β be the standard inner product of signature (p, q), i.e.,

$$\beta(x, y) = \sum_{i=1}^{p} x_i y_i - \sum_{i=p+1}^{n} x_i y_i .$$

Then $\mathrm{GL}(V)_\beta = \mathrm{O}(p, q)$ and $\mathrm{SL}(V)_\beta = \mathrm{SO}(p, q)$. In particular, we see that the *Lorentz group* $\mathrm{O}(3, 1)$ is a Lie group.

Example 2.24 Let $V = \mathbb{R}^{2n}$ and let β be the standard *symplectic form* given by

$$\beta(x, y) = \sum_{i=1}^{n} x_i y_{n+i} - \sum_{i=1}^{n} x_{n+i} y_i .$$

Then $\mathrm{GL}(V)_\beta$ is the *real symplectic group* $\mathrm{Sp}(n, \mathbb{R})$.

Example 2.25 Let V be a finite dimensional complex linear space, equipped with a complex inner product β. This inner product is not a complex bilinear form, since it is skew linear in its second component (this will always be our convention with complex inner products). However, as a map $V \times V \to \mathbb{C}$ it is bilinear over $\mathbb{R}$; in particular, it is continuous. As in the proof of Lemma 2.21 we infer that the associated *unitary group* $\mathrm{U}(V) = \mathrm{GL}(V)_\beta$ is a closed subgroup of $\mathrm{GL}(V)$, hence a Lie group. Likewise, the *special unitary group* $\mathrm{SU}(V) := \mathrm{U}(V) \cap \mathrm{SL}(V)$ is a Lie group.

Via the standard basis of $\mathbb{C}^n$ we identify $\mathrm{End}(\mathbb{C}^n) \simeq \mathrm{M}(n, \mathbb{C})$ and $\mathrm{GL}(\mathbb{C}^n) \simeq \mathrm{GL}(n, \mathbb{C})$, see also Remark 2.12. We equip $\mathbb{C}^n$ with the standard inner product given by

$$\langle z\,,\, w\rangle = \sum_{i=1}^{n} z_i \bar{w}_i \qquad (z, w \in \mathbb{C}^n).$$

The associated unitary group $\mathrm{U}(\mathbb{C}^n)$ may be identified with the group $\mathrm{U}(n)$ of unitary $n \times n$-matrices. Similarly, $\mathrm{SU}(\mathbb{C}^n)$ corresponds with the special unitary matrix group $\mathrm{SU}(n)$.

Remark 2.26 It is possible to immediately apply Lemma 2.21 in the above example, in order to conclude that $U(n)$ is closed. For this we observe that we may forget the complex structure of V and view it as a real linear space. We write $V_{(\mathbb{R})}$ for V viewed as a linear space. If $n = \dim_{\mathbb{C}} V$ and if $v_1, \ldots, v_n$ is a basis of V, then $v_1, iv_1, \ldots, v_n, iv_n$ is a basis of the real linear space $V_{(\mathbb{R})}$. In particular we see that $\dim_{\mathbb{R}} V_{(\mathbb{R})} = 2n$. Any complex linear map $T \in \mathrm{End}(V)$ may be viewed as a real linear map from V to itself, hence as an element of $\mathrm{End}(V_{(\mathbb{R})})$, which we denote by $T_{(\mathbb{R})}$. We note that $T \mapsto T_{(\mathbb{R})}$ is a real linear embedding of $\mathrm{End}(V)$ into $\mathrm{End}(V_{(\mathbb{R})})$. Accordingly we may view $\mathrm{End}(V)$ as a real linear subspace of $\mathrm{End}(V_{(\mathbb{R})})$. Let J denote multiplication by i, viewed as a real linear endomorphism of $V_{(\mathbb{R})}$. We leave it to the reader to verify that

$$\mathrm{End}(V) = \{A \in \mathrm{End}(V_{(\mathbb{R})}) \mid A \circ J = J \circ A\}.$$

Accordingly,

$$\mathrm{GL}(V) = \{a \in \mathrm{GL}(V_{(\mathbb{R})}) \mid a \circ J = J \circ A\}.$$

From this one readily deduces that $\mathrm{GL}(V)$ is a closed subgroup of $\mathrm{GL}(V_{(\mathbb{R})})$. In the situation of Example 2.25, $H := \mathrm{GL}(V_{(\mathbb{R})})_\beta$ is a closed subgroup of $\mathrm{GL}(V_{(\mathbb{R})})$, by Lemma 2.21. Hence $\mathrm{U}(V) = \mathrm{GL}(V) \cap H$ is a closed subgroup as well.

We end this section with useful descriptions of the orthogonal, unitary and symplectic groups.

Example 2.27 For a matrix $A \in \mathrm{M}(n, \mathbb{R})$ we define its transpose $A^t \in \mathrm{M}(n, \mathbb{R})$ by $(A^t)_{ij} = A_{ji}$. Let $\beta = \langle \cdot\,,\, \cdot \rangle$ be the standard inner product on $\mathbb{R}^n$. Then $\langle Ax\,,\, y\rangle = \langle x\,,\, A^t y\rangle$. Let $a \in \mathrm{GL}(n, \mathbb{R})$. Then for all $x, y \in \mathbb{R}^n$,

$$a^{-1} \cdot \beta(x, y) = \langle ax\,,\, ay\rangle = \langle a^t ax\,,\, y\rangle.$$

Since $\mathrm{O}(n) = \mathrm{GL}(n, \mathbb{R})_\beta$, we infer that

$$\mathrm{O}(n) = \{a \in \mathrm{GL}(n, \mathbb{R}) \mid a^t a = I\}.$$

Example 2.28 If $A \in \mathrm{M}(n, \mathbb{C})$ we denote its complex adjoint by $(A^*)_{ij} = \bar{A}_{ji}$. Let $\langle \cdot\,,\, \cdot\rangle$ be the complex standard inner product on $\mathbb{C}^n$. Then $\langle Ax\,,\, y\rangle = \langle x\,,\, A^* y\rangle$ for all $x, y \in \mathbb{C}^n$. As in the previous example we now deduce that

$$\mathrm{U}(n) = \{a \in \mathrm{GL}(n, \mathbb{C}) \mid a^* a = I\}.$$

Example 2.29 Let β be the standard symplectic form on $\mathbb{R}^{2n}$, see Example 2.24. Let $J \in \mathrm{M}(2n, \mathbb{R})$ be defined by

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

where the indicated blocks are of size $n \times n$.

Let $\langle \cdot\,,\, \cdot\rangle$ denote the standard inner product on $\mathbb{R}^{2n}$. Then for all $x, y \in \mathbb{R}^{2n}$, we have $\beta(x, y) = \langle x\,,\, Jy\rangle$. Let $a \in \mathrm{GL}(n, \mathbb{R})$, then

$$a^{-1} \cdot \beta(x, y) = \langle ax\,,\, Jay\rangle = \langle x\,,\, a^t Jay\rangle.$$

From this we see that $\mathrm{Sp}(n,\mathbb{R}) = \mathrm{GL}(2n,\mathbb{R})_\beta$ consists of all $a \in \mathrm{GL}(2n,\mathbb{R})$ with $a^t J a = J$, or, equivalently, with

$$(a^t)^{-1} = J a J^{-1} \tag{1}$$

This description motivates the following definition. The map $A \mapsto A^t$ uniquely extends to a complex linear endomorphism of $\mathrm{M}(2n,\mathbb{C})$. This extension is given by the usual formula $(A^t)_{ij} = A_{ji}$. We now define $\mathrm{Sp}(n,\mathbb{C})$ to be the collection of $a \in \mathrm{GL}(2n,\mathbb{C})$ satisfying condition (1). One readily verifies that $\mathrm{Sp}(n,\mathbb{C})$ is a closed subgroup of $\mathrm{GL}(2n,\mathbb{C})$ hence a Lie group. We call it the *complex symplectic group*.

Note that $\mathrm{GL}(2n,\mathbb{R})$ is a closed subgroup of $\mathrm{GL}(2n,\mathbb{C})$ and that $\mathrm{Sp}(n,\mathbb{R}) = \mathrm{GL}(2n,\mathbb{R}) \cap \mathrm{Sp}(n,\mathbb{C})$.

Finally, we define the *compact symplectic group* by

$$\mathrm{Sp}(n) := \mathrm{U}(2n) \cap \mathrm{Sp}(n,\mathbb{C}).$$

Clearly, this is a closed subgroup of $\mathrm{GL}(2n,\mathbb{C})$, hence a Lie group.

Remark 2.30 In this section we have frequently used the following principle. If G is a Lie group, and if $H, K \subset G$ are closed subgroups, then $H \cap K$ is a closed subgroup, hence a Lie group.

3 Invariant vector fields and the exponential map

If M is a manifold, we denote by $\mathcal{V}(M)$ the real linear space of smooth *vector fields* on M. A vector field $v \in \mathcal{V}(G)$ is called *left invariant*, if $(l_x)_* v = v$ for all $x \in G$, or, equivalently if

$$v(xy) = T_y(l_x)\, v(y) \qquad (x, y \in G). \tag{2}$$

The collection of smooth left invariant vector fields is a linear subspace of $\mathcal{V}(G)$, which we denote by $\mathcal{V}_L(G)$. From the above equation with $y = e$ we see that a left invariant vector field is completely determined by its value $v(e) \in T_e G$ at e. Differently said, $v \mapsto v(e)$ defines an injective linear map from $\mathcal{V}_L(G)$ into $T_e G$. The next result asserts that this map is surjective as well. If $X \in T_e G$, we define the vector field v_X on G by

$$v_X(x) = T_e(l_x) X, \qquad (x \in G). \tag{3}$$

Lemma 3.1 *The map $X \mapsto v_X$ defines a linear isomorphism from $T_e G$ onto $\mathcal{V}_L(G)$. Its inverse is given by $v \mapsto v(e)$.*

Proof: From the fact that $(x, y) \mapsto l_x(y)$ is a smooth map $G \times G \to G$, it follows by differentiation with respect to y at $y = e$ in the direction of $X \in T_e G$ that $x \mapsto T_e(l_x)X$ is smooth as a map $G \to TG$. This implies that v_X is a smooth vector field on G. Hence $X \mapsto v_X$ defines a real linear map $T_e G \to \mathcal{V}(G)$. We claim that it maps into $\mathcal{V}_L(G)$.

Fix $X \in T_e G$. Differentiating the relation $l_{xy} = l_x \circ l_y$ and applying the chain rule we see that $T_e(l_{xy}) = T_y(l_x) T_e(l_y)$. Applying this to the definition of v_X we see that v_X satisfies (2), hence is left invariant. This establishes the claim.

From $v_X(e) = X$ we see that the map $\epsilon : v \mapsto v(e)$ from $\mathcal{V}_L(G)$ to $T_e G$ is not only injective, but also surjective. Thus, ϵ is a linear isomorphism, with inverse $X \mapsto v_X$. $\square$

If $X \in T_eG$, we define α_X to be the *maximal integral curve* of v_X with initial point e.

Lemma 3.2 *Let $X \in T_eG$. Then the integral curve α_X has domain $\mathbb{R}$. Moreover, we have $\alpha_X(s+t) = \alpha_X(s)\alpha_X(t)$ for all $s, t \in \mathbb{R}$. Finally the map $(t, X) \mapsto \alpha_X(t),\ \mathbb{R} \times T_eG \to G$ is smooth.*

Proof: Let α be any integral curve for v_X, let $y \in G$, and put $\alpha_1(t) = y\alpha(t)$. Differentiating this relation with respect to t we obtain:

$$\frac{d}{dt}\alpha_1(t) = T_{\alpha(t)}l_y \frac{d}{dt}\alpha(t) = T_{\alpha(t)}l_y v_X(\alpha(t)) = v_X(\alpha_1(t)),$$

by left invariance of v_X. Hence α_1 is an integral curve for v_X as well.

Let now I be the domain of α_X, fix $t_1 \in I$, and put $x_1 = \alpha_X(t_1)$. Then $\alpha_1(t) := x_1\alpha_X(t)$ is an integral curve for v_X with starting point x_1 and domain I. On the other hand, the maximal integral curve for v_X with starting point x_1 is given by $\alpha_2 : t \mapsto \alpha_X(t+t_1)$. It has domain $I - t_1$. We infer that $I \subset I - t_1$. It follows that $s + t_1 \in I$ for all $s, t_1 \in I$. Hence, $I = \mathbb{R}$.

Fix $s \in \mathbb{R}$, then by what we saw above $c : t \mapsto \alpha_X(s)\alpha_X(t)$ is the maximal integral curve for v_X with initial pont $\alpha_X(s)$. On the other hand, the same holds for $d : t \mapsto \alpha_X(s+t)$. Hence, by uniqueness of the maximal integral curve, $c = d$.

The final assertion is a consequence of the fact that the vector field v_X depends linearly, hence smoothly on the parameter X. Let φ_X denote the flow of v_X. Then it is a well known (local) result that the map $(X, t, x) \mapsto \varphi_X(t, x)$ is smooth. In particular, $(t, X) \mapsto \alpha_X(t) = \varphi_X(t, e)$ is a smooth map $\mathbb{R} \times T_eG \to G$. $\square$

Definition 3.3 Let G be a Lie group. The *exponential map* $\exp = \exp_G : T_eG \to G$ is defined by

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

where α_X is defined as above; i.e., α_X is the maximal integral curve with initial point e of the left invariant vector field v_X on G determined by $v_X(e) = X$.

Example 3.4 We return to the example of the group $\mathrm{GL}(V)$, with V a finite dimensional real linear space. Its neutral element e equals $I = I_V$. Since $\mathrm{GL}(V)$ is open in $\mathrm{End}(V)$, we have $T_e\mathrm{GL}(V) = \mathrm{End}(V)$. If $x \in \mathrm{GL}(V)$, then l_x is the restriction of the linear map $L_x : A \mapsto xA,\ \mathrm{End}(V) \to \mathrm{End}(V)$, to $\mathrm{GL}(V)$, hence $T_e(l_x) = L_x$, and we see that for $X \in \mathrm{End}(V)$ the invariant vectorfield v_X is given by $v_X(x) = xX$. Hence, the integral curve α_X satisfies the equation:

$$\frac{d}{dt}\alpha(t) = \alpha(t)X.$$

Since $t \mapsto e^{tX}$ is a solution to this equation with the same initial value, we must have that $\alpha_X(t) = e^{tX}$. Thus in this case exp is the ordinary exponential map $X \mapsto e^X,\ \mathrm{End}(V) \to \mathrm{GL}(V)$.

Remark 3.5 In the above example we have used the exponential e^A of an endomorphism $A \in \mathrm{End}(V)$. One way to define this exponential is precisely by the method of differential equations just described. Another way is to introduce it by its power series

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n.$$

From the theory of power series it follows that $A \to e^A$ is a smooth map $\mathrm{End}(V) \to \mathrm{End}(V)$. Moreover,

$$\frac{d}{dt} e^{tA} = Ae^{tA} = e^{tA}A,$$

by termwise differentiation of power series. By multiplication of power series we obtain

$$e^X e^Y = e^{X+Y} \quad \text{if} \quad X, Y \in \mathrm{End}(V) \quad \text{commute, i.e.,} \quad XY = YX. \tag{4}$$

Applying this with $X = sA$ and $Y = tA$, we obtain $e^{(s+t)A} = e^{sA}e^{tA}$, for all $A \in \mathrm{End}(V)$ and $s, t \in \mathbb{R}$. This formula will be established in general in Lemma 3.6 (b) below.

Lemma 3.6 *For all* $s, t \in \mathbb{R}$, $X \in T_eG$ *we have*

(a) $\exp(sX) = \alpha_X(s)$.

(b) $\exp(s + t)X = \exp sX \ \exp tX$.

Moreover, the map $\exp : T_eG \to G$ *is smooth and a local diffeomorphism at* 0. *Its tangent map at the origin is given by* $T_0 \exp = I_{T_eG}$.

Proof: Consider the curve $c(t) = \alpha_X(st)$. Then $c(0) = e$, and

$$\frac{d}{dt} c(t) = s\dot{\alpha}_X(st) = s\, v_X(\alpha_X(st)) = v_{sX}(c(t)).$$

Hence c is the maximal integral curve of v_{sX} with initial point e, and we conclude that $c(t) = \alpha_{sX}(t)$. Now evaluate at $t = 1$ to obtain the equality.

Formula (b) is an immediate consequence of (a) and Lemma 3.2. Finally, from Lemma 3.2 we have that $(t, X) \mapsto \alpha_X(t)$ is a smooth map $\mathbb{R} \times T_eG \to G$. Substituting $t = 1$ we obtain smoothness of $\exp$. Moreover,

$$T_0(\exp)X = \frac{d}{dt} \exp(tX)|_{t=0} = \dot{\alpha}_X(0) = v_X(e) = X.$$

Hence $T_0(\exp) = I_{T_eX}$, and from the *inverse function theorem* it follows that exp is a local diffeomorphism at 0, i.e., there exists an open neighborhood U of 0 in T_eG such that exp maps U diffeomorphically onto an open neighborhood of e in G. $\square$

Definition 3.7 A smooth group homomorphism $\alpha : (\mathbb{R}, +) \to G$ is called a *one-parameter subgroup* of G.

Lemma 3.8 *If $X \in T_e G$, then $t \mapsto \exp tX$ is a one-parameter subgroup of G. Moreover, all one-parameter subgroups are obtained in this way. More precisely, let α be a one-parameter subgroup in G, and put $X = \dot\alpha(0)$. Then $\alpha(t) = \exp(tX)$ $(t \in \mathbb{R})$.*

Proof: The first assertion follows from Lemma 3.2. Let $\alpha : \mathbb{R} \to G$ be a one-parameter subgroup. Then $\alpha(0) = e$, and

$$\frac{d}{dt}\alpha(t) = \frac{d}{ds}\alpha(t+s)|_{s=0} = \frac{d}{ds}\alpha(t)\alpha(s)|_{s=0} = T_e(l_{\alpha(t)})\dot\alpha(0) = v_X(\alpha(t)),$$

hence α is an integral curve for v_X with initial point e. Hence $\alpha = \alpha_X$ by the uniqueness of integral curves. Now apply Lemma 3.6. $\square$

We now come to a very important application.

Lemma 3.9 *Let $\varphi : G \to H$ be a homomorphism of Lie groups. Then the following diagram commutes:*

$$\begin{array}{ccc} G & \xrightarrow{\varphi} & H \\ {\scriptstyle \exp_G}\uparrow & & \uparrow{\scriptstyle \exp_H} \\ T_e G & \xrightarrow{T_e\varphi} & T_e H \end{array}$$

Proof: Let $X \in T_e G$. Then $\alpha(t) = \varphi(\exp_G(tX))$ is a one-parameter subgroup of H. Differentiating at $t = 0$ we obtain $\dot\alpha(0) = T_e(\varphi)T_0(\exp_G)X = T_e(\varphi)X$. Now apply the above lemma to conclude that $\alpha(t) = \exp_H(tT_e(\varphi)X)$. The result follows by specializing to $t = 1$. $\square$

4 The Lie algebra of a Lie group

In this section we assume that G is a Lie group. If $x \in G$ then the translation maps $l_x : y \mapsto xy$ and $r_x : y \mapsto yx$ are diffeomorphisms from G onto itself. Therefore, so is the conjugation map $\mathcal{C}_x = l_x \circ r_x^{-1} : y \mapsto xyx^{-1}$. The latter map fixes the neutral element e; therefore, its tangent map at e is a linear automorphism of $T_e G$. Thus, $T_e\mathcal{C}_x \in \mathrm{GL}(T_e G)$.

Definition 4.1 If $x \in G$ we define $\mathrm{Ad}(x) \in \mathrm{GL}(T_e G)$ by $\mathrm{Ad}(x) := T_e\mathcal{C}_x$. The map $\mathrm{Ad} : G \to \mathrm{GL}(T_e G)$ is called the *adjoint representation* of G in $T_e G$.

Example 4.2 We return to the example of $\mathrm{GL}(V)$, with V a finite dimensional real linear space. Since $\mathrm{GL}(V)$ is an open subset of the linear space $\mathrm{End}(V)$ we may identify its tangent space at I with $\mathrm{End}(V)$. If $x \in \mathrm{GL}(V)$, then $\mathcal{C}_x$ is the restriction of the linear map $C_x : A \mapsto xAx^{-1}$, $\mathrm{End}(V) \to \mathrm{End}(V)$. Hence $\mathrm{Ad}(x) = T_e(\mathcal{C}_x) = C_x$ is conjugation by x.

The above example suggests that $\mathrm{Ad}(x)$ should be looked at as an action of x on T_eG by conjugation. The following result is consistent with this point of view.

Lemma 4.3 *Let $x \in G$, then for every $X \in T_eG$ we have*

$$x \exp X\, x^{-1} = \exp(\mathrm{Ad}(x)\, X).$$

Proof: We note that $\mathcal{C}_x : G \to G$ is a Lie group homomorphism. Hence we may apply Lemma 3.9 with $H = G$ and $\varphi = \mathcal{C}_x$. Since $T_e\mathcal{C}_x = \mathrm{Ad}(x)$, we see that the following diagram commutes:

$$\begin{tikzcd} G \arrow[r, "\mathcal{C}_x"] & G \\ T_eG \arrow[u, "\exp"] \arrow[r, "\mathrm{Ad}(x)"] & T_eG \arrow[u, "\exp"'] \end{tikzcd}$$

The result follows. $\square$

Lemma 4.4 *The map $\mathrm{Ad} : G \to \mathrm{GL}(T_eG)$ is a Lie group homomorphism.*

Proof: From the fact that $(x, y) \mapsto xyx^{-1}$ is a smooth map $G \times G \to G$ it follows by differentiation with respect to y at $y = e$ that $x \mapsto \mathrm{Ad}(x)$ is a smooth map from G to $\mathrm{End}(T_eG)$. Since $\mathrm{GL}(V)$ is open in $\mathrm{End}(T_eG)$ it follows that $\mathrm{Ad} : G \to \mathrm{GL}(T_eG)$ is smooth.

From $\mathcal{C}_e = I_G$ it follows that $\mathrm{Ad}(e) = I_{T_eG}$. Moreover, differentiating the relation $\mathcal{C}_{xy} = \mathcal{C}_x\mathcal{C}_y$ at e, we find, by application of the chain rule, that $\mathrm{Ad}(xy) = \mathrm{Ad}(x)\mathrm{Ad}(y)$ for all $x, y \in G$. $\square$

Since $\mathrm{Ad}(e) = I = I_{T_eG}$ and $T_I\mathrm{GL}(T_eG) = \mathrm{End}(T_eG)$, we see that the tangent map of Ad at e is a linear map $T_eG \to \mathrm{End}(T_eG)$.

Definition 4.5 The linear map $\mathrm{ad} : T_eG \to \mathrm{End}(T_eG)$ is defined by

$$\mathrm{ad} := T_e\mathrm{Ad}.$$

We note that, by the chain rule, for all $X \in T_eG$,

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

Lemma 4.6 *For all $X \in T_eG$ we have:*

$$\mathrm{Ad}(\exp X) = e^{\,\mathrm{ad} X}.$$

Proof: In view of Lemma 4.4, we may apply Lemma 3.9 with $H = \mathrm{GL}(T_eG)$ and $\varphi = \mathrm{Ad}$. Since $T_eH = T_I\mathrm{GL}(T_eG) = \mathrm{End}(T_eG)$, whereas $\exp_H$ is given by $X \mapsto e^X$, we see that the following diagram commutes:

$$\begin{tikzcd} G \arrow[r, "\mathrm{Ad}"] & \mathrm{GL}(T_eG) \\ T_eG \arrow[u, "\exp"] \arrow[r, "\mathrm{ad}"] & \mathrm{End}(T_eG) \arrow[u, "e^{(\cdot)}"'] \end{tikzcd}$$

The result follows. $\square$

Example 4.7 Let V be finite dimensional real linear space. Then for $x \in \mathrm{GL}(V)$ the linear map $\mathrm{Ad}(x) : \mathrm{End}(V) \to \mathrm{End}(V)$ is given by $\mathrm{Ad}(x)Y = xYx^{-1}$. Substituting $x = e^{tX}$ and differentiating the resulting expression with respect to t at $t = 0$ we obtain:

$$(\mathrm{ad}X)Y = \frac{d}{dt}[e^{tX}Ye^{-tX}]_{t=0} = XY - YX.$$

Hence in this case $(\mathrm{ad}X)Y$ is the commutator bracket of X and Y.

Motivated by the above example we introduce the following notation.

Definition 4.8 For $X, Y \in T_eG$ we define the Lie bracket $[X, Y] \in T_eG$ by

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

Lemma 4.9 *The map $(X, Y) \mapsto [X, Y]$ is bilinear $T_eG \times T_eG \to T_eG$. Moreover, it is* anti-symmetric, *i.e.,*

$$[X, Y] = -[Y, X] \qquad (X, Y \in T_eG).$$

Proof: The bilinearity is an immediate consequence of the fact that $\mathrm{ad} : T_eG \to \mathrm{End}(T_eG)$ is linear. Let $Z \in T_eG$. Then for all $s, t \in \mathbb{R}$ we have

$$\exp(tZ) = \exp(sZ)\exp(tZ)\exp(-sZ) = \exp(t\mathrm{Ad}(\exp(sZ))\,Z),$$

by Lemmas 3.6 and 4.3. Differentiating this relation with respect to t at $t = 0$ we obtain:

$$Z = \mathrm{Ad}(\exp(sZ))\,Z \qquad (s \in \mathbb{R}).$$

Differentiating this with respect to s at $s = 0$ we obtain:

$$0 = \mathrm{ad}(Z)T_0(\exp)Z = \mathrm{ad}(Z)Z = [Z, Z].$$

Now substitute $Z = X + Y$ and use the bilinarity to arrive at the desired conclusion. $\square$

Lemma 4.10 *Let $\varphi : G \to H$ be a homomorphism of Lie groups. Then*

$$T_e\varphi([X, Y]_G) = [T_e\varphi X, T_e\varphi Y]_H, \qquad (X, Y \in T_eG). \tag{5}$$

Proof: One readily verifies that $\varphi \circ \mathcal{C}^G_x = \mathcal{C}^H_{\varphi(x)} \circ \varphi$. Taking the tangent map of both sides of this equation at e, we obtain that the following diagram commutes:

$$\begin{array}{ccc} T_eG & \xrightarrow{T_e\varphi} & T_eH \\ {\scriptstyle \mathrm{Ad}_G(x)}\uparrow & & \uparrow{\scriptstyle \mathrm{Ad}_H(\varphi(x))} \\ T_eG & \xrightarrow{T_e\varphi} & T_eH \end{array}$$

Differentiating once more at $x = e$, in the direction of $X \in T_eG$, we obtain that the following diagram commutes:

$$\begin{array}{ccc} T_eG & \xrightarrow{T_e\varphi} & T_eH \\ {\scriptstyle \mathrm{ad}_G(X)}\uparrow & & \uparrow{\scriptstyle \mathrm{ad}_H(T_e\varphi X)} \\ T_eG & \xrightarrow{T_e\varphi} & T_eH \end{array}$$

We now agree to write $[X, Y] = \mathrm{ad}(X)Y$. Then by applying $T_e\varphi \circ \mathrm{ad}_G X$ to $Y \in T_eG$ the commutativity of the above diagram yields (5). $\square$

Corollary 4.11 *For all $X, Y, Z \in T_eG$,*

$$[[X, Y], Z] = [X, [Y, Z]] - [Y, [X, Z]]. \tag{6}$$

Proof: Put $\varphi = \mathrm{Ad}$ and $H = \mathrm{GL}(T_eG)$. Then $e_H = I$ and $T_I H = \mathrm{End}(T_eG)$. Moreover, $[A, B]_H = AB - BA$ for all $A, B \in \mathrm{End}(T_eG)$. Applying Lemma 4.10 and using that $[\cdot\,,\,\cdot]_G = [\cdot\,,\,\cdot]$ and $T_e\varphi = \mathrm{ad}$, we obtain

$$\mathrm{ad}([X, Y]) = [\,\mathrm{ad}X,\ \mathrm{ad}Y]_H = \mathrm{ad}X\,\mathrm{ad}Y - \mathrm{ad}Y\,\mathrm{ad}X.$$

Applying the latter relation to $Z \in T_eG$, we obtain (6). □

Definition 4.12 A real *Lie algebra* is a real linear space $\mathfrak{a}$ equipped with a bilinear map $[\cdot, \cdot] : \mathfrak{a} \times \mathfrak{a} \to \mathfrak{a}$, such that for all $X, Y, Z \in \mathfrak{a}$ we have:

(a) $[X, Y] = -[Y, X]$ (anti-symmetry);

(b) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ (Jacobi identity).

Remark 4.13 Note that condition (a) may be replaced by the equivalent condition (a'): $[X, X] = 0$ for all $X \in \mathfrak{a}$. In view of the anti-symmetry (a), condition (b) may be replaced by the equivalent condition (6). We leave it to the reader to check that another equivalent form of the Jacobi identity is given by the Leibniz type rule

$$[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]]. \tag{7}$$

Corollary 4.14 *Let G be a Lie group. Then T_eG equipped with the bilinear map $(X, Y) \mapsto [X, Y] := (\,\mathrm{ad}X)Y$ is a Lie algebra.*

Proof: The anti-linearity was established in Lemma 4.9. The Jacobi identity follows from (6) combined with the anti-linearity. □

Definition 4.15 Let $\mathfrak{a}, \mathfrak{b}$ be Lie algebras. A *Lie algebra homomorphism* from $\mathfrak{a}$ to $\mathfrak{b}$ is a linear map $\varphi : \mathfrak{a} \to \mathfrak{b}$ such that

$$\varphi([X, Y]_{\mathfrak{a}}) = [\varphi(X), \varphi(Y)]_{\mathfrak{b}},$$

for all $X, Y \in \mathfrak{a}$.

From now on we will adopt the convention that Roman capitals denote Lie groups. The corresponding Gothic lower case letters will denote the associated Lie algebras. If $\varphi : G \to H$ is a Lie group homomorphism then the associated tangent map $T_e\varphi$ will be denoted by φ_*. We now have the following.

Lemma 4.16 *Let $\varphi : G \to H$ be a homomorphism of Lie groups. Then the associated tangent map $\varphi_* : \mathfrak{g} \to \mathfrak{h}$ is a homomorphism of Lie algebras. Moreover, the following diagram commutes:*

$$\begin{array}{ccc} G & \xrightarrow{\varphi} & H \\ {\scriptstyle \exp_G}\uparrow & & \uparrow{\scriptstyle \exp_H} \\ \mathfrak{g} & \xrightarrow{\varphi_*} & \mathfrak{h} \end{array}$$

Proof: The first assertion follows from Lemma 4.10, the second from Lemma 3.9. □

Example 4.17 We consider the Lie group $G = \mathbb{R}^n$. Its Lie algebra $\mathfrak{g} = T_0\mathbb{R}^n$ may be identified with $\mathbb{R}^n$. From the fact that G is commutative, it follows that $\mathcal{C}_x = I_G$, for all $x \in G$. Hence, $\mathrm{Ad}(x) = I_{\mathfrak{g}}$, for all $x \in G$. It follows that $\mathrm{ad}(X) = 0$ for all $X \in \mathfrak{g}$. Hence $[X, Y] = 0$ for all $X, Y \in \mathfrak{g}$.

Let $X \in \mathfrak{g} \simeq \mathbb{R}^n$. Then the associated one-parameter subgroup α_X is given by $\alpha_X(t) = tX$. Hence $\exp(X) = X$, for all $X \in \mathfrak{g}$.

We consider the Lie group homomorphism $\varphi = (\varphi_1, \dots, \varphi_n) : \mathbb{R}^n \to \mathbb{T}^n$ given by $\varphi_j(x) = e^{2\pi i x_j}$. One readily verifies that φ is a local diffeomorphism. Its kernel equals $\mathbb{Z}^n$. Hence, by the isomorphism theorem for groups, the map φ factors through an isomorphism of groups $\bar{\varphi} : \mathbb{R}^n/\mathbb{Z}^n \to \mathbb{T}^n$. Via this isomorphism we transfer the manifold structure of $\mathbb{T}^n$ to a manifold structure on $\mathbb{R}^n/\mathbb{Z}^n$. Thus, $\mathbb{R}^n/\mathbb{Z}^n$ becomes a Lie group, and $\bar{\varphi}$ an isomorphism of Lie groups. Note that the manifold structure on $H := \mathbb{R}^n/\mathbb{Z}^n$ is the unique manifold structure for which the canonical projection $\pi : \mathbb{R}^n \to \mathbb{R}^n/\mathbb{Z}^n$ is a local diffeomorphism. The projection π is a Lie group homomorphism. The associated homomorphism of Lie algebras $\pi_* : \mathfrak{g} \to \mathfrak{h}$ is bijective, since π is a local diffeomorphism. Hence, π_* is an isomorphism of Lie algebras. We adopt the convention to identify $\mathfrak{h}$ with $\mathfrak{g} \simeq \mathbb{R}^n$ via π_*. It then follows from Lemma 4.16 that the exponential map $\exp_H : \mathbb{R}^n \to H = \mathbb{R}^n/\mathbb{Z}^n$ is given by $\exp_H(X) = \pi(X) = X + \mathbb{Z}^n$.

5 Commuting elements

In the following we assume that G is a Lie group with Lie algebra $\mathfrak{g}$. Two elements $X, Y \in \mathfrak{g}$ are said to *commute* if $[X, Y] = 0$. The Lie algebra $\mathfrak{g}$ is called *commutative* if every pair of elements $X, Y \in \mathfrak{g}$ commutes.

Example 5.1 If $G = \mathrm{GL}(V)$, with V a finite dimensional real or complex linear space, then $\mathfrak{g} = \mathrm{End}(V)$. In this case the Lie bracket of two elements $A, B \in \mathrm{End}(V)$ equals the commutator bracket $[A, B] = AB - BA$. Hence, the assertion that A and B commute means that $AB = BA$, as we are used to. In this case we know that the associated exponentials e^A and e^B commute as linear maps, hence as elements of G; moreover, $e^A e^B = e^{A+B}$. The following lemma generalizes this fact to arbitrary Lie algebras.

Lemma 5.2 *Let $X, Y \in \mathfrak{g}$ be commuting elements. Then the elements* $\exp X$ *and* $\exp Y$ *of G commute. Moreover,*

$$\exp(X + Y) = \exp X \exp Y.$$

Proof: We will first show that $x = \exp X$ and $y = \exp Y$ commute. For this we observe that, by Lemma 4.3, $xyx^{-1} = \exp(\mathrm{Ad}(x)Y)$. Now $\mathrm{Ad}(x)Y = e^{\,\mathrm{ad}X}Y$, by Lemma 4.6. Since $\mathrm{ad}(X)Y = [X, Y] = 0$, it follows that $\mathrm{ad}(X)^n Y = 0$ for all $n \geq 1$. Hence, $\mathrm{Ad}(x)Y = e^{\,\mathrm{ad}X}Y = Y$. Therefore, $xyx^{-1} = y$ and we see that x and y commute.

For every $s, t \in \mathbb{R}$ we have that $[sX, tY] = st[X, Y] = 0$. Hence by the first part of this proof the elements $\exp(sX)$ and $\exp(tY)$ commute for all $s, t \in \mathbb{R}$. Define the map $\alpha : \mathbb{R} \to G$ by

$$\alpha(t) = \exp(tX)\exp(tY) \qquad (t \in \mathbb{R}).$$

Then $\alpha(0) = e$. Moreover, for $s, t \in \mathbb{R}$ we have

$$\begin{aligned}\alpha(s+t) &= \exp(s+t)X \exp(s+t)Y \\ &= \exp sX \exp tX \exp sY \exp tY \\ &= \exp sX \exp sY \exp tX \exp tY = \alpha(s)\alpha(t).\end{aligned}$$

It follows that α is a one-parameter subgroup of G. Hence $\alpha = \alpha_Z$ with $Z = \alpha'(0)$, by Lemma 3.8. Now, by Lemma 5.3 below,

$$\alpha'(0) = \left(\frac{d}{dt}\right)_{t=0} \exp(tX)\exp(0) + \left(\frac{d}{dt}\right)_{t=0} \exp(0)\exp(tY) = X + Y.$$

From this it follows that $\alpha(t) = \alpha_Z(t) = \exp(tZ) = \exp(t(X+Y))$, for $t \in \mathbb{R}$. The desired equality follows by substituting $t = 1$. $\square$

The following lemma gives a form of the chain rule for differentiation that has been used in the above, and will often be useful to us.

Lemma 5.3 *Let M be a smooth manifold, U a neighborhood of $(0,0)$ in $\mathbb{R}^2$ and $\varphi : U \to M$ a map that is differentiable at $(0,0)$. Then*

$$\left(\frac{d}{dt}\right)_{t=0} \varphi(t,t) = \left(\frac{d}{dt}\right)_{t=0} \varphi(t,0) + \left(\frac{d}{dt}\right)_{t=0} \varphi(0,t).$$

Proof: Let $D_1\varphi(0,0)$ denote the tangent map of $s \mapsto \varphi(s,0)$ at zero. Similarly, let $D_2\varphi(0,0)$ denote the tangent map of $s \mapsto \varphi(0,s)$ at zero. Then the tangent $T_{(0,0)}\varphi : \mathbb{R}^2 \to T_{\varphi(0,0)}M$ of φ at the origin is given by $T_{(0,0)}\varphi(X,Y) = D_1\varphi(0,0)X + D_2\varphi(0,0)Y$, for $(X,Y) \in \mathbb{R}^2$.

Let $d : \mathbb{R} \to \mathbb{R}^2$ be defined by $d(t) = (t,t)$. Then the tangent map of d at 0 is given by $T_0 d : \mathbb{R} \mapsto \mathbb{R}^2$, $X \mapsto (X,X)$. By application of the chain rule, it follows that

$$\begin{aligned}(d/dt)_{t=0}\, \varphi(t,t) &= (d/dt)_{t=0}\, \varphi(d(t)) = T_0(\varphi \circ d)\, 1 \\ &= [T_{(0,0)}\varphi \circ T_0 d]\, 1 = [T_{(0,0)}\varphi](1,1) \\ &= D_1\varphi(0,0)1 + D_2\varphi(0,0)1 \\ &= (d/dt)_{t=0}\, \varphi(t,0) + (d/dt)_{t=0}\, \varphi(0,t).\end{aligned}$$

$\square$

Definition 5.4 The subgroup G_e generated by the elements $\exp X$, for $X \in \mathfrak{g}$, is called the *component of the identity* of G.

Remark 5.5 From this definition it follows that

$$G_e = \{\exp(X_1)\cdots\exp(X_k) \mid k \geq 1,\ X_1, \ldots, X_k \in \mathfrak{g}\}.$$

In general it is not true that $G_e = \exp(\mathfrak{g})$. Nevertheless, many properties of $\mathfrak{g}$ can be lifted to analogous properties of G_e. As we will see in this section, this is in particular true for the property of commutativity.

By an *open subgroup* of a Lie group G we mean a subgroup H of G that is an open subset of G in the sense of topology.

Lemma 5.6 *G_e is an open subgroup of G.*

Proof: Let $a \in G_e$. Then there exists a positive integer $k \geq 1$ and elements $X_1, \dots, X_k \in \mathfrak{g}$ such that $a = \exp(X_1) \dots \exp(X_k)$. The map $\exp : \mathfrak{g} \to G$ is a local diffeomorphism at 0 hence there exists an open neighborhood Ω of 0 in $\mathfrak{g}$ such that exp is a diffeomorphism of Ω onto an open subset of G. Since l_a is a diffeomorphism, it follows that $l_a(\exp(\Omega))$ is an open neighborhood of a. We now observe that $l_a(\exp(\Omega)) = \{\exp(X_1) \dots \exp(X_k) \exp(X) \mid X \in \Omega\} \subset G_e$. Hence a is an interior point of G_e. It follows that G_e is open in G. $\square$

Lemma 5.7 *Let H be an open subgroup of G. Then H is closed as well.*

Proof: For all $x, y \in G$ we have $xH = yH$ or $xH \cap yH = \emptyset$. Hence there exists a subset $S \subset G$ such that G is the disjoint union of the sets $sH,\ s \in S$. The complement of H in G is the disjoint union of the sets sH with $s \in S, s \notin H$. Being the union of open sets, this complement is open. Hence H is closed. $\square$

Lemma 5.8 *G_e equals the connected component of G containing e. In particular, G is connected if and only if $G_e = G$.*

Proof: The set G_e is open and closed in G, hence a (disjoint) union of connected components. On the other hand G_e is arcwise connected. For let $a \in G_e$, then we may write $a = \exp(X_1) \dots \exp(X_k)$ with $k \geq 1$ and $X_1, \dots, X_k \in \mathfrak{g}$. It follows that $c : [0, 1] \to G,\ t \mapsto \exp(tX_1) \dots \exp(tX_k)$ is a continuous curve with initial point $c(0) = e$ and end point $c(1) = a$. This establishes that G_e is arcwise connected, hence connected. Therefore G_e is a connected component; it obviously contains e. $\square$

Lemma 5.9 *Let G be a Lie group, $x \in G$. Then the following assertions are equivalent.*

(a) *x commutes with G_e;*

(b) $\mathrm{Ad}(x) = I$.

Proof: Assume (a). Then for every $Y \in \mathfrak{g}$ and $t \in \mathbb{R}$ we have $\exp(tY) \in G_e$, hence

$$\exp(t\mathrm{Ad}(x)Y) = x \exp tYx^{-1} = \exp tY$$

Differentiating this expression at $t = 0$ we see that $\mathrm{Ad}(x)Y = Y$. This holds for any $Y \in \mathfrak{g}$, hence (b).

For the converse implication, assume (b). If $Y \in \mathfrak{g}$, then

$$x \exp Yx^{-1} = \exp \mathrm{Ad}(x)Y = \exp Y.$$

Hence x commutes with $\exp(\mathfrak{g})$. Since the latter set generates the subgroup G_e, it follows that x commutes with G_e. $\square$

Remark 5.10 Note that the point of the above proof is that one does not need $\exp : \mathfrak{g} \to G$ to be surjective in order to derive properties of a connected Lie group G from properties of its Lie algebra. It is often enough that G is *generated* by $\exp \mathfrak{g}$. Another instance of this principle is given by the following theorem.

Theorem 5.11 *Let G be a Lie group. The following conditions are equivalent.*

(a) *The Lie algebra $\mathfrak{g}$ is commutative.*

(b) *The group G_e is commutative.*

In particular, if G is connected then $\mathfrak{g}$ is commutative if and only if G is commutative.

Proof: Assume (a). Then $[X, Y] = 0$ for all $X, Y \in \mathfrak{g}$. Hence $\exp X$ and $\exp Y$ commute for all $X, Y \in \mathfrak{g}$ and it follows that G_e is commutative.

Conversely, assume (b). Let $x \in G_e$. Then it follows by the previous lemma that $\mathrm{Ad}(x) = I$. In particular this holds for $x = \exp(tX)$, with $X \in \mathfrak{g}$ and $t \in \mathbb{R}$. It follows that $e^{\mathrm{ad}(tX)} = \mathrm{Ad}(\exp(tX)) = I$. Differentiating at $t = 0$ we obtain $\mathrm{ad}(X) = 0$. Hence $[X, Y] = 0$ for all $X, Y \in \mathfrak{g}$ and (a) follows.

Finally, if G is connected, then $G_e = G$ and the last assertion follows. $\square$

6 Commutative Lie groups

From Example 4.17, we recall that the group $\mathbb{R}^p/\mathbb{Z}^p$ ($p \in \mathbb{N}$) has a unique structure of manifold which turns the natural projection $\pi : \mathbb{R}^p \to \mathbb{R}^p/\mathbb{Z}^p$ into a local diffeomorphism. With this structure of manifold, the group $\mathbb{R}^p/\mathbb{Z}^p$ is a commutative Lie group. It is isomorphic with the p-dimensional torus $\mathbb{T}^p$.

In this section we will prove the following classification of commutative Lie groups.

Theorem 6.1 *Let G be a commutative connected Lie group. Then there exist $p, q \in \mathbb{N}$ such that $G \simeq \mathbb{T}^p \times \mathbb{R}^q$.*

Before we give the proof, we need to collect some results on discrete subgroups of a Lie group. A subgroup H of a Lie group is called discrete if it is discrete as a topological space for the restriction topology. Equivalently, this means that for every $h \in H$ there exists an open neighborhood U in G such that $U \cap H = \{h\}$.

Proposition 6.2 *Let G be a Lie group and H a subgroup. Then the following statements are equivalent.*

(a) *There exists an open neighborhood U of e such that $U \cap H = \{e\}$.*

(b) *The group H is discrete.*

(c) *For every compact subset $C \subset G$, the intersection $C \cap H$ is finite.*

(d) *The group H is a closed Lie subgroup with Lie algebra $\{0\}$.*

Proof: '(a) $\Rightarrow$ (b)': Let $h \in H$. Then $U_h = hU$ is an open neighborhood of h in G. Moreover, $U_h \cap H = hU \cap H = h(U \cap h^{-1}H) = h(U \cap H) = \{h\}$.

'(b) $\Rightarrow$ (c)': We first prove that H is closed in G. Let U be an open neighborhood of e in G such that $U \cap H = \{e\}$. Let $g \in G$ be a point in the closure of H. Then it suffices to show that $g \in H$. There exists a sequence $\{h_j\}$ in H converging to g. It follows that $h_{j+1}h_j^{-1} \to gg^{-1} = e$, as $j \to \infty$. Hence for j sufficiently large we have $h_{j+1}h_j^{-1} \in U \cap H = \{e\}$, hence $h_j = h_{j+1}$. It follows that the sequence h_j becomes stationary after a certain index; hence $h_j = g$ for j sufficiently large and we conclude that $g \in H$.

It follows from the above that the set $H \cap C$ is closed in C, hence compact. For $h \in H \cap C$ we select an open subset of U_h of G such that $U_h \cap H = \{h\}$. Then $\{U_h \mid h \in H \cap C\}$ is an open cover of $H \cap C$ which does not contain a proper subcover. By compactness of $H \cap C$ this cover must therefore be finite, and we conclude that $H \cap C$ is finite.

'(c) $\Rightarrow$ (d)' Let $g \in G$ be a point in the closure of H. The point g has a compact neighborhood C. Now g lies in the closure of $H \cap C$; the latter set is finite, hence closed. Hence $g \in H \cap C \subset H$ and we conclude that the closure of H is contained in H. Therefore, H is closed.

It follows that H is a closed Lie subgroup. Its Lie algebra $\mathfrak{h}$ consists of the $X \in \mathfrak{g}$ with $\exp(\mathbb{R}X) \subset H$. Since $\exp : \mathfrak{g} \to G$ is a local diffeomorphism at 0, there exists an open neighborhood Ω of 0 if $\mathfrak{g}$ such that exp is injective on Ω. Let $X \in \mathfrak{g} \setminus \{0\}$. Then there exists an $\epsilon > 0$ such that $[-\epsilon, \epsilon]X \subset \Omega$. The curve $c : [-\epsilon, \epsilon] \to G,\ t \mapsto \exp tX$ has compact image; this image has a finite intersection with H. Hence $\{t \in [-\epsilon, \epsilon] \mid \exp tX \in H\}$ is finite, and we see that $X \notin \mathfrak{h}$. It follows that $\mathfrak{h} = \{0\}$.

'(d) $\Rightarrow$ (a)' Assume (d). Then H is a closed smooth submanifold of X of dimension 0. By definition this implies that there exists an open neighborhood U of e in G such that $U \cap H = \{e\}$. Hence (a). $\square$

Proof of Theorem 6.1: Assume that G is a connected Lie group that is commutative. Then its Lie algebra $\mathfrak{g}$ is commutative, i.e., $[X, Y] = 0$ for all $X, Y \in \mathfrak{g}$. From this it follows that $\exp X \exp Y = \exp(X + Y)$ for all $X, Y \in \mathfrak{g}$. Therefore, the map $\exp : \mathfrak{g} \to G$ is a homomorphism of the Lie group $(\mathfrak{g}, +, 0)$ to G. It follows that $\exp(\mathfrak{g})$ is already a subgroup of G, hence equals the subgroup G_e generated by it. Since G is connected, $G_e = G$, and it follows that exp has image G, hence is a surjective Lie group homomorphism. Let Γ be the closed subgroup $\ker(\exp)$ of $\mathfrak{g}$. By the isomorphism theorem for groups we have $G \simeq \mathfrak{g}/\Gamma$ as groups.

Since exp is a local diffeomorphism at 0, there exists an open neighborhood Ω of 0 in $\mathfrak{g}$ on which exp is injective. In particular this implies that $\Omega \cap \Gamma = \{0\}$. By Proposition 6.2 it follows that Γ is a discrete subgroup of $\mathfrak{g}$. In view of Lemma 6.4 below there exists a collection $\gamma_1, \ldots, \gamma_p$ of linear independent elements in $\mathfrak{g}$ such that $\Gamma = \mathbb{Z}\gamma_1 \oplus \cdots \oplus \mathbb{Z}\gamma_p$. We may extend the above set to a basis $\gamma_1, \ldots, \gamma_n$ of $\mathfrak{g}$; here $n = \dim\mathfrak{g} = p + q$ for some $q \in \mathbb{N}$. Via the basis $\gamma_1, \ldots, \gamma_n$ we obtain a linear isomorphism $\varphi : \mathfrak{g} \to \mathbb{R}^p \times \mathbb{R}^q$. Let $\epsilon = \exp \circ \varphi^{-1}$, then $\epsilon : \mathbb{R}^n \to G$ is a surjective Lie group homomorphism, and a local diffeomorphism everywhere. Moreover, its kernel equals $\varphi(\Gamma) = \mathbb{Z}^p \times \{0\}$. It follows that ϵ factors through a bijective group homomorphism $\bar{\epsilon} : (\mathbb{R}/\mathbb{Z})^p \times \mathbb{R}^q \simeq \mathbb{R}^n/(\mathbb{Z}^p \times \{0\}) \to G$. The canonical map $\pi : \mathbb{R}^n \to (\mathbb{R}/\mathbb{Z})^p \times \mathbb{R}^q$ is a local diffeomorphism. Moreover, $\epsilon = \bar{\epsilon} \circ \pi$ is a local diffeomorphism as well. Hence $\bar{\epsilon}$ is a local diffeomorphism. Since $\bar{\epsilon}$ is a bijective as well, we conclude that $\bar{\epsilon}$ is a diffeomorphism, hence an isomorphism of Lie groups. $\square$

Lemma 6.3 *Let $\varphi : G \to H$ be a homomorphism of Lie groups. If φ is a local diffeomorphism at e, then φ is a local diffeomorphism at every point of G.*

Proof: We prove this by homogeneity. Let $a \in G$. Then from $\varphi(ax) = \varphi(a)\varphi(x)$ we see that $\varphi \circ l_\alpha = l_{\varphi(a)} \circ \varphi$, hence $\varphi = l_{\varphi(a)} \circ \varphi \circ l_a^{-1}$. Now l_a and $l_{\varphi(a)}$ are diffeomorphisms. Since l_a^{-1} maps a to e, whereas φ is a local diffeomorphism at e it follows that $l_{\varphi(a)} \circ \varphi \circ l_a^{-1}$ is a local diffeomorphism at a; hence φ is a local diffeomorphism at a. $\square$

Lemma 6.4 *Let V be a finite dimensional real linear space. Let Γ be a discrete subgroup of V. Then there exists a collection of linearly independent elements $\gamma_1, \ldots, \gamma_p$ of V such that $\Gamma = \mathbb{Z}\gamma_1 \oplus \cdots \oplus \mathbb{Z}\gamma_p$.*

Proof: We prove the lemma by induction on the dimension of V.

First assume that $\dim V = 1$. Via a choice of basis we may identify V with $\mathbb{R}$; then Γ becomes a discrete subgroup of $\mathbb{R}$. Suppose $\Gamma \neq \{0\}$. Then there exists an element $a \in \Gamma \setminus \{0\}$. Passing to $-a$ is necessary, we may assume that $a > 0$. Now $\Gamma \cap [0, a]$ is finite (cf. Prop. 6.2), hence contains a smallest element γ. We note that $\Gamma \cap]\,0, 1\,[\,\gamma = \emptyset$. Now $\Gamma \supset \mathbb{Z}\gamma$. On the other hand, if $g \in \Gamma$, then $g \notin \mathbb{Z}\gamma$ would imply that $g \in]\,m, m+1\,[\,\gamma$ for a suitable $m \in \mathbb{Z}$. This would imply that $g - m\gamma \in \Gamma \cap]\,0, 1\,[\,\gamma = \emptyset$, contradiction. It follows that $\Gamma \subset \mathbb{Z}\gamma$. Hence $\Gamma = \mathbb{Z}\gamma$. This completes the proof of the result for $\dim V = 1$.

Now assume that $\dim V > 1$ and that the result has been established for spaces of strictly smaller dimension. If $\Gamma = \{0\}$ we may take $p = 0$ and we are done. Thus, assume that $\gamma \in \Gamma \setminus \{0\}$. Then the intersection of $\mathbb{R}\gamma$ with Γ is discrete in $\mathbb{R}\gamma$ and non-trivial, hence of the form $\mathbb{Z}\gamma_1$ by the first part of the proof. Select a linear subspace W of V such that $\mathbb{R}\gamma_1 \oplus W = V$. Let π denote the corresponding projection $V \to W$. Let C be a compact subset of W. Then

$$\pi^{-1}(C) = \mathbb{R}\gamma_1 + C = ([0, 1]\gamma_1 + C) + \mathbb{Z}\gamma_1.$$

From this it follows that

$$\begin{aligned} C \cap \pi(\Gamma) \subset \pi(\pi^{-1}(C) \cap \Gamma) &= \pi((C + [0, 1]\gamma_1) \cap \Gamma + \mathbb{Z}\gamma_1) \\ &= \pi((C + [0, 1]\gamma_1) \cap \Gamma); \end{aligned}$$

the latter set is finite by compactness of $C + [0, 1]\gamma_1$. Thus we see that $\pi(\Gamma) \cap C$ is finite for every compact subset of W. By Prop. 6.2 this implies that $\pi(\Gamma)$ is a discrete subgroup of W. By the induction hypothesis there exist linearly independent elements $\bar{\gamma}_2, \ldots, \bar{\gamma}_p$ of $\pi(\Gamma)$ such that $\pi(\Gamma) = \mathbb{Z}\bar{\gamma}_2 \oplus \cdots \mathbb{Z}\bar{\gamma}_p$. Fix $\gamma_2, \ldots, \gamma_p \in \Gamma$ such that $\pi(\gamma_j) = \bar{\gamma}_j$. Then the elements $\gamma_1, \ldots, \gamma_p$ are readily seen to be linear independent; moreover, $\Gamma = \mathbb{Z}\gamma_1 \oplus \cdots \oplus \mathbb{Z}\gamma_p$. $\square$

7 Lie subgroups

Definition 7.1 A *Lie subgroup* of a Lie group G is a subgroup H, equipped with the structure of a Lie group, such that the inclusion map $\iota : H \to G$ is a Lie group homomorphism.

The above definition allows examples of Lie subgroups that are not submanifolds. This is already so if we restrict ourselves to one-parameter subgroups.

Lemma 7.2 *Let G be a Lie group, and let $X \in \mathfrak{g}$. The image of the one-parameter subgroup α_X is a Lie subgroup of G.*

Proof: The result is trivial for $X = 0$. Thus, assume that $X \neq 0$. The map $\alpha_X : \mathbb{R} \to G$ is a Lie group homorphism. Its image H is a subgroup of G.

Assume first that α_X is injective. Then H has a unique structure of smooth manifold for which the bijection $\alpha_X : \mathbb{R} \to H$ is a diffeomorphism. Clearly, this structure turns H into a Lie group and the inclusion map $i : H \to G$ is a Lie group homomorphism.

Next, assume that α_X is not injective. As $\alpha'_X(0) = X \neq 0$, the map $t \mapsto \alpha_X(t) = \exp tX$ is injective on a suitable open interval I containing 0. It follows that $\ker \alpha_X$ is a discrete subgroup of $\mathbb{R}$. Hence $\ker \alpha_X = \mathbb{Z}\gamma$ for some $\gamma \in \mathbb{R}$. This implies that there exists a unique group homomorphism $\bar{\alpha} : \mathbb{R}/\mathbb{Z}\gamma \to \mathbb{R}$ such that $\alpha_X = \bar{\alpha} \circ \mathrm{pr}$. Since pr is a local diffeomorphism, the map $\bar{\alpha}$ is smooth, hence a Lie group homomorphism. Therefore, $H = \mathrm{im}\, \bar{\alpha}$ is compact. By homogeneity, $\bar{\alpha}$ is an injective immersion. This implies that $\bar{\alpha}$ is an embedding of $\mathbb{R}/\mathbb{Z}\gamma$ onto a smooth submanifold of G. We conclude that $H = \mathrm{im}\, \bar{\alpha}$ is a smooth submanifold of G, hence a Lie subgroup. $\square$

We will give an example of a one-parameter subgroup of $\mathbb{T}^2$ whose image is everywhere dense. The following lemma is needed as a preliminary.

Lemma 7.3 *Let S be an infinite subgroup of $\mathbb{T}$. Then S is everywhere dense.*

Proof: If the subgroup S were discrete at 1, it would be finite, by compactness of $\mathbb{T}$. It follows that there exists a sequence σ_n in $S \setminus \{1\}$ such that $\sigma_n \to 1$. We consider the surjective Lie group homomorphism $p : \mathbb{R} \to \mathbb{T}$ given by $p(t) = e^{2\pi i t}$. Since $p : \mathbb{R} \to \mathbb{T}$ is a local diffeomorphism at 0, there exists a sequence s_n in $\mathbb{R} \setminus \{0\}$ such that $p(s_n) = \sigma_n$ and $s_n \to 0$.

Let $\xi \in \mathbb{T}$. Fix $x \in \mathbb{R}$ with $p(x) = \xi$. For each n there exists a unique $k_n \in \mathbb{Z}$ such that $x \in [k_n s_n, (k_n + 1)s_n)$. Thus, $|k_n s_n - x| < |s_n|$ and it follows that $k_n s_n \to x$. Therefore, in $\mathbb{T}$ we have $\sigma_n^{k_n} = p(k_n s_n) \to \xi$. Since $\sigma_n^{k_n} \in S$ for every n, we conclude that x belongs to the closure of S. Hence, S is dense. $\square$

Corollary 7.4 *Let $\alpha : \mathbb{R} \to \mathbb{T}^2$ be an injective one-parameter subgroup of $\mathbb{T}^2$. Then the image of α is dense in $\mathbb{T}^2$.*

Proof: Let H denote the image of α. For $j = 1, 2$, let $p_j : \mathbb{T}^2 \to \mathbb{T}$ denote the projection onto the j-th component and consider the one-parameter subgroup $\alpha_j := p_j \circ \alpha : \mathbb{R} \to \mathbb{T}$. Its kernel Γ_j is an additive subgroup of $\mathbb{R}$, hence either trivial or infinite. If $\alpha_j'(0) = 0$ then $\Gamma_j = \mathbb{R}$. If $\alpha_j'(0) \neq 0$ then α_j is immersive at 0, hence everywhere by homogeneity. It follows that the image of α_j is an open subgroup of $\mathbb{T}$, hence equal to $\mathbb{T}$, by connectedness of the latter group. It follows that α_j is a local diffeomorphism from $\mathbb{R}$ onto $\mathbb{T}$. On the other hand, α_j cannot be a diffeomorphism, as $\mathbb{T}$ is compact and $\mathbb{R}$ is not. It follows that Γ_j is not trivial, hence infinite. Thus, in all cases Γ_1 and Γ_2 are infinite additive subgroups of $\mathbb{R}$.

As α is injective, we observe that $\Gamma_1 \cap \Gamma_2 = \ker \alpha = \{0\}$. Hence, α maps Γ_2 injectively to $H \cap (\mathbb{T} \times \{1\})$. Since Γ_2 is infinite, it follows that $H \cap (\mathbb{T} \times \{1\})$ must be infinite. By Lemma 7.3 it follows that $H \cap (\mathbb{T} \times \{1\})$ is dense in $\mathbb{T} \times \{1\}$.

Likewise, $H \cap (\{1\} \times \mathbb{T})$ is dense in $\{1\} \times \mathbb{T}$. Let $z = (s, t) \in \mathbb{T}$. Then there exists a sequence x_n in $H \cap (\mathbb{T} \times \{1\})$ with limit $(s, 1)$. Similarly, there exists a sequence y_n in $H \cap (\{1\} \times T)$ converging to $(1, t)$. It follows that $x_n y_n$ is a sequence in H with limit z. Hence, H is dense. $\square$

We finally come to our example.

Example 7.5 We consider the group $G = \mathbb{R}^2/\mathbb{Z}^2$. The canonical projection $\pi : \mathbb{R}^2 \to \mathbb{R}^2/\mathbb{Z}^2$ is a homomorphism of Lie groups. We recall from Example 4.17 that π is a local diffeomorphism. Accordingly, we use its tangent map π_* to identify $\mathbb{R}^2 = T_0\mathbb{R}^2$ with $\mathfrak{g}$. Let $X \in \mathbb{R}^2$; then the associated one-parameter subgroup $\alpha = \alpha_X$ in G is given by

$$\alpha(t) = tX + \mathbb{Z}^2, \qquad (t \in \mathbb{R}).$$

From Lemma 7.2 it follows that the image H of α_X is a Lie subgroup of G. If $X = 0$, then α_X is constant, and its image is the trivial group. We now assume that $X \neq 0$. If X_1, X_2 have a rational ratio, and $X_1 \neq 0$, then $X_2 = pX_1/q$, with $p, q \in \mathbb{Z}$, $q > 0$. Hence $qX_1^{-1}X \in \mathbb{Z}^2$, and it follows that α_X is not injective. In the proof of Lemma 7.2 we saw that H is a compact submanifold of G, diffeomorphic to the circle. A similar assertion holds in case $X_1/X_2 \in \mathbb{Q}$.

If X_1, X_2 have an irrational ratio, then $tX \notin \mathbb{Z}^2$ for all $t \in \mathbb{R}$, so that α_X is injective. From Corollary 7.4 it follows that H is dense in G in this case.

Lemma 7.6 *Let $\varphi : H \to G$ be an injective homomorphism of Lie groups. Then φ is immersive everywhere. In particular, the tangent map $\varphi_* = T_e\varphi : \mathfrak{h} \to \mathfrak{g}$ is injective.*

Proof: We will first establish the last assertion. There exists an open neighborhood Ω of 0 in $\mathfrak{h}$ such that $\exp_H$ maps Ω diffeomorphically onto an open neighborhood of e in H. The following diagram commutes:

$$\begin{array}{ccc} H & \xrightarrow{\varphi} & G \\ {\scriptstyle \exp_H} \uparrow & & \uparrow {\scriptstyle \exp_G} \\ \mathfrak{h} & \xrightarrow{\varphi_*} & \mathfrak{g} \end{array}$$

Since $\exp_H$ is injective on Ω, it follows that $\varphi \circ \exp_H$ is injective on Ω; hence so is $\exp_G \circ \varphi_*$. It follows that φ_* is injective on Ω. Hence $\ker(\varphi_*) \cap \Omega = \{0\}$. But $\ker(\varphi_*)$ is a linear subspace of $\mathfrak{h}$; it must be trivial, since its intersection with an open neighborhood of 0 is a point.

We have shown that φ is immersive at e. We may complete the proof by homogeneity. Let $h \in H$ be arbitrary. Then $l_{\varphi(h)} \circ \varphi \circ l_{h^{-1}} = \varphi$. Hence, by taking tangent maps at h it follows that $T_h\varphi$ is injective. $\square$

In the following we assume that H is a Lie subgroup of G. The inclusion map is denoted by $\iota : H \to G$. As usual we denote the Lie algebras of these Lie groups by $\mathfrak{h}$ and $\mathfrak{g}$, respectively. The following result is an immediate consequence of the above lemma.

Corollary 7.6' The tangent map $\iota_* := T_e\iota : \mathfrak{h} \to \mathfrak{g}$ is injective.

We recall that ι_* is a homomorphism of Lie algebras. Thus, via the embedding ι_* the Lie algebra $\mathfrak{h}$ may be identified with a *Lie subalgebra* of $\mathfrak{g}$, i.e., a linear subspace that is closed under the Lie bracket. We will make this identification from now on. Note that after this identification the map ι_* of the above diagram becomes the inclusion map.

Lemma 7.7 *As a subalgebra of* $\mathfrak{g}$, *the Lie algebra of* H *is given by:*

$$\mathfrak{h} = \{X \in \mathfrak{g} \mid \forall t \in \mathbb{R} : \quad \exp_G(tX) \in H\}.$$

Proof: We denote the set on the right-hand side of the above equation by V.

Let $X \in \mathfrak{h}$. Then $\exp_G(tX) = \iota(\exp_H tX)$ by commutativity of the above diagram with $\varphi = \iota$. Hence, $\exp_G(tX) \in \iota(H) = H$ for all $t \in \mathbb{R}$. This shows that $\mathfrak{h} \subset V$.

To prove the converse inclusion, let $X \in \mathfrak{g}$, and assume that $X \notin \mathfrak{h}$. We consider the map $\varphi : \mathbb{R} \times \mathfrak{h} \to G$ defined by

$$\varphi(t, Y) = \exp(tX)\exp(Y).$$

The tangent map of φ at $(0, 0)$ is the linear map $T_{(0,0)}\varphi : \mathbb{R} \times \mathfrak{h} \to \mathfrak{g}$ given by

$$T_{(0,0)}\varphi : (\tau, Y) \mapsto \tau X + Y.$$

Since $X \notin \mathfrak{h}$, its kernel is trivial. By the immersion theorem there exists a constant $\epsilon > 0$ and an open neighbourhood Ω of 0 in $\mathfrak{h}$, such that φ maps $]-\epsilon, \epsilon[\times\Omega$ injectively into G. Shrinking Ω if necessary, we may in addition assume that $\exp_H$ maps Ω diffeomorphically onto an open neighborhood U of e in H.

The map $m : (x, y) \mapsto x^{-1}y,\ H \times H \to H$ is continuous, and maps (e, e) to e. Since U is an open neighborhood of e in H, there exists an open neighborhood U_0 of e in H such that $m(U_0 \times U_0) \subset U$, or, written differently,

$$U_0^{-1}U_0 \subset U.$$

Since H is a union of countably many compact sets, there exists a countable collection $\{h_j \mid j \in \mathbb{N}\} \subset H$ such that the open sets h_jU_0 cover H. For every $j \in \mathbb{N}$ we define

$$T_j = \{t \in \mathbb{R} \mid \exp tX \in h_jU_0\}.$$

Let now $j \in \mathbb{N}$ be fixed for the moment, and assume that $s, t \in T_j,\ |s - t| < \epsilon$. Then it follows from the definition of T_j that $\exp[(t - s)X] = \exp(-sX)\exp(tX) \in U_0^{-1}U_0 \subset U$.

Hence $\exp[(t-s)X] = \exp Y$ for a unique $Y \in \Omega$, and we see that $\varphi(t-s, 0) = \varphi(0, Y)$. By injectivity of φ on $]-\epsilon, \epsilon[\times\Omega$ it follows that $Y = 0$ and $s = t$. From the above we conclude that different elements $s, t \in T_j$ satify $|s-t| \geq \epsilon$. Hence T_j is countable.

The union of countably many countable sets is countable. Hence the union of the sets T_j is properly contained in $\mathbb{R}$ and we see that there exists a $t \in \mathbb{R}$ such that $t \notin T_j$ for all $j \in \mathbb{N}$. This implies that $\exp tX \notin \cup_{j\in\mathbb{N}} h_j U_0 = H$. Hence $X \notin V$. Thus we see that $\mathfrak{g} \setminus \mathfrak{h} \subset \mathfrak{g} \setminus V$ and it follows that $V \subset \mathfrak{h}$. $\square$

Example 7.8 Let V be a finite dimensional linear space (with $k = \mathbb{R}$ or $\mathbb{C}$). In Example 2.14 we saw that $\mathrm{SL}(V)$ is a submanifold of $\mathrm{GL}(V)$, hence a Lie subgroup. The Lie algebra of $\mathrm{GL}(V)$ is equal to $\mathfrak{gl}(V) = \mathrm{End}(V)$, equipped with the commutator brackets. We recall from Example 2.14 that $\det : \mathrm{GL}(V) \to k$ is a submersion at I. Hence the tangent space $\mathfrak{sl}(V)$ of $\mathrm{SL}(V) = \det^{-1}(1)$ at I is equal to $\ker(T_I \det) = \ker \mathrm{tr}$. We conclude that the Lie algebra of $\mathrm{SL}(V)$ is given by

$$\mathfrak{sl}(V) = \{X \in \mathrm{End}(V) \mid \mathrm{tr}X = 0\}; \tag{8}$$

in particular, it is a subalgebra of $\mathfrak{gl}(V)$. The validity of (8) may also be derived by using the methods of this section, as follows.

If $X \in \mathfrak{sl}(V)$, then by Lemma 7.7, $\exp(tX) \in \mathrm{SL}(V)$ for all $t \in \mathbb{R}$, hence

$$\mathrm{tr}X = \left.\frac{d}{dt}\right|_{t=0} \det(e^{tX}) = \left.\frac{d}{dt}\right|_{t=0} 1 = 0.$$

It follows that $\mathfrak{sl}(V)$ is contained in the set on the right-hand side of (8).

For the converse inclusion, let $X \in \mathrm{End}(V)$, and assume that $\mathrm{tr}X = 0$. Then for every $t \in \mathbb{R}$ we have $\det e^{tX} = e^{\mathrm{tr}(tX)} = 1$, hence $\exp tX = e^{tX} \in \mathrm{SL}(V)$. Using Lemma 7.7 we conclude that $X \in \mathfrak{sl}(V)$.

Example 7.9 We consider the subgroup $\mathrm{O}(n)$ of $\mathrm{GL}(n, \mathbb{R})$ consisting of real $n \times n$ matrices x with $x^t x = I$. Being a closed subgroup, $\mathrm{O}(n)$ is a Lie subgroup. We claim that its Lie algebra is given by

$$\mathfrak{o}(n) = \{X \in \mathrm{M}(n, \mathbb{R}) \mid X^t = -X\}, \tag{9}$$

the space of anti-symmetric $n \times n$ matrices. Indeed, let $X \in \mathfrak{o}(n)$. Then by Lemma 7.7, $\exp sX \in \mathrm{O}(n)$, for all $s \in \mathbb{R}$. Hence,

$$I = (e^{sX})^t e^{sX} = e^{sX^t} e^{sX}.$$

Differentiating with respect to s at $s = 0$ we obtain $X^t + X = 0$, hence X belongs to the set on the right-ghand side of (9).

For the converse inclusion, assume that $X \in \mathrm{M}(n, \mathbb{R})$ and $X^t = -X$. Then, for every $s \in \mathbb{R}$,

$$(e^{sX})^t e^{sX} = e^{sX^t} e^{sX} = e^{-sX} e^{sX} = I.$$

Hence $\exp sX \in \mathrm{O}(n)$ for all $s \in \mathbb{R}$, and it follows that $X \in \mathfrak{o}(n)$.

If $X \in \mathfrak{o}(n)$ then its diagonal elements are zero. Hence $\mathrm{tr}X = 0$ and we conclude that $X \in \mathfrak{sl}(n, \mathbb{R})$. Therefore, $\mathfrak{o}(n) \subset \mathfrak{sl}(n, \mathbb{R})$. It follows that $\exp(\mathfrak{o}(n)) \subset \mathrm{SL}(n, \mathbb{R})$, hence $\mathrm{O}(n)_e \subset$

$\mathrm{SL}(n, \mathbb{R})$. We conclude that $\mathrm{O}(n)_e \subset \mathrm{SO}(n) \subset \mathrm{O}(n)$. Since $\mathrm{SO}(n)$ is connected, see exercises, it follows that

$$\mathrm{O}(n)_e = \mathrm{SO}(n).$$

The determinant $\det : \mathrm{O}(n) \to \mathbb{R}^*$ has image $\{-1, 1\}$ and kernel $\mathrm{SO}(n)$, hence induces a group isomorphism $\mathrm{O}(n)/\mathrm{O}(n)_e \simeq \{-1, 1\}$. It follows that $\mathrm{O}(n)$ consists of two connected components, $\mathrm{O}(n)_e$ and $x\mathrm{O}(n)_e$, where x is any orthogonal matrix with determinant -1. Of course, one may take x to be the diagonal matrix with -1 in the bottom diagonal entry, and 1 in the remaining diagonal entries, i.e., x is the reflection in the hyperplane $x_n = 0$.

Lemma 7.10 *Let G be a Lie group and $H \subset G$ a subgroup. Then H allows at most one structure of Lie subgroup.*

Proof: See exercises. □

We now come to a result that is the main motivation for allowing Lie subgroups that are not closed.

Theorem 7.11 *Let G be a Lie group with Lie algebra $\mathfrak{g}$. If $\mathfrak{h} \subset \mathfrak{g}$ is a Lie subalgebra, then the subgroup $\langle \exp \mathfrak{h} \rangle$ generated by* $\exp \mathfrak{h}$ *has a unique structure of Lie subgroup. Moreover, the map $\mathfrak{h} \mapsto \langle \exp \mathfrak{h} \rangle$ is a bijection from the collection of Lie subalgebras of $\mathfrak{g}$ onto the collection of connected Lie subgroups of G.*

Proof: See next section. □

Remark 7.12 In the literature, the group $\langle \exp \mathfrak{h} \rangle$ is usually called the analytic subgroup of G with Lie algebra $\mathfrak{h}$.

8 Proof of the analytic subgroup theorem

The proof of Theorem 7.11 will be based on the following result. Throughout this section we assume that G is a Lie group and that $\mathfrak{h}$ is a subalgebra of its Lie algebra $\mathfrak{g}$.

Lemma 8.1 *There exists an open neighborhood Ω of 0 in $\mathfrak{g}$ such that $M = \exp(\mathfrak{h} \cap \Omega)$ is a submanifold of G with tangent space equal to*

$$T_m M = T_e(l_m)\mathfrak{h}, \tag{10}$$

for every $m \in M$. If Ω is any such neighborhood, then also $T_m M = T_e(r_m)\mathfrak{h}$ for all $m \in M$.

In the literature one usually proves this result by using the Frobenius integrability theorem for subbundles of the tangent bundle. We will first recall this proof, and then give an independent proof based on a calculation of the derivative of the exponential map.

Proof: We consider the subbundle S of TG given by $S_x = T_e(l_x)$. Then for all $X \in \mathfrak{h}$ the left invariant vector field v_X is a section of S. We note that $[v_X, v_Y] = \pm v_{[X,Y]}$ (see one of the exercises). Hence for all $X, Y \in \mathfrak{h}$ the Lie bracket of v_X and v_Y defines a section of S as well. Let now $X_1, \dots, X_k$ and put $v_j = v_{X_j}$.

Let now ξ, η be any pair of smooth sections of S; then $\xi = \sum_{i=1}^k \xi^i v_i$ and $\eta = \sum_{j=1}^k \eta^j v_j$ for uniquely defined smooth functions ξ^i and η^j on G. Since

$$[\xi, \eta] = \sum_{i,j} \xi^i \eta^j [v_i, v_j] + \xi^i v_i(\eta^j) v_j + \eta^j v_j(\xi^i) v_i,$$

it follows that $[\xi, \eta]$ is a section of S. By the Frobenius integrability theorem it follows that the bundle S is integrable. In particular, there exists a k-dimensional submanifold N of G containing e, such that $T_x N = S_x$ for all $x \in N$.

For $X \in \mathfrak{h}$, the vector field v_X is everywhere tangent to N, hence restricts to a smooth vector field v_X^N on N. By smooth parameter dependence of this vector field on X there exists an open neighborhood U of 0 in $\mathfrak{h}$ and a positive constant $\delta > 0$ such that for every $X \in U$ the integral curve γ_X of v_X^N with initial point e is defined on $I_\delta :=]-\delta, \delta[$. This integral curve is also an integral curve for v_X, hence equals $\alpha_X : t \mapsto \exp tX$ on I_δ. It follows that $\exp \delta U \subset N$. We may now select an open neighborhood Ω of 0 in $\mathfrak{g}$ such that exp is a diffeomorphism from Ω onto an open subset of G and such that $\Omega \cap \mathfrak{h}$ is contained in δU. Then $M := \exp(\Omega \cap \mathfrak{h})$ is a k-dimensional hence open submanifold of N. It follows that $T_x M = T_x N = T_e(l_x)\mathfrak{h}$ for all $x \in M$.

For the last assertion, we note that $r_m = r_m l_m^{-1} l_m$ so that $r_m = l_m \mathcal{C}_{m^{-1}}$ and $T_e(r_m) = T_e(l_m)\mathrm{Ad}(m^{-1})$, and it suffices to show that $\mathrm{Ad}(m^{-1})$ leaves $\mathfrak{h}$ invariant. Write $m = \exp X$ with $X \in \Omega \cap \mathfrak{h}$. Then

$$\mathrm{Ad}(m^{-1}) = \mathrm{Ad}(\exp(-X)) = e^{-\,\mathrm{ad} X}$$

and the result follows, since $\mathrm{ad}(X)$ leaves the closed subspace $\mathfrak{h}$ invariant. □

We shall now give a different proof of Lemma 8.1. The following result plays a crucial role.

Lemma 8.2 *Let $X \in \mathfrak{g}$. Then*

$$\begin{aligned} T_X \exp &= T_e(l_{\exp X}) \circ \int_0^1 e^{-s\,\mathrm{ad} X}\, ds \\ &= T_e(r_{\exp X}) \circ \int_0^1 e^{s\,\mathrm{ad} X}\, ds. \end{aligned}$$

Proof: For $X, Y \in \mathfrak{g}$, we define

$$F(X, Y) = [T_e(l_{\exp X})]^{-1} \circ T_X(\exp) Y \in \mathfrak{g}$$

and note that, by the chain rule,

$$F(X, Y) = T_{\exp X}(l_{\exp(-X)}) T_X(\exp) Y = \left.\frac{\partial}{\partial t}\right|_{t=0} \exp(-X) \exp(X + tY).$$

From this it follows by interchanging partial derivatives, that

$$\frac{\partial}{\partial s} F(sX, sY) = \left.\frac{\partial}{\partial t}\right|_{t=0} \frac{\partial}{\partial s} \exp(-sX)\exp(sX + tsY).$$

Now,

$$\begin{aligned}
\frac{\partial}{\partial s} \exp(-sX)\exp(sX + tsY) &= \left.\frac{\partial}{\partial \sigma}\right|_{\sigma=0} \exp(-(-s+\sigma)X)\exp((s+\sigma)(X+tY)) \\
&= \left.\frac{\partial}{\partial \sigma}\right|_{\sigma=0} \exp(-sX)\exp(-\sigma X)\exp(\sigma(X+tY))\exp(sX+tsY) \\
&= T_e(l_{\exp(-sX)} r_{\exp(sX+stY)}) \left.\frac{\partial}{\partial \sigma}\right|_{\sigma=0} \exp(-\sigma X)\exp(\sigma X + \sigma t Y) \\
&= T_e(l_{\exp(-sX)} r_{\exp(sX+stY)})(tY),
\end{aligned}$$

and we conclude that

$$\begin{aligned}
\frac{\partial}{\partial s} F(sX, sY) &= \left.\frac{\partial}{\partial t}\right|_{t=0} T_e(l_{\exp(-sX)} r_{\exp(sX+stY)})(tY) \\
&= \left.\frac{\partial}{\partial t}\right|_{t=0} T_e(l_{\exp(-sX)} r_{\exp(sX+stY)})(0) + \left.\frac{\partial}{\partial t}\right|_{t=0} T_e(l_{\exp(-sX)} r_{\exp(sX)})(tY) \\
&= \mathrm{Ad}(\exp(-sX))Y = e^{-s\,\mathrm{ad}X} Y.
\end{aligned}$$

It follows that

$$F(X,Y) = \int_0^1 \frac{\partial}{\partial s} F(sX, sY)\, ds = \int_0^1 e^{-s\,\mathrm{ad}X} Y\, ds,$$

whence the first identity. The second identity may be obtained in a similar manner. It can also be derived from the first as follows. We have $T_e(l_{\exp X}) = T_e(r_{\exp X}) \circ \mathrm{Ad}(\exp X)$, hence

$$\begin{aligned}
T_e(l_{\exp X}) \circ \int_0^1 e^{-s\,\mathrm{ad}X}\, ds &= T_e(r_{\exp X}) \circ e^{\mathrm{ad}X} \int_0^1 e^{-s\,\mathrm{ad}X}\, ds \\
&= T_e(r_{\exp X}) \circ \int_0^1 e^{(1-s)\,\mathrm{ad}X}\, ds \\
&= T_e(r_{\exp X}) \circ \int_0^1 e^{s\,\mathrm{ad}X}\, ds.
\end{aligned}$$

$\square$

Remark 8.3 The integral in the above expression may be expressed as a power series as follows. Let V be a finite dimensional linear space, and $A \in \mathrm{End}(V)$. Then using the power series expansion for e^{sA}, we obtain

$$\begin{aligned}
\int_0^1 e^{sA}\, ds &= \sum_{n=0}^{\infty} \int_0^1 \frac{1}{n!} A^n s^n\, ds \\
&= \sum_{n=0}^{\infty} \frac{1}{(n+1)!} A^n.
\end{aligned}$$

For obvious reasons, the sum of the latter series is also denoted by $(e^A - I)/A$.

Alternative proof of Lemma 8.1: Let Ω be an open neighborhood of 0 in $\mathfrak{g}$ such that $\exp|_\Omega$ is a diffeomorphism onto an open subset of G. Then $M := \exp(\mathfrak{h} \cap \Omega)$ is a smooth submanifold of G of dimension $\dim M = \dim \mathfrak{h}$. For (10), we put $m = \exp X$, with $X \in \mathfrak{h} \cap \Omega$. Since $\mathfrak{h}$ is a subalgebra, $(e^{-\operatorname{ad} X} - I)/\operatorname{ad} X$ leaves $\mathfrak{h}$ invariant. Hence

$$T_m M = T_X(\exp)\mathfrak{h} = T_e(l_m) \circ \left(\frac{I - e^{-\operatorname{ad} X}}{\operatorname{ad} X} \right) \mathfrak{h} \subset T_e(l_m)\mathfrak{h}.$$

Equality follows for dimensional reasons. The identity with $T_e r_m$ is proved in a similar manner. $\square$

We shall now proceed with the proof of Theorem 7.11, starting with the the result of Lemma 8.1, with Ω and M as given there.

Lemma 8.4 *Let C be a compact subset of M. Then there exists an open neighborhood U of* 0 *in* $\mathfrak{g}$ *such that m* $\exp(\mathfrak{h} \cap U)$ *is open in M for all $m \in C$. In particular, C* $\exp(\mathfrak{h} \cap U)$ *is an open neighborhood of C in M.*

Proof: For every $X \in \mathfrak{h}$, we denote by $\Phi_X : \mathbb{R} \times G \to G$ the flow of the left invariant vectorfield v_X. We recall that $\Phi_X(t, x) = x \exp tX$, for all $X \in \mathfrak{h}$, $t \in \mathbb{R}$ and $x \in G$. For fixed $X \in \mathfrak{h},\ x \in G$, the map $t \mapsto \Phi_X(x, t)$ is the maximal integral curve of v_X with initial point x.

Since the vector field v_X is everywhere tangent to M, it follows that $v_X|_M$ is a vector field on M. For every $X \in \mathfrak{h}$, and $m \in M$, we denote by $t \mapsto \varphi_X(t, m)$ the maximal integral curve of $v_X|_M$ in M, with initial point m. By the theory of systems of ordinary differential equations with parameter dependence, it follows that $(X, t, m) \mapsto \varphi_X(t, m)$ is smooth on its domain, which is an open subset $D \subset \mathfrak{h} \times \mathbb{R} \times M$ containing $\mathfrak{h} \times \{0\} \times M$. Clearly, $t \mapsto \varphi_X(t, m)$ is also an integral curve of v_X in G with initial point m. Therefore, $\varphi_X(t, m) = \Phi_X(t, m)$ for all $(X, t, m) \in D$. We conclude that $\Phi_X(t, m) \in M$ for all $(X, t, m) \in D$. Using that $\Phi_{sX}(t, m) = \Phi_X(st, m)$ and that C is compact, we now readily deduce that there exists an open neighborhood U_0 of 0 in $\mathfrak{h}$ such that $m \exp(tX) = \Phi_X(t, m) \in M$, for all $X \in U_0$, $t \in [0, 1]$ and $m \in C$. We may now select an open neighborhood U of 0 in $\mathfrak{g}$ such that $\exp|_U$ is a diffeomorphism. Moreover, replacing U by a smaller subset if necessary, we may in addition assume that $\mathfrak{h} \cap U \subset U_0$. Then, for every $m \in C$, the map $X \mapsto m \exp X$ is an injective immersion of $\mathfrak{h} \cap U$ into M. Since $\dim M = \dim \mathfrak{h}$, the map is a diffeomorphism onto an open subset of M. The final assertion follows as $C \exp(\mathfrak{h} \cap U)$ is the union of the open sets $m \exp(\mathfrak{h} \cap U)$, for $m \in C$. $\square$

Corollary 8.5 *Let $M \subset G$ be as in Lemma 8.1. Then for every $x_1, x_2 \in G$, the intersection $x_1 M \cap x_2 M$ is open in both $x_1 M$ and $x_2 M$.*

Proof: Let $y \in x_1 M \cap x_2 M$. Then by Lemma 8.4, there exists an open neighborhood U of 0 in $\mathfrak{g}$, such that the sets $x_j^{-1} y \exp(U \cap \mathfrak{h})$ are open in M, for $j = 1, 2$. It follows that $y \exp(U \cap \mathfrak{h})$ is open in both $x_1 M$ and $x_2 M$. $\square$

Proof of Theorem 7.11. Let H be the group generated by $\exp \mathfrak{h}$. We will first equip H with the structure of a manifold.

We fix Ω and M as in Lemma 8.1. Replacing Ω by a smaller neighborhood if necessary, we may assume that $\exp|_\Omega$ is a diffeomorphism of Ω onto an open subset of G. Then exp restricts to a diffeomorphism of $\Omega_0 := \Omega \cap \mathfrak{h}$ onto the submanifold M of G. Accordingly, its inverse $\chi : M \to \Omega_0$ is a diffeomorphism of manifolds as well.

Since $M \subset H$, it follows that H is covered by the submanifolds hM of G, where $h \in H$. We equip H with the finest topology that makes the inclusions $hM \hookrightarrow H$ continuous. Then by definition a subset $U \subset H$ is open if and only if $U \cap hM$ is open in hM for every $h \in H$. We note that by Corollary 8.5, each set hM, for $h \in H$, is open in H. For each open subset $\mathcal{O} \subset G$ and each $h \in H$, the set $\mathcal{O} \cap hM$ is open in hM; hence $\mathcal{O} \cap H$ is open in H. It follows that the inclusion map $H \hookrightarrow G$ is continuous. Since G is Hausdorff, it now follows that H, with the defined topology, is Hausdorff. For each $h \in H$, the map $\chi_h = \chi \circ l_h^{-1} : hM \to \Omega_0$ is a diffeomorphism. This automatically implies that the transition maps are smooth. Hence $\{\chi_h \mid h \in H\}$ is an atlas.

Fix a compact neighborhood C_0 of 0 in $\Omega \cap \mathfrak{h}$. Then $C = \exp C_0$ is a compact neighborhood of e in M. It follows that C is compact in H. Since $\mathfrak{h}$ is the union of the sets nC_0 for $n \in \mathbb{N}$, it follows that $\exp \mathfrak{h}$ is the union of the sets $\{c^n \mid c \in C\}$, for $n \in \mathbb{N}$. One now readily sees that H is the union of the sets C^n, for $n \in \mathbb{N}$. Each of the sets C^n is compact, being the image of the compact Cartesian product $C \times \cdots \times C$ (n factors) under the continuous multiplication map. Hence the manifold H is a countable union of compact subsets, which in turn implies that its topology has a countable basis.

We will finish the proof by showing that H with the manifold structure just defined is a Lie group. If $h \in H$ then the map $l_h : H \to H$ is a diffeomorphism by definition of the atlas. We will first show that right multiplication $r_h : H \to H$ is a diffeomorphism as well.

If $X \in \mathfrak{h}$ then the linear endomorphism $\mathrm{Ad}(\exp X) : \mathfrak{g} \to \mathfrak{g}$ equals $e^{\,\mathrm{ad} X}$ hence leaves $\mathfrak{h}$ invariant. Since H is generated by elements of the form $\exp X$ with $X \in \mathfrak{h}$, it follows that for every $h \in H$ the linear endomorphism $\mathrm{Ad}(h)$ of $\mathfrak{g}$ leaves $\mathfrak{h}$ invariant. Fix $h \in H$. Then there exists an open neighborhood $\mathcal{O}$ of 0 in $\Omega \subset \mathfrak{g}$ such that $\mathrm{Ad}(h^{-1})(\mathcal{O}) \subset \Omega$ hence also $\mathrm{Ad}(h^{-1})(\mathfrak{h} \cap \mathcal{O}) \subset \mathfrak{h} \cap \Omega$. From $\exp Xh = h \exp \mathrm{Ad}(h^{-1})X$ we now see that

$$\chi_h \circ r_h = \mathrm{Ad}(h)^{-1} \circ \chi_e \qquad \text{on} \quad \exp(\mathfrak{h} \cap \mathcal{O}).$$

This implies that $r_h : \exp(\mathfrak{h} \cap \mathcal{O}) \to M$ is smooth. Hence, $r_h : H \to H$ is smooth at e. By left homogeneity it follows that $r_h : H \to H$ is smooth everywhere. Since r_h is bijective with inverse $r_{h^{-1}}$, it follows that r_h is a diffeomorphism from H to itself.

We will finish by showing that the multiplication map $\mu_H : H \times H \to H, (h, h') \mapsto hh'$ and the inversion map $\iota_H : H \to H, h \mapsto h^{-1}$ are both smooth. If $h_1, h_2 \in H$ then $\mu_H \circ (l_{h_1} \times r_{h_2}) = l_{h_1} r_{h_2} \circ \mu_H$. Since l_{h_1} and r_{h_2} are diffeomorphisms, smoothness of μ_H at (h_1, h_2) follows from smoothness of μ_H at (e, e). Thus, it suffices to show smoothness of μ_H at (e, e). From $\iota_H \circ l_h = r_h^{-1} \circ \iota_H$, we see that it also suffices to prove smoothness of the inversion map ι_H at e.

Fix an open neighborhood N_e of e in M such that $\bar{N}_e$ is a compact subset of M. Then by Lemma 8.4, there exists an open neighborhood U of 0 in $\mathfrak{g}$ such that $N_e \exp(\mathfrak{h} \cap U) \subset M$.

Replacing U by its intersection with Ω, we see that $N_0 = \exp(\mathfrak{h} \cap U)$ is an open neighborhood of e in M and that $N_e N_0 \subset M$. It follows that the smooth map μ_G maps $N_e \times N_0$ into M, hence its restriction $\mu_G|_{N_e \times N_0}$, which equals $\mu_H|_{N_e \times N_0}$, maps $N_e \times N_0$ smoothly into the smooth submanifold M of G. This implies that μ_H is smooth in an open neighborhood of (e, e) in $H \times H$.

For the inversion, we note that $\Omega_1 := \Omega \cap (-\Omega)$ is an open neighborhood of 0 in $\mathfrak{g}$ that is stable under reflection in the origin. It follows that ι_G maps the open neighborhood $N_1 := \exp(\Omega_1 \cap \mathfrak{h})$ of e in M into itself. Hence its restriction to N_1, which equals $\iota_H|_{N_1}$, maps N_1 smoothly into the smooth manifold M. It follows that ι_H is smooth in an open neighborhood of e. $\square$

9 Closed subgroups

Theorem 9.1 *Let H be a subgroup of a Lie group G. Then the following assertions are equivalent:*

(a) *H is a C^∞-submanifold of G at the point e;*

(b) *H is a C^∞-submanifold of G;*

(c) *H is a closed subset of G.*

Note that condition (b) implies that H is a Lie subgroup of G. Indeed, the map $\iota_H : H \to H, h \mapsto h^{-1}$ is the restriction of the smooth map ι_G to the smooth manifold H, hence smooth. Similarly μ_H is the restriction of μ_G to the smooth submanifold $H \times H$ of $G \times G$, hence smooth.

In the proof of the theorem we will need the following result. If G is a Lie group we shall use the notation log for the map $G \supset\!\!\to \mathfrak{g}$, defined on a sufficiently small neighborhood U of e, that inverts the exponential map, i.e., $\exp \circ \log = I$ on U.

Lemma 9.2 *Let $X, Y \in \mathfrak{g}$. Then*

$$X + Y = \lim_{n \to \infty} n \log(\exp(n^{-1}X) \exp(n^{-1}Y)).$$

Proof: Being the local inverse to exp, the map log is a local diffeomorphism at e. Its tangent map at e is given by $T_e \log = (T_0 \exp)^{-1} = I_{\mathfrak{g}}$.

The map $\psi : \mathfrak{g} \times \mathfrak{g} \to G$, $(X, Y) \mapsto \exp X \exp Y$ has tangent map at $(0, 0)$ given by $T_{(0,0)}\psi : (X, Y) \mapsto X + Y$. The composition $\log \circ \psi$ is well defined on a sufficiently small neighborhood of $(0, 0)$ in $\mathfrak{g} \times \mathfrak{g}$. Moreover, by the chain rule its derivative at $(0, 0)$ is given by $(X, Y) \mapsto X + Y$.

It follows that, for $(X, Y) \in \mathfrak{g} \times \mathfrak{g}$ sufficiently close to $(0, 0)$,

$$\log(\exp X \exp Y) = X + Y + \rho(X, Y), \tag{11}$$

where $\rho(X,Y) = o(\|X\| + \|Y\|)$ as $(X,Y) \to (0,0)$ (here $\|\cdot\|$ is any choice of norm on $\mathfrak{g}$). Hence

$$\begin{aligned} n\log(\exp(n^{-1}X)\exp(n^{-1}Y)) &= n[n^{-1}X + n^{-1}Y + \rho(n^{-1}X, n^{-1}Y)] \\ &= X + Y + n\,o(n^{-1}) \to X + Y \qquad (n \to \infty). \end{aligned}$$

□

Proof of Theorem 9.1. Let $n = \dim G$. We first show that (a) $\Rightarrow$ (b). Let k be the dimension of H at e. By the assumption there exists an open neighborhood U of e in G and a diffeomorphism κ onto an open subset of $\mathbb{R}^n$, such that $\kappa(U \cap H) = \kappa(U) \cap \mathbb{R}^k \times \{0\}$. Let now $h \in H$. Then $U_h := hU = l_h(U)$ is an open neighborhood of h in G, and $\kappa_h := \kappa \circ l_h^{-1}$ is a diffeomorphism from U_h onto $\kappa(U)$. Moreover, since $hU \cap H = h(U \cap H)$, it follows that $\kappa_h(U_h \cap H) = \kappa(U \cap H) = \kappa(U) \cap \mathbb{R}^k \times \{0\}$. It follows that H is a C^∞-submanifold at any of its points.

Next we show that (b) $\Rightarrow$ (c). Assume (b). Then there exists an open neighborhood U of e in G such that $U \cap \bar{H} = U \cap H$. Let $y \in \bar{H}$. Then yU is an open neighborhood of $y \in \bar{H}$ in G, hence there exists a $h \in yU \cap H$. Hence $y^{-1}h \in U$. On the other hand, from $y \in \bar{H}, h \in H$ it follows that $y^{-1}h \in \bar{H}$. Hence $y^{-1}h \in U \cap \bar{H} = U \cap H$, and we see that $y \in H$. We conclude that $\bar{H} \subset H$, hence H is closed.

Finally, we show that (c) $\Rightarrow$ (a). We call an element $X \in \mathfrak{g}$ tangential to H if there exist sequences $X_n \in \mathfrak{g}$, $\xi_n \in \mathbb{R}$ such that $\lim_{n\to\infty} X_n = 0$, $\exp X_n \in H$, and $\lim_{n\to\infty} \xi_n X_n = X$. Let T be the set of $X \in \mathfrak{g}$ that are tangential to H. From the definition it is obvious that for all $X \in T$ we have $\mathbb{R}X \subset T$.

We claim that for every $X \in T$ we have $\exp X \in H$. Indeed, let $X \in T$, and let X_n, ξ_n be sequences as above. If $X = 0$ then obviously $\exp X \in H$. If $X \neq 0$, then $|\xi_n| \to \infty$. Choose $m_n \in \mathbb{Z}$ such that $\xi_n \in [m_n, m_n + 1]$. Then $m_n \to \infty$ hence

$$|\xi_n/m_n - 1| \leq 1/|m_n| \to 0$$

and it follows that $m_n/\xi_n \to 1$. Thus, $m_n X_n = (m_n/\xi_n)\xi_n X_n \to X$. Hence

$$\exp X = \lim_{n\to\infty} (\exp X_n)^{m_n} \in \bar{H} \subset H.$$

We also claim that T is a linear subspace of $\mathfrak{g}$. Let $X, Y \in T$. Put $X_n = \frac{1}{n}X$ and $Y_n \in \frac{1}{n}Y$. Then $X_n, Y_n \to 0$ and $X_n, Y_n \in T$, hence $\exp X_n, \exp Y_n \in H$. For $n \in \mathbb{N}$ sufficiently large we may define $Z_n = \log(\exp X_n \exp Y_n)$. Then $\exp Z_n = \exp(X_n)\exp(Y_n) \in H$. Moreover, by (11) we have $Z_n \to 0$ and $nZ_n \to X + Y$. It follows from this that $X + Y \in T$.

We will finish the proof by showing that H is a C^∞ submanifold at the point e. Fix a linear subspace $S \subset \mathfrak{g}$ such that

$$\mathfrak{g} = S \oplus T.$$

Then $\varphi : (X,Y) \mapsto \exp X \exp Y$ is smooth as a map $S \times T \to G$ and has tangent map at $(0,0)$ given by $T_{(0,0)}\varphi : (\xi, \eta) \mapsto \xi + \eta$, $S \times T \to \mathfrak{g}$. This tangent map is bijective, and it follows that φ is a local diffeomorphism at $(0,0)$. Hence, there exist open neighborhoods Ω_S and Ω_T of the origins in S and T respectively such that φ is a diffeomorphism from $\Omega_S \times \Omega_T$ onto an open neighborhood U of e in G.

We will finish the proof by establishing the claim that for Ω_S and Ω_T sufficiently small, we have

$$\varphi(\{0\} \times \Omega_T) = U \cap H.$$

Assume the latter claim to be false. Fix decreasing sequences of neighborhoods Ω_S^k and Ω_T^k of the origins in Ω_S and Ω_T, respectively, with $\Omega_S^k \times \Omega_T^k \to \{0\}$. By the latter assertion we mean that for every open neighborhood $\mathcal{O}$ of $(0, 0)$ in $S \times T$, there exists a k such that $\Omega_S^k \times \Omega_T^k \subset \mathcal{S}$. By the assumed falseness of the claim, we may select $h_k \in \varphi(\Omega_S^k \times \Omega_T^k) \cap H$ such that $h_k \notin \varphi(\{0\} \times \Omega_T^k)$, for all k.

There exist unique $X_k \in \Omega_S^k$ and $Y_k \in \Omega_T^k$ such that $h_k = \varphi(X_k, Y_k) = \exp X_k \exp Y_k$. From the above it follows that X_k is a sequence in $S \setminus \{0\}$ converging to 0. Moreover, from $\exp X_k = h_k \exp(-Y_k)$ we see that $\exp X_k \in H$ for all k. Fix a norm $\|\cdot\|$ on S. Then the sequence $X_k/\|X_k\|$ is contained in the closed unit ball in S, which is compact. Passing to a suitable subsequence we may arrange that $\|X_k\|^{-1} X_k$ converges to an element $X \in S$ of norm 1. Applying the definition of T with $\xi_n = \|X_k\|^{-1}$, we see that also $X \in T$. This contradicts the assumption that $S \cap T = \{0\}$. □

Corollary 9.3 *Let G, H be Lie groups, and let $\varphi : G \to H$ be a continuous homomorphism of groups. Then φ is a C^∞-map (hence a homomorphism of Lie groups).*

Proof: Let $\Gamma = \{(x, \varphi(x)) \mid x \in G\}$ be the graph of φ. Then obviously Γ is a subgroup of the Lie group $G \times H$. From the continuity of φ it follows that Γ is closed. Indeed, let (g, h) belong to the closure of Γ in $G \times H$. Let $\gamma_n = (g_n, h_n)$ be a sequence in Γ converging to (g, h). Then $g_n \to g$ and $h_n \to h$ as $n \to \infty$. Note that $h_n = \varphi(g_n)$. By the continuity of φ it follows that $h_n = \varphi(g_n) \to \varphi(g)$. Hence $\varphi(g) = h$ and we see that $(g, h) \in \Gamma$. Hence Γ is closed.

It follows that Γ is a C^∞-submanifold of $G \times H$. Let $p_1 : G \times H \to G$ and $p_2 : G \times H \to H$ the natural projection maps. Then $p = p_1|\Gamma$ is a smooth map from the Lie group Γ onto G. Note that p is a bijective Lie group homomorphism with inverse $p^{-1} : g \mapsto (g, \varphi(g))$. Thus p^{-1} is continuous. By the lemma below p is a diffeomorphism, hence $p^{-1} : G \to \Gamma$ is C^∞. It follows that $\varphi = p_2 \circ p^{-1}$ is a C^∞-map. □

Lemma 9.4 *Let G, H be Lie groups, and $p : G \to H$ a bijective Lie homomorphism. If p is a homeomorphism (i.e., p^{-1} is continuous), then p is a diffeomorphism (i.e., p^{-1} is C^∞).*

Proof: Consider the commutative diagram

$$\begin{array}{ccc} G & \xrightarrow{p} & H \\ \exp_G \uparrow & & \uparrow \exp_H \\ \mathfrak{g} & \xrightarrow{p_*} & \mathfrak{h} \end{array}$$

where $p_* = T_e p$. Fix open neighborhoods Ω_G, Ω_H of the origins in $\mathfrak{g}, \mathfrak{h}$, respectively, such that $\exp_G|_{\Omega_G}, \exp_H|_{\Omega_H}$ are diffeomorphisms onto open subsets U_G of G and U_H of H respectively.

Replacing Ω_G by a smaller neigborhood if necessary we may assume that $p(U_G) \subset U_H$ (use continuity of p). Since p is a homeomorphism, $p(U_G)$ is an open subset of U_H, containing e. Thus $\Omega'_H := (\exp_H)^{-1}(p(U_G)) \cap \Omega_H$ is an open neighborhood of 0 in $\mathfrak{h}$, contained in Ω_H. Note that $\exp_H$ is a diffeomorphism from Ω'_H onto $U'_H = p(U_G) \subset U_H$. From the commutativity of the diagram and the bijectivity of $\exp_G : \Omega_G \to U_G, \exp_H : \Omega'_H \to U'_H$ and $p : U_G \to U'_H$ it follows that p_* is a bijection of Ω_G onto Ω'_H. It follows from this that p_* is a bijective linear map. Its inverse p_*^{-1} is linear, hence C^∞. Lifting via the exponential maps we see that p^{-1} maps U'_H smoothly onto U_G; it follows that p^{-1} is C^∞ at e. By homogeneity it follows that p^{-1} is C^∞ everywhere. Indeed, let $h \in H$. Then $p^{-1} = l_{p(h)} \circ p^{-1} \circ l_h^{-1}$, since p^{-1} is a homomorphism. But l_h^{-1} maps hU'_H smoothly onto U'_H; hence p^{-1} is C^∞ on hU'_H. $\square$

10 The groups SU(2) and SO(3)

We recall that $\mathrm{SU}(2)$ is the closed subgroup of matrices $x \in \mathrm{SL}(2,\mathbb{C})$ satisfying $x^*x = I$. Here x^* denotes the Hermitian adjoint of x. By a simple calculation we find that $\mathrm{SU}(2)$ consists of all matrices

$$\begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix}, \tag{12}$$

with $(\alpha,\beta) \in \mathbb{C}^2$, $|\alpha|^2 + |\beta|^2 = 1$.

Let $j : \mathbb{C}^2 \to \mathrm{M}(2,\mathbb{C})$ be the map assigning to any point (α,β) the matrix given in (12). Then j is an injective real linear map from $\mathbb{C}^2 \simeq \mathbb{R}^4$ into $\mathrm{M}(2,\mathbb{C}) \simeq \mathbb{C}^4 \simeq \mathbb{R}^8$. In particular, it follows that j is an embedding. Hence, the restriction of j to the unit sphere $S \subset \mathbb{C}^2 \simeq \mathbb{R}^4$ is an embedding of S onto a compact submanifold of $\mathrm{M}(2,\mathbb{C}) \simeq \mathbb{R}^8$. On the other hand, it follows from the above that $j(S) = \mathrm{SU}(2)$. Hence, as a manifold, $\mathrm{SU}(2)$ is diffeomorphic to the 3-dimensional sphere. In particular, $\mathrm{SU}(2)$ is a compact and connected Lie group.

By a calculation which is completely analogous to the calculation in Example 7.9 we find that the Lie algebra $\mathfrak{su}(2)$ of $\mathrm{SU}(2)$ is the algebra of $X \in \mathrm{M}(2,\mathbb{C})$ with

$$X^* = -X, \qquad \mathrm{tr} X = 0.$$

From this one sees that as a real linear space $\mathfrak{su}(2)$ is generated by the elements

$$\mathbf{r}_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \mathbf{r}_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathbf{r}_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Note that $\mathbf{r}_j = i\sigma_j$, where $\sigma_1, \sigma_2, \sigma_3$ are the famous Pauli spin matrices. One readily verifies that $\mathbf{r}_1^2 = \mathbf{r}_2^2 = \mathbf{r}_3^2 = -I$ and $\mathbf{r}_1\mathbf{r}_2 = -\mathbf{r}_2\mathbf{r}_1 = \mathbf{r}_3$, and $\mathbf{r}_2\mathbf{r}_3 = -\mathbf{r}_3\mathbf{r}_2 = \mathbf{r}_1$.

Remark 10.1 One often sees the notation $\mathbf{i} = \mathbf{r}_1, \mathbf{j} = \mathbf{r}_2, \mathbf{k} = \mathbf{r}_3$. Indeed, the real linear span $\mathbb{H} = \mathbb{R}I \oplus \mathbb{R}\mathbf{i} \oplus \mathbb{R}\mathbf{j} \oplus \mathbb{R}\mathbf{k}$ is a realization of the quaternion algebra. The latter is the unique (up to isomorphism) associative $\mathbb{R}$ algebra with unit, on the generators $\mathbf{i}, \mathbf{j}, \mathbf{k}$, subject to the above well known quaternionic relations.

It follows from the above product rules that the commutator brackets are given by

$$[\mathbf{r}_1, \mathbf{r}_2] = 2\mathbf{r}_3, \quad [\mathbf{r}_2, \mathbf{r}_3] = 2\mathbf{r}_1, \quad [\mathbf{r}_3, \mathbf{r}_1] = 2\mathbf{r}_2.$$

From this it follows that the endomorphisms $\mathrm{ad}\mathbf{r}_j \in \mathrm{End}(\mathfrak{su}(2))$ have the following matrices with respect to the basis $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$:

$$\mathrm{mat}\,\mathrm{ad}\mathbf{r}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & 2 & 0 \end{pmatrix}, \quad \mathrm{mat}\,\mathrm{ad}\mathbf{r}_2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ -2 & 0 & 0 \end{pmatrix}, \quad \mathrm{mat}\,\mathrm{ad}\mathbf{r}_3 = \begin{pmatrix} 0 & -2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The above elements belong to

$$\mathfrak{so}(3) = \{X \in \mathrm{M}(3, \mathbb{R}) \mid X^* = -X\},$$

the Lie algebra of the group SO(3).

If $a \in \mathbb{R}^3$, then the exterior product map $X \mapsto a \times X,\ \mathbb{R}^3 \to \mathbb{R}^3$ has matrix

$$R_a = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}$$

with respect to the standard basis e_1, e_2, e_3 of $\mathbb{R}^3$. Clearly $R_a \in \mathfrak{so}(3)$.

Lemma 10.2 *Let $t \in \mathbb{R}$. Then $\exp tR_a$ is the rotation with axis a and angle $t|a|$.*

Proof: Let $r \in \mathrm{SO}(3)$. Then one readily verifies that $R_a = r \circ R_{r^{-1}a} \circ r^{-1}$, and hence

$$\exp tR_a = r \circ \exp[tR_{r^{-1}a}] \circ r^{-1}.$$

Selecting r such that $r^{-1}a = |a|e_1$, we see that we may reduce to the case that $a = |a|e_1$. In that case one readily computes that:

$$\exp tR_a = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t|a| & -\sin t|a| \\ 0 & \sin t|a| & \cos t|a| \end{pmatrix}.$$

□

Write $R_j = R_{e_j}$, for $j = 1, 2, 3$. Then by the above formulas for $\mathrm{mat}\,\mathrm{ad}(\mathbf{r}_j)$ we have

$$\mathrm{mat}\,\mathrm{ad}(\mathbf{r}_j) = 2R_j \qquad (j = 1, 2, 3). \tag{13}$$

We now define the map $\varphi : \mathrm{SU}(2) \to \mathrm{GL}(3, \mathbb{R})$ by $\varphi(x) = \mathrm{matAd}(x)$, the matrix being taken with respect to the basis $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$. Then φ is a homomorphism of Lie groups.

Proposition 10.3 *The map* $\varphi : \mathrm{SU}(2) \to \mathrm{GL}(3,\mathbb{R})$, $x \mapsto \mathrm{matAd}(x)$ *is a surjective group homomorphism onto* $\mathrm{SO}(3)$, *and induces an isomorphism:*

$$\mathrm{SU}(2)/\{\pm I\} \simeq \mathrm{SO}(3).$$

Proof: From

$$\varphi(\exp X) = \mathrm{mat} e^{\,\mathrm{ad} X} = e^{\mathrm{mat\,ad} X}$$

we see that φ maps $\mathrm{SU}(2)_e$ into $\mathrm{SO}(3)$. Since $\mathrm{SU}(2)$ is obviously connected, we have $\mathrm{SU}(2) = \mathrm{SU}(2)_e$, so that φ is a Lie group homomorphism from $\mathrm{SU}(2)$ to $\mathrm{SO}(3)$. The tangent map of φ is given by $\varphi_* : X \mapsto \mathrm{mat\,ad} X$. It maps the basis $\{\mathbf{r}_j\}$ of $\mathfrak{su}(2)$ onto the basis $\{2R_j\}$ of $\mathfrak{so}(3)$, hence is a linear isomorphism. It follows that φ is a local diffeomorphism at I, hence its image $\mathrm{im}\varphi$ contains an open neighborhood of I in $\mathrm{SO}(3)$. By homogeneity, $\mathrm{im}\varphi$ is an open connected subgroup of $\mathrm{SO}(3)$, and we see that $\mathrm{im}\varphi = \mathrm{SO}(3)_e$. As $\mathrm{SO}(3)$ is connected, it follows that $\mathrm{im}\varphi = \mathrm{SO}(3)$. From this we conclude that $\varphi : \mathrm{SU}(2) \to \mathrm{SO}(3)$ is a surjective group homomorphism. Hence $\mathrm{SO}(3) \simeq \mathrm{SU}(2)/\ker\varphi$. The kernel of φ may be computed as follows. If $x \in \ker\varphi$, then $\mathrm{Ad}(x) = I$. Hence $x\mathbf{r}_j = \mathbf{r}_j x$ for $j = 1, 2, 3$. From this one sees that $x \in \{-I, I\}$. Hence $\ker\varphi = \{-I, I\}$. $\square$

It is of particular interest to understand the restriction of φ to one-parameter subgroups of $\mathrm{SU}(2)$. We first consider the one-parameter group $\alpha : t \mapsto \exp(t\mathbf{r}_1)$. Its image T in $\mathrm{SU}(2)$ consists of the matrices

$$u_t = \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}, \qquad (t \in \mathbb{R}).$$

Obviously, T is the circle group. The image of u_t under φ is given by

$$\varphi(u_t) = \varphi(e^{t\mathbf{r}_1}) = e^{\varphi_*(t\mathbf{r}_1)} = e^{2tR_1}.$$

By a simple calculation, we deduce that, for $\theta \in \mathbb{R}$,

$$R_\theta := e^{\theta R_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix},$$

the rotation with angle θ around the x_1-axis. Let D be the group consisting of these rotations. Then φ maps T onto D. Moreover, from $\varphi(u_t) = R_{2t}$ we see that φ restricts to a double covering from T onto D.

More generally, if X is any element of $\mathfrak{su}(2)$, different from 0, there exists a $x \in \mathrm{SU}(2)$ such that $\mathrm{Ad}(x^{-1})X = x^{-1}Xx = \tau\mathbf{r}_1$, for some $\tau > 0$. It follows that the one-parameter subgroup α_X has image $\exp \mathbb{R}X = \exp[\mathrm{Ad}(x)\mathbb{R}\mathbf{r}_1] = xTx^{-1}$ in $\mathrm{SU}(2)$. The image of xTx^{-1} under φ equals rDr^{-1}, with $r = \varphi(x)$. Moreover, the following diagram commutes:

$$\begin{array}{ccc} T & \xrightarrow{\mathcal{C}_x} & xTx^{-1} \\ \downarrow \varphi & & \downarrow \varphi \\ D & \xrightarrow{\mathcal{C}_r} & rDr^{-1}. \end{array}$$

The horizontal arrows being diffeomorphisms, it follows that $\varphi|_{xTx^{-1}}$ is a double covering from xTx^{-1} onto rDr^{-1}.

11 Group actions and orbit spaces

Definition 11.1 Let M be a set and G a group. A *(left) action* of G on M is a map $\alpha : G \times M \to M$ such that

(a) $\alpha(g_1, \alpha(g_2, m)) = \alpha(g_1 g_2, m) \qquad (m \in M,\ g_1, g_2 \in G);$

(b) $\alpha(e, m) = m \qquad (m \in M).$

Instead of the cumbersome notation α we usually exploit the notation $g \cdot m$ or gm for $\alpha(g, m)$. Then the above rules (a) and (b) become: $g_1 \cdot (g_2 \cdot m) = (g_1 g_2) \cdot m$, and $e \cdot m = m$.

If $g \in G$, then we sometimes use the notation α_g for the map $m \mapsto \alpha(g, m) = gm,\ M \to M$. From (a) and (b) we see that α_g is a bijection with inverse map equal to $\alpha_{g^{-1}}$. Let $\mathrm{Sym}(M)$ denote the set of bijections from M onto itself. Then $\mathrm{Sym}(M)$, equipped with the composition of maps, is a group. According to (a) and (b) the map $\alpha : g \mapsto \alpha_g$ is a group homomorphism of G into $\mathrm{Sym}(M)$. Conversely, any group homomorphism $G \to \mathrm{Sym}(M)$ comes from a unique left action of G on M in the above fashion.

Let M_1, M_2 be two sets equipped with (left) G-actions. A map $\varphi : M_1 \to M_2$ is said to *intertwine* the G-actions, or to be *equivariant,* if $\varphi(gm) = g\varphi(m)$ for all $m \in M_1$ and $g \in G$.

Remark 11.2 Similarly, a *right action* of a group G on a set M is defined to be a map $\alpha : M \times G \to M$, $(m, g) \mapsto mg$, such that $me_G = m$ and $(mg_1)g_2 = m(g_1 g_2)$ for all $m \in M$ and $g_1, g_2 \in G$. Notice that these requirements on α are equivalent to the requirement that the map $\alpha^\vee : G \times M \to M$ defined by $\alpha^\vee(g, m) = mg^{-1}$ is a left action. Thus, all results for left actions have natural counterparts for right actions.

Our goal is to study smooth actions of a Lie group on a manifold. As a first step we concentrate on continuous actions. This is most naturally done for topological groups.

Definition 11.3 A *topological group* is a group G equipped with a topology such that the multiplication map $\mu : G \times G \to G$, $(x, y) \mapsto xy$ and the inversion map $\iota : G \to G$, $x \mapsto x^{-1}$ are continuous.

Note that a Lie group is in particular a topological group.

Definition 11.4 Let G be a topological group. By a *continuous right action* of G on a topological space M we mean an action $\alpha : M \times G \to M$ that is continuous as a map between topological spaces. A (right) G-space is a topological space equipped with a continuous (right) G-action.

We assume that H is a topological group and that M is a topological space equipped with a continuous right action α of H. Given $h \in H$ we denote by α_h the map $M \to M$ given by $m \mapsto mh$. Then α_h is continuous and so is its inverse $\alpha_{h^{-1}}$. Therefore, α_h is a homeomorphism of M onto itself.

Sets of the form mH $(m \in M)$ are called *orbits* for the action α. Note that for two orbits m_1H, m_2H either $m_1H = m_2H$ or $m_1H \cap m_2H = \emptyset$. Thus, the orbits constitue a partition of M. The set of all orbits, called the *orbit space*, is denoted by M/H. The canonical projection $M \to M/H,\ m \mapsto mH$ is denoted by π.

The orbit space $X = M/H$ is equipped with the quotient topology. This is the finest topology for which the map $\pi : M \to M/H$ is continuous. Thus, a subset $\mathcal{O}$ of X is open if and only if its preimage $\pi^{-1}(\mathcal{O})$ is open in M.

In general this topology need not be Hausdorff even if M is Hausdorff. We will return to this issue later.

The following result is useful, but particular for group actions. It is not true for quotient topologies in general.

Lemma 11.5 *The natural map $\pi : M \to M/H$ is open.*

Proof: Let $U \subset M$ be open and put $\mathcal{O} = \pi(U)$. Then the preimage $\pi^{-1}(\mathcal{O})$ equals the union of the sets $Uh = \alpha_h(U)$, which are open in M. It follows that $\pi^{-1}(\mathcal{O})$ is open, hence $\mathcal{O}$ is open by definition of the quotient topology. $\square$

We denote by $\mathcal{F}(M)$ the complex linear space of functions $M \to \mathbb{C}$. Let $\mathcal{F}(M)^H$ denote the subspace of $\mathcal{F}(M)$ consisting of functions $g : M \to \mathbb{C}$ that are H-invariant, i.e., $g(mh) = g(m)$ for all $m \in M,\ h \in H$.

If $f : M/H \to \mathbb{C}$ is a function, then the pull-back of f by π, defined by $\pi^*(f) := f \circ \pi$, is a function on M that is H-invariant, i.e., it belongs to $\mathcal{F}(M)^H$. One readily verifies that π^* is a linear isomorphism from $\mathcal{F}(M/H)$ onto $\mathcal{F}(M)^H$.

Let $C(M)^H$ be the space $C(M) \cap \mathcal{F}(M)^H$ of continuous functions $M \to \mathbb{C}$ which are H-invariant.

Lemma 11.6 *The pull-back map $\pi^* : f \mapsto f \circ \pi$ maps $C(M/H)$ bijectively onto $C(M)^H$.*

Proof: Obviously π^* maps $C(M/H)$ injectively into $C(M)^H$. It remains to establish surjectivity. Let $f \in C(M)^H$. Then $f = \pi^*(g)$ for a unique function $g : M/H \to \mathbb{C}$. We must show that g is continuous. Let Ω be an open subset of $\mathbb{C}$. Then $U = f^{-1}(\Omega)$ is open in M. From the H-invariance of f it follows that U is right H-invariant. Hence $U = \pi^{-1}(\pi(U))$ and it follows that $\pi(U)$ is open in M/H. But $\pi(U) = g^{-1}(\Omega)$. Thus, g is continuous. $\square$

Remark 11.7 With exactly the same proof it follows: if X is an arbitrary topological space, then π^* maps $C(M/H, X)$ (bijectively) onto $C(M, X)^H$. In fact, the quotient topology on M/H is uniquely characterized by this property for all X.

In what follows we shall mainly be interested in actions on locally compact Hausdorff spaces. Recall that the topological space M is said to be Hausdorff if for each pair of distinct points m_1, m_2 of M there exist open neighborhoods U_j of m_j such that $U_1 \cap U_2 = \emptyset$. The space M is said to be locally compact if each point in M has a compact neighborhood. Note that in a Hausdorff space M each compact subset is closed. Moreover, if M is locally compact Hausdorff, then for every point $m \in M$ and every open neighborhood U of m there exists a compact neighborhood N of m contained in U.

Lemma 11.8 *Let M be a locally compact Hausdorff space, equipped with a continuous right action of a topological group H. Then the following assertions are equivalent.*

(a) *The orbit space M/H is Hausdorff.*

(b) *For each compact subset $C \subset M$ the set CH is closed.*

Proof: Assume (a) and let $C \subset M$ be compact. Then $\pi(C)$ is compact. As M/H is Hausdorff, it follows that $\pi(C)$ is closed. As $CH = \pi^{-1}(\pi(C))$, it follows that CH is closed.

Next, assume (b). From the fact that $\{m\}$ is compact, for $m \in M$, it follows that the orbit mH is closed. Let $x_1, x_2 \in X = M/H$ be distinct points. Select $m_j \in \pi^{-1}(x_j)$. Then $x_j = m_j H$, with $m_1 H \cap m_2 H = \emptyset$. The complement V of $m_2 H$ in M is open, right H-invariant, and contains $m_1 H$. Select an open neighborhood U_1 of m_1 in M such that $\bar{U}_1$ is compact and contained in V. Then by (a), the set $\bar{U}_1 H$ is closed and still contained in V. Its complement V_2 is open and containes $m_2 H$. Hence, $\pi(V_2)$ is open in X and contains x_2.

On the other hand, $V_1 = U_1 H$ is the union of the open sets $U_1 h$, hence open in M. Moreover, V_1 contains m_1, so that $\pi(V_1)$ is an open neighborhood of x_1 in X. Clearly, the sets V_1 and V_2 are right H-invariant and disjoint. It follows that the sets $\pi(V_1)$ and $\pi(V_2)$ are disjoint open subsets of X containing the points x_1 and x_2, respectively. This establishes the Hausdorff property. $\square$

12 Smooth actions and principal fiber bundles

Definition 12.1 Let M be a smooth manifold and H a Lie group. An action of H on M is said to be *smooth* if the action map $\alpha : M \times H \to M$, $(m, h) \mapsto mh$ is a C^∞ map of manifolds.

In the rest of this section we will always assume that M is a smooth manifold on which H has a smooth right action. We will first study smooth actions for which the quotient M/H allows a natural structure of smooth manifold.

If Ω is a smooth manifold, then H has a right action on the manifold $\Omega \times H$, given by $(x, g) \cdot h = (x, gh)$. We will say that such an action is of trivial principal fiber bundle (or trivial PFB) type.[1]

More generally, the right action of H on a manifold M is called of trivial PFB type if there exist a smooth manifold Ω and a diffeomorphism $\tau : M \to \Omega \times H$ that intertwines the H-actions. Such a map τ is called a trivialization of the action. Note that $\dim\Omega = \dim M - \dim H$.

Definition 12.2 The right action of H on M is called of principal fiber bundle (PFB) type if the following two conditions are fulfilled.

(a) Every point m of M possesses an open H-invariant neighborhood U such that the right H-action on U is of trivial PFB type.

(b) If C is a compact subset of M, then CH is closed.

[1]The terminology 'principal fiber bundle type' is not standard, but used here for purposes of exposition.

In view of Lemma 11.8, the second condition is equivalent to the condition that the quotient space M/H is Hausdorff.

We call the pair (U, τ) of condition (a) a local trivialization of the right H-space M at the point m. Clearly, if the right H-space M is of PFB type, then there exists a collection $\{(U_\alpha, \tau_\alpha) \mid \alpha \in \mathcal{A}\}$ of local trivializations such that the open sets U_α cover M. Such a covering is called a trivializing covering.

Remark 12.3 If H is a closed subgroup of a Lie group G, then the map $(h, g) \mapsto gh$, $H \times G \to G$ defines a smooth right action of H on G. At a later stage we will see that this action is of PFB-type.

If the right H-action on M is of PFB type, then the quotient M/H admits a unique natural structure of smooth manifolds. In order to understand the uniqueness, the following preliminary result about submersions will prove to be very useful.

Lemma 12.4 *Let X, Y, Z be smooth manifolds, and let $\pi : X \to Y$, $\varphi : X \to Z$ and $\psi : Y \to Z$ be maps such that the following diagram commutes*

$$\begin{array}{ccc} X & \xrightarrow{\varphi} & Z \\ {\scriptstyle\pi}\downarrow & \nearrow {\scriptstyle\psi} & \\ Y & & \end{array}$$

If φ is smooth and π a submersion, then ψ is smooth on $\pi(X)$.

Proof: The map π, being a submersion, is open. In particular, $\pi(X)$ is an open subset of Y. Let $y_0 \in \pi(X)$. Fix $x_0 \in X$ such that $\pi(x_0) = y_0$. Since π is a submersion, there exists an open neighborhood U of x_0 and a diffeomorphism $\varphi : U \to \pi(U) \times F$, with F a smooth manifold of dimension $\dim X - \dim Y$, such that $\pi = \mathrm{pr}_2 \circ \varphi$. Here pr_2 denotes the projection $\pi(U) \times F \to F$ onto the second component. Let $b = \mathrm{pr}_2 \varphi(x_0)$. Then the smooth map $\sigma : \pi(U) \to X$ defined by $\sigma(y) = \varphi^{-1}(y, b)$ satisfies $\pi \circ \sigma = I$ on $\pi(U)$ and $\sigma(y_0) = x_0$. In other words, π admits a smooth locally defined section σ with $\sigma(y_0) = x_0$. From this it follows that $\psi = \psi \circ \pi \circ \sigma = \varphi \circ \sigma$ on $\pi(U)$. Hence, ψ is smooth on $\pi(U)$. $\square$

Theorem 12.5 *Let the right H action on M be of PFB type. Then M/H carries a unique structure of C^∞-manifold (compatible with the topology) such that the canonical projection $\pi : M \to M/H$ is a smooth submersion.*

If $m \in M$, then the tangent map $T_m\pi : T_m M \to T_{\pi(m)}(M/H)$ has kernel $T_m(mH)$, the tangent space of the orbit mH at m. Accordingly, it induces a linear isomorphism from $T_m M / T_m(mH)$ onto $T_{\pi(m)}(M/H)$.

Finally, $\pi^* : f \mapsto f \circ \pi$ restricts to a bijective linear map from $C^\infty(M/H)$ onto $C^\infty(M)^H$.

Remark 12.6 It follows from the assertion on the tangent maps that the dimension of M/H equals $\dim M - \dim H$.

Proof: We will first show that the manifold structure, if it exists, is unique. Let X_j denote M/H, equipped with a manifold structure labeled by $j \in \{1, 2\}$ and assume that the projection map $\pi : M \to M/H$ is submersive for both manifold structures. We will show the identity map $I : X_1 \to X_2$ is smooth for the given manifold structures. The following diagram commutes

$$\begin{array}{ccc} & M & \\ \pi \quad \swarrow & & \searrow \quad \pi \\ X_1 & \xrightarrow{I} & X_2 \end{array}$$

Since $\pi : M \to X_1$ is a submersion and $\pi : M \to X_2$ smooth, it follows by Lemma 12.4 that $I : X_1 \to X_2$ is smooth. By symmetry of the argument it follows that the inverse to I is smooth as well. Hence, X_1 and X_2 are diffeomorphic manifolds. This establishes uniqueness of the manifold structure.

We defer the treatment of existence until the end of the proof, and will first derive the other assertions as consequences.

We first address the assertion about the tangent map of π. Let $m \in M$. Since π is a submersion, $T_m\pi : T_mM \to T_{\pi(m)}(M/H)$ is a surjective linear map, with kernel equal to the tangent space of the fiber $\pi^{-1}(\pi(m))$. This fiber equals mH. Hence $\ker T_m\pi = T_m(mH)$.

Finally, it is obvious that π^* restricts to a linear injection from $C^\infty(M/H)$ into $C^\infty(M)^H$. Let $g \in C^\infty(M)^H$. Then $g = f \circ \pi$ for a unique function $f : M \to \mathbb{C}$. Since g is smooth and π a smooth submersion, it follows by application of Lemma 12.4 that f is smooth. This establishes the surjectivity, and hence the bijectivity of π^*.

We end the proof by establishing the existence of a manifold structure on $X = M/H$ for which π becomes a smooth submersion. First of all, X is a topological space, which is Hausdorff because of Lemma 11.8.

Let $\{(U_\alpha, \tau_\alpha)\}_{\alpha \in \mathcal{A}}$ be a trivializing covering of M as above. Thus, τ_α is a diffeomorphism of U_α onto $\Omega_\alpha \times H$ which intertwines the right H-actions. Writing the manifolds Ω_α as unions of charts, we see that we may replace the trivializing covering by one for which each Ω_α equals an open subset of $\mathbb{R}^n$. We write i_α for the injection $x \mapsto (x, e)$, $\Omega_\alpha \to \Omega_\alpha \times H$ and p_α for the projection $\Omega_\alpha \times H \to \Omega_\alpha$ onto the first coordinate.

We will use the trivializing covering to define a smooth atlas of X. The map $\tau_\alpha : U_\alpha \to \Omega_\alpha \times H$ is a diffeomorphism intertwining the H-actions, hence induces a homeomorphism $\chi_\alpha : \pi(U_\alpha) \to (\Omega_\alpha \times H)/H$. The projection map p_α induces a homeomorphism of the latter space onto Ω_α, by which we shall identify. Put $V_\alpha = \pi(U_\alpha)$. Then the following diagram commutes:

$$\begin{array}{ccc} U_\alpha & \xrightarrow{\tau_\alpha} & \Omega_\alpha \times H \\ \pi \downarrow & & \downarrow p_\alpha \\ V_\alpha & \xrightarrow{\chi_\alpha} & \Omega_\alpha \end{array} \tag{14}$$

The sets V_α, for $\alpha \in \mathcal{A}$, constitute an open covering of X, and the maps $\chi_\alpha : V_\alpha \to \Omega_\alpha$ are homeomorphisms. We will show that the pairs (V_α, χ_α), for $\alpha \in \mathcal{A}$, constitute a smooth atlas. Put $\Omega^\alpha_\beta = \chi_\beta(V_\alpha \cap V_\beta)$. Then the transition map

$$\chi_{\beta\alpha} := \chi_\beta \circ \chi_\alpha^{-1}$$

is a homeomorphism from Ω_α^β onto Ω_β^α. We must show it is smooth.

The transition map $\tau_{\beta\alpha} = \tau_\beta \circ \tau_\alpha^{-1}$ is a diffeomorphism from $\Omega_\alpha^\beta \times H$ onto $\Omega_\beta^\alpha \times H$. Moreover, the diagram

$$\begin{array}{ccc} \Omega_\alpha^\beta \times H & \stackrel{\tau_{\beta\alpha}}{\longrightarrow} & \Omega_\beta^\alpha \times H \\ {\scriptstyle p_\alpha} \downarrow & & \downarrow {\scriptstyle p_\beta} \\ \Omega_\alpha^\beta & \stackrel{\chi_{\beta\alpha}}{\longrightarrow} & \Omega_\beta^\alpha \end{array}$$

commutes. As the vertical arrow represent smooth submersions, it follows by application of Lemma 12.4 that $\chi_{\beta\alpha}$ is smooth.

Let X be equipped with the structure of C^∞-manifold determined by the atlas defined above. The map π maps U_α onto V_α. Moreover, from the commutativity of the diagram (14) we see that $\pi|_{U_\alpha}$ corresponds via the horizontal diffeomorphisms χ_α and τ_α with the smooth projection p_α. Hence π is smooth and submersive on each U_α; it follows that π is a smooth submersion. $\square$

The following terminology is standard in the literature, and explains the terminology 'PFB type' used so far. We assume that X is a smooth manifold.

Definition 12.7 A *principal fiber bundle* over X with *structure group* H is a pair (P, π) consisting of a smooth right H-manifold P and a smooth map $\pi : P \to X$ with the following property. For every point $x \in X$ there exists an open neighborhood V of x in X and a diffeomorphism $\tau : \pi^{-1}(V) \to V \times H$ such that

(a) $\pi = \mathrm{pr}_V \circ \tau$ on $\pi^{-1}(V)$, where pr_V denotes the projection $V \times H \to V$;

(b) τ intertwines the right H-actions.

The manifold P is called the *total space*, X is called the *base space* of the bundle. A map τ as above is called a *local trivialization* of the bundle.

The terminology 'action of PFB type' is finally justified by the following result.

Lemma 12.8 *Let H be a Lie group.*

(a) *If $\pi : P \to X$ is a principal bundle with structure group H, then the right action of H on P is of PFB-type. Moreover, π factors through a diffeomorphism $P/H \xrightarrow{\simeq} X$.*

(b) *Conversely, if M is a smooth manifold equipped with a smooth right-action of H that is of PFB type, then $\pi : M \to M/H$ is a principal fiber bundle with structure group H.*

Proof: Assertion (a) is a straightforward consequence of Definition 12.7. Assertion (b) is easily seen from the proof of Theorem 12.5. $\square$

Example 12.9 (Frame bundle of a vector bundle) Let V be a finite dimensional real vector space of dimension k. Let $\mathrm{Hom}(\mathbb{R}^k, V)$ denote the linear space of linear maps $\mathbb{R}^k \to V$. A frame in V is defined to be an injective linear map $f : \mathbb{R}^k \to V$. The set of frames, denoted $F(V)$, is a dense open subset of $\mathrm{Hom}(\mathbb{R}^k, V)$. Let $e_1, \dots, e_k$ be the standard basis of $\mathbb{R}^k$. Then the map

$f \mapsto (f(e_1), \ldots, f(e_k))$ is a bijection from $F(V)$ onto the set of ordered bases of V. Thus, a frame may be specified by giving an ordered basis of V.

The group $H := \mathrm{GL}(k, \mathbb{R}) \simeq \mathrm{GL}(\mathbb{R}^k)$ acts on $F(V)$ from the right; indeed the action is given by $(f, a) \mapsto f \circ a$. This action is free and transitive; see the text preceding Theorem 13.5 and Proposition 15.5 for the definitions of these notions. Thus, for each $f \in F(V)$ the map $a \mapsto fa$ is a diffeomorphism from H onto $F(V)$.

Let now $p : E \to M$ be a vector bundle of rank k over a smooth manifold M. For an open subset $U \subset M$ we write $E_U := p^{-1}(U)$. Then $p : E_U \to U$ is a vector bundle over U, called the restriction of E to U. A trivialization of E over an open subset $U \subset M$ is defined to be an isomorphism $\tau : E_U \to U \times \mathbb{R}^k$ of vector bundles. For $x \in U$ we define the linear isomorphism $\tau_x : E_x \to \mathbb{R}^k$ by $\tau = (x, \tau_x)$ on E_x. Let $\mathbb{R}^k_M$ denote the trivial vector bundle $M \times \mathbb{R}^k$ over M. Then the vector bundle $\mathrm{Hom}(\mathbb{R}^k_M, E)$ has fiber $\mathrm{Hom}(\mathbb{R}^k, E_x)$ at the point $x \in M$. A trivialization (U, τ) of E induces a trivialization τ' of $\mathrm{Hom}(\mathbb{R}^k, E)$ given by $\tau'_x(T_x) = \tau_x \circ T_x$ for $T_x \in \mathrm{Hom}(\mathbb{R}^k, E_x)$.

We define $F(E)$ to be the subset $\cup_{x \in M} F(E_x)$ of $\mathrm{Hom}(\mathbb{R}^k_M, E)$. This subset is readily seen to be open; the natural map $F(E) \to M$ mapping $F(E)_x$ to x defines a sub fiber bundle of $\mathrm{Hom}(\mathbb{R}^k_M, E)$. A trivialization τ of E over U induces a trivialization τ'' of $F(E)$ over U given by $\tau''_x(f) = \tau_x \circ f$ for $f \in F(E_x)$. The group H acts from the right on each fiber $F(E_x)$. By looking at trivializations we see that these actions together constitute a smooth right action of H on $F(E)$ which turns $F(E)$ into a principal fiber bundle with structure group H.

13 Proper free actions

In this section we discuss a useful criterion for smooth actions to be of PFB type.

We recall that a continuous map $f : X \to Y$ between locally compact Hausdorff (topological) spaces X and Y is said to be *proper* if for every compact subset $C \subset Y$ the preimage $f^{-1}(C)$ is compact.

For the moment we assume that M is a locally compact Hausdorff space equipped with a continuous right action of a locally compact Hausdorff topological group H.

Definition 13.1 The action of H on M is called *proper* if $(m, h) \mapsto (m, mh)$ is a proper map $M \times H \to M \times M$.

Remark 13.2 Note that a continuous action of a compact (in particular of a finite) group is always proper.

Lemma 13.3 *The following conditions are equivalent.*

(a) *The action is proper.*

(b) *For every pair of compact subsets $C_1, C_2 \subset M$ the set $H_{C_1,C_2} := \{h \in H \mid C_1 h \cap C_2 \neq \emptyset\}$ is compact.*

Proof: Let $\varphi : M \times H \to M \times M,\ (m,h) \mapsto (m, mh)$. Assume (a) and let $C_1, C_2 \subset M$ be compact sets. Then $C_1 \times C_2$ is compact, hence $\varphi^{-1}(C_1 \times C_2)$ is a compact subset of $M \times H$. Now

$$\varphi^{-1}(C_1 \times C_2) = \{(m,h) \mid m \in C_1, mh \in C_2\},$$

hence $H_{C_1,C_2} = p_2(\varphi^{-1}(C_1 \times C_2))$, with p_2 denoting the projection $M \times H \to H$. It follows that H_{C_1,C_2} is compact. Hence (b).

Now assume that (b) holds, and let C be a compact subset of $M \times M$. Then there exist compact subsets $C_1, C_2 \subset M$ such that $C \subset C_1 \times C_2$. Now $\varphi^{-1}(C)$ is a closed subset of $\varphi^{-1}(C_1 \times C_2)$, hence it suffices to show that the latter set is compact. The latter set is clearly closed; moreover, it is contained in $C_1 \times H_{C_1,C_2}$, hence compact. $\square$

Remark 13.4 We leave it to the reader to verify that condition (b) is equivalent to the condition that $\{h \in H \mid Ch \cap C \neq \emptyset\}$ be compact, for any compact set $C \subset M$.

The action of H on M is called *free* if for all $m \in M, h \in H$ we have $mh = m \Rightarrow h = e$. From now on we assume that H is a Lie group.

Theorem 13.5 *Let M be a smooth manifold equipped with a smooth right H-action. Then the following statements are equivalent.*

(a) *the action of H on M is proper and free;*

(b) *the action of H on M is of PFB type.*

As a preparation for the proof we need the following lemma.

Lemma 13.6 *Let M be a smooth right H-manifold. If $C \subset H$ is compact, and $m \in M$ a point such that $m \notin mC$, then there exists an open neighborhood U of m in M such that $Uh \cap U = \emptyset$ for all $h \in C$.*

Proof: Since mC is compact, there exist disjoint open neighborhoods Ω_1, Ω_2 of m and mC in M. By continuity of the action and compactness of C there exists an open neighborhood U of m in Ω_1 such that $UC \subset \Omega_2$. It follows that $UC \cap U = \emptyset$. $\square$

The following lemma is the key to the proof of Theorem 13.5.

Lemma 13.7 (Slice Lemma). *Let M be a smooth manifold equipped with a smooth right H action which is proper and free. Then for each $m \in M$ there exists a smooth submanifold S of M containing m such that the map $(s,h) \mapsto sh$ maps $S \times H$ diffeomorphically onto an open H-invariant neighborhood of m in M.*

Remark 13.8 The manifold S is called a *slice* for the H-action at the point m.

Proof: Fix $m \in M$ and define the map $\alpha_m : H \to M$ by $h \mapsto mh$. By freeness of the action, this map is injective. We claim that its tangent map at e is an injective linear map $\mathfrak{h} \to T_m M$.

Given $X \in \mathfrak{h}$ we define the smooth vector field v_X on M by

$$v_X(m) = \left.\frac{d}{dt}\right|_{t=0} m \exp tX.$$

By application of the chain rule we see that

$$v_X(m) = T_e(\alpha_m)(X).$$

One readily sees that the integral curve of v_X with initial point m is given by $c : t \mapsto m \exp tX$. From $v_X(m) = 0$ it follows that c is constant; by freeness of the action this implies that $\exp tX = e$ for all $t \in \mathbb{R}$, hence $X = 0$. Thus $v_X(m) = 0 \Rightarrow X = 0$ and it follows that the linear map $T_e(\alpha_m)$ has trivial kernel, hence is injective.

We now select a linear space $\mathfrak{s} \subset T_m M$ such that $\mathfrak{s} \oplus T_e(\alpha_m)(\mathfrak{h}) = T_m M$. Moreover, we select a submanifold S' of M of dimension $\dim M - \dim H$ which has tangent space at m equal to $\mathfrak{s}$. Consider the map $\varphi : S' \times H \to M, (s,h) \mapsto sh$. Then $T_{(m,e)}\varphi : \mathfrak{s} \times \mathfrak{h} \to T_m M$ is given by $(X,Y) \mapsto X + T_e(\alpha_m)Y$, hence bijective. Replacing S' by a neighborhood (in S') of its point m we may as well assume that S' has compact closure and that there exists an open neighborhood $\mathcal{O}$ of e in H such that φ maps $S' \times \mathcal{O}$ diffeomorphically onto an open subset of M. In particular it follows that the tangent map $T_{(s,e)}\varphi$ is injective for every $s \in S'$. Using the homogeneity $\alpha_h \circ \varphi \circ (I \times r_h^{-1}) = \varphi$ for all $h \in H$ we see that φ has bijective tangent map at every point of $S' \times H$.

Let $C = H_{\bar{S}',\bar{S}'}$. Then C is a compact subset of H. Hence $C_0 = C \setminus \mathcal{O}$ is a compact subset of H, not containing e. Note that $m \notin mC_0$ by freeness of the action. Hence there exists an open subset S of S' containing m such that $S \cap Sh = \emptyset$ for all $h \in C_0$ (use Lemma 13.6).

We claim that the map φ is injective on $S \times H$. Indeed, assume $\varphi(s_1, h_1) = \varphi(s_2, h_2)$, for $s_1, s_2 \in S$, $h_1, h_2 \in H$. Then $s_2 = s_1(h_1 h_2^{-1})$, hence $h_1 h_2^{-1}$ belongs to the compact set $C = H_{\bar{S}',\bar{S}'}$. From the definition of S it follows that $h_1 h_2^{-1} \in C \setminus C_0 \subset \mathcal{O}$. From the injectivity of φ on $S' \times \mathcal{O}$ it now follows that $s_1 = s_2$ and $h_1 h_2^{-1} = e$. Hence φ is injective on $S \times H$.

Since we established already that φ has a bijective tangent map at every point of $S \times H$ it now follows that φ is a diffeomorphism from $S \times H$ onto an open subset U of M. As $\varphi(m,e) = m$, it follows that $m \in U$. Moreover, φ intertwines the H-action on $S \times H$ with the H-action on U. Therefore, U is H-invariant. $\square$

Proof of Theorem 13.5. '(a) $\Rightarrow$ (b)': Assume (a). We shall first prove that the first condition of Definition 12.2 holds. Let $m \in M$ and let S be a slice through m as in the above lemma. Then the map $\varphi : S \times H \to M$ given in the lemma is an H-equivariant diffeomorphism onto an H-invariant open neighborhood U of m in M. It follows that the inverse map $\tau = \varphi^{-1} : U \to S \times H$ is a trivialization of the H-action on U.

We now turn to the second condition of Definition 12.2. Let $C \subset M$ be compact and let x be a point in the closure of CH. Fix a compact neighborhood C' of x in M. Then there exists a sequence $(x_n)_{n \geq 1}$ in $C' \cap CH$ such that $x_n \to x$ as $n \to \infty$. Write $x_n = c_n h_n$, with $c_n \in C$

and $h_n \in H$. Then h_n is contained in $H_{C,C'}$; the latter set is compact by condition (a). By passing to subsequences if necessary, we arrive in the situation that the sequences (c_n) and (h_n) are convergent, say with limits $c \in C$ and $h \in H$, respectively. Now $x = \lim c_n h_n = ch \in CH$. It follows CH contains its closure, hence is closed. This establishes the second condition of Definition 12.2. Thus, (b) follows.

'(b) $\Rightarrow$ (a)': Assume (b) holds. To see that the action of H on M is free, let $x \in M,\ h \in H$ and assume that $xh = x$. There exists an H-invariant open neighborhood U of x on which the H-action is of trivial PFB-type. Let $\tau : U \to \Omega \times H$ be a trivialization of the action. Then from $\tau(xh) = \tau(x)$ and $\tau(xh) = \tau(x)h$ it follows that $\tau(x)h = \tau(x)$. Hence $h = e$. This establishes freeness of the action.

To see that the action of H on M is proper, let $C, C' \subset M$ be compact subsets. Then it suffices to show that $H_{C,C'} = \{h \in H \mid Ch \cap C' \neq \emptyset\}$ is compact. For every $x \in C$ there exists an H-invariant open neighborhood U_x of x on which the action is of trivial type. Moreover, there exists a compact neighborhood C_x of x contained in U_x. The interiors of the sets C_x form an open cover of C, hence contain a finite subcover, parametrized by finitely many elements $x_1, \dots, x_n \in M$. Put $C_i = C_{x_i}$, then $C \subset \cup_{i=1}^n C_i$ where C_i is contained in U_{x_i}. One easily verifies that $H_{\cup_i C_i, C'} = \cup_i H_{C_i, C'}$. Therefore it suffices to prove that $H_{C,C'}$ is compact under the assumption that C is contained in an H-invariant open set U on which the action is of trivial type. Now CH is closed, hence $C'' = CH \cap C'$ is compact and contained in U. Moreover, $H_{C,C'} = H_{C,C''}$. Thus, we may as well assume that $C' \subset U$. Using a trivializing diffeomorphism we see that we may as well assume that M is of the form $\Omega \times H$. Let D and D' be the projections of C and C' onto H, respectively. Then D and D' are compact. Moreover, $H_{C,C'}$ is a closed subset of $\{h \in H \mid Dh \cap D' \neq \emptyset\} = D^{-1}D'$. The latter set is the image of the compact set $D \times D'$ under the continuous map $H \times H \to H,\ (h_1, h_2) \mapsto h_1^{-1}h_2$, hence compact. It follows that $H_{C,C'}$ is compact as well. This establishes (a). $\square$

Example 13.9 We return to the setting of Example 12.9, with $p : E \to M$ a rank k-vector bundle. The frame bundle $\pi : F(E) \to M$ is a principal fiber bundle with structure group $H = \mathrm{GL}(k, \mathbb{R})$. Thus, the action of H on $F(E)$ is proper and free, with quotient space $F(E)/H \simeq M$.

We observe that the bundle E can be retrieved from $F(E)$ as follows. The map $\varphi : F(E) \times \mathbb{R}^k \to E$ defined by $(f, v) \mapsto f(v)$ on $F(E)_x \times \mathbb{R}^k$, for $x \in M$, is a surjective smooth map. Using trivializations of E one sees that φ is a submersion. Two elements (f_1, v_1) and (f_2, v_2) have the same image if and only if they belong to the same fiber $F(E)_x \times \mathbb{R}^k$ and there exists a $h \in H$ such that $(f_2, v_2) = (f_1 \circ h, h^{-1}v_1)$. Define the right action of H on $F(E) \times \mathbb{R}^k$ by $(f, v)a = (fh, h^{-1}v)$. Then it follows that the fibers of φ are precisely the orbits for the right action of H on $F(E) \times \mathbb{R}^k$.

Via the projection $q : F(E) \times \mathbb{R}^k \to F(E)$ we view $F(E) \times \mathbb{R}^k$ as a trivial vector bundle over $F(E)$. The map q intertwines the given right actions of H. As the action of H on $F(E)$ is proper and free, so is the action of H on $F(E) \times \mathbb{R}^k$ (argument left to the reader). It follows that the induced map $\bar{q} : (F(E) \times \mathbb{R}^k)/H \to F(E)/H = M$ is smooth (show this). Using trivializations of E, hence of $F(E)$, one readily checks that the projection $\bar{q}$ defines a smooth rank k vector bundle over M.

The map $\varphi : F(E)\times\mathbb{R}^k \to E$ defined above induces a smooth map $\bar\varphi : (F(E)\times\mathbb{R}^k)/H \to E$ (give the argument). Again using trivializations of E one checks that $\bar\varphi$ is an isomorphism of vector bundles. Thus, the vector bundle $\bar q : (F(E) \times \mathbb{R}^k)/\mathrm{GL}(k,\mathbb{R}) \to M$ is naturally isomorphic to E.

14 Coset spaces

We now consider a type of proper and free action that naturally occurs in many situations. Let G be a Lie group and H a closed subgroup. The map $(g,h) \mapsto gh$ defines a smooth right action of H on G. The associated orbit space is the coset space G/H, consisting of the right cosets $gH,\ g \in G$.

Lemma 14.1 *Let H be a closed subgroup of the Lie group G. Then the right action of H on G is proper and free.*

Proof: It is clear that the action is free. To prove it is proper, let C_1, C_2 be compact subsets of G. Then $H_{C_1,C_2} = C_1^{-1}C_2 \cap H$. Now $C_1^{-1}C_2$ is the image of $C_1 \times C_2$ under the continuous map $(x,y) \mapsto x^{-1}y$, hence compact. Moreover, H is closed, hence H_{C_1,C_2} is compact. □

Corollary 14.2 *Let G be a Lie group and H a closed subgroup. Then the coset space G/H has a unique structure of smooth manifold such that the canonical projection $\pi : G \to G/H$ is a smooth submersion. Relative to this manifold structure, the following hold.*

(a) *The map $\pi : G \to G/H$ is a principal fiber bundle with structure group H.*

(b) *The left action of G on G/H given by $(g, xH) \mapsto gxH$ is smooth.*

Proof: From Lemma 14.1 and Theorem 13.5 it follows that the right action of H on G is of PFB type. Hence, the first assertion is an immediate consequence of Theorem 12.5. Moreover, assertion (a) follows from Lemma 12.8 (b). Finally, put $X = G/H$ and let α denote the action map $G \times X \to X$. Then the following diagram commutes:

$$\begin{array}{ccc} G \times G & \xrightarrow{\mu} & G \\ \downarrow {\scriptstyle I\times\pi} & & \downarrow {\scriptstyle \pi} \\ G \times X & \xrightarrow{\alpha} & X. \end{array}$$

Since the vertical map on the left side of the diagram is a submersion, whereas μ and π are smooth, it follows that α is smooth (see Lemma 12.4). □

Corollary 14.3 *Let G be a Lie group and H a closed subgroup. The tangent map $T_e\pi$ of $\pi : G \to G/H$ is surjective and has kernel equal to $\mathfrak{h}$.*

Proof: This is an immediate consequence of the fact that π is a submersion with fiber $\pi^{-1}(eH) = H$. □

Remark 14.4 It follows from the above that the tangent map $T_e\pi$ induces a linear isomorphism from $\mathfrak{g}/\mathfrak{h}$ onto $T_{eH}(G/H)$; we agree to identify the two spaces via this isomorphism from now on. With this identification, $T_e\pi$ becomes identified with the canonical projection $\mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$.

15 Orbits of smooth actions

In this section we assume that G is a Lie group and that M is a smooth manifold equipped with a smooth left G-action α.

Given $X \in \mathfrak{g}$, we denote by α_X the smooth vector field on M defined by

$$\alpha_X(m) := \frac{d}{dt}\Big|_{t=0} (\exp tX)\, m$$

We leave it to the reader to verify that for every $m \in M$, the curve $t \mapsto (\exp tX)\, m$ is the maximal integral curve of α_X with initial point m.

Lemma 15.1 *The map $X \mapsto \alpha_X$ is a Lie algebra anti-homomorphism from $\mathfrak{g}$ into the Lie algebra $\mathcal{V}(M)$ of smooth vector fields on M.*

Proof: Fix $m \in M$, and let $\alpha_m : G \to M,\ g \mapsto gm$. Then $\alpha_X(m) = T_e(\alpha_m)X$. It follows that $X \mapsto \alpha_X(m)$ is a linear map $\mathfrak{g} \to T_mM$. This shows that $X \mapsto \alpha_X$ is a linear map $\mathfrak{g} \to \mathcal{V}(M)$.

It remains to be shown that $[\alpha_X, \alpha_Y] = \alpha_{[Y,X]}$, for all $X, Y \in \mathfrak{g}$. Since $(t, m) \mapsto (\exp tX)\, m = \alpha_{\exp tX}(m)$ is the flow of α_X, the Lie bracket of the vector fields α_X and α_Y is given by

$$\begin{aligned}
[\alpha_X, \alpha_Y](m) &= \frac{d}{dt}\Big|_{t=0} \alpha^*_{\exp tX}\alpha_Y(m) \\
&= \frac{d}{dt}\Big|_{t=0} T_{(\exp tX)\, m}(\alpha_{\exp -tX})\alpha_Y((\exp tX)\, m) \\
&= \frac{d}{dt}\Big|_{t=0} \frac{d}{ds}\Big|_{s=0} (\exp -tX)(\exp sY)(\exp tX)\, m \\
&= \frac{d}{dt}\Big|_{t=0} \frac{d}{ds}\Big|_{s=0} (\exp se^{-t\,\mathrm{ad}X}Y)m \\
&= \frac{d}{dt}\Big|_{t=0} \alpha_{e^{-t\,\mathrm{ad}X}Y}(m).
\end{aligned}$$

By linearity of $Z \mapsto \alpha_Z(m)$ it follows from this that $[\alpha_X, \alpha_Y](m) = \alpha_Z(m)$, where $Z = (d/dt)e^{-t\,\mathrm{ad}X}Y|_{t=0} = -[X, Y]$. $\square$

Remark 15.2 Right multiplication $x \mapsto r_x$ defines a right action of G on itself. The associated map $\mathfrak{g} \to \mathcal{V}(G)$ is given by the map $X \mapsto v_X$ of Lemma 3.1 and defines a linear isomorphism of $\mathfrak{g}$ onto the space $\mathcal{V}_L(G)$ of left invariant vector fields on $\mathfrak{g}$. It follows from the above that $\mathcal{V}_L(G)$ is a Lie subalgebra of $\mathcal{V}(G)$ and that $X \mapsto v_X$ is an isomorphism of Lie algebras from $\mathfrak{g}$ onto $\mathcal{V}_L(G)$.

If $x \in M$, then the *stabilizer* G_x of x in G is defined by

$$G_x = \{g \in G \mid gx = x\}.$$

Being the pre-image of x under the continuous map $g \mapsto gx$, the stabilizer is a closed subgroup of G.

Lemma 15.3 *Let $x \in M$. The Lie algebra $\mathfrak{g}_x$ of G_x is given by*

$$\mathfrak{g}_x = \{X \in \mathfrak{g} \mid \alpha_X(x) = 0\}. \tag{15}$$

Proof: Let $\mathfrak{g}_x$ denote the Lie algebra of G_x. Then for all $t \in \mathbb{R}$ we have $\exp tX \in G_x$, hence $(\exp tX)x = x$. Differentiating this expression with respect to t at $t = 0$ we see that $\alpha_X(x) = 0$. It follows that $\mathfrak{g}_x$ is contained in the set on the right-hand side of (15).

To establish the converse inclusion, assume that $\alpha_X(x) = 0$. Then $c(t) = (\exp tX)\,x$ is the maximal integral curve of the vector field α_X with initial point x. On the other hand, since $\alpha_X(x) = 0$, the constant curve $d(t) = x$ is also an integral curve. It follows that $\exp tX\,x = c(t) = d(t) = x$, hence $\exp tX \in G_x$ for all $t \in \mathbb{R}$. In view of Lemma 7.7 it now follows that $X \in \mathfrak{g}_x$. □

As G_x is a closed subgroup of G, it follows from Corollary 14.2 that the coset space G/G_x has the structure of a smooth manifold. Moreover, let $\pi : G \to G/G_x$ denote the canonical projection. Then π is a submersion, and the tangent space of G/G_x at $\bar{e} := \pi(e)$ is given by $T_{\bar{e}}(G/G_x) \simeq \mathfrak{g}/\ker T_e\pi = \mathfrak{g}/\mathfrak{g}_x$.

The map $\alpha_x : g \mapsto gx$ factors through a bijection $\bar{\alpha}_x$ of G/G_x onto the orbit Gx.

Lemma 15.4 *The map $\bar{\alpha}_x : G/G_x \to M$ is a smooth immersion.*

Proof: It follows from Corollary 14.2 that the natural projection $\pi : G \to G/G_x$ is a smooth submersion. Since $\alpha_x = \bar{\alpha}_x \circ \pi$, it follows by application of Lemma 12.4 that $\bar{\alpha}_x$ is smooth.

From $\alpha_x = \bar{\alpha}_x \circ \pi$ it follows by taking tangent maps at e and application of the chain rule that

$$T_e\alpha_x = T_{\bar{e}}(\bar{\alpha}_x) \circ T_e\pi. \tag{16}$$

Now $T_e\pi$ is identified with the canonical projection $\mathfrak{g} \to \mathfrak{g}/\mathfrak{g}_x$. Moreover, if $X \in \mathfrak{g}$, then $T_e(\alpha_x)(X) = d/dt\ \alpha_x(\exp tX)|_{t=0} = \alpha_X(x)$. Hence $\ker T_e(\alpha_x) = \mathfrak{g}_x = \ker T_e\pi$. Combining this with (16) we conclude that $T_{\bar{e}}\bar{\alpha}_x$ is injective $\mathfrak{g}/\mathfrak{g}_x \to T_xM$. Hence, $\bar{\alpha}_x$ is immersive at $\bar{e}$.

We finish the proof by applying homogeneity. For $g \in G$, let l_g denote the left action of g on G/G_x, and let α_g denote the left action of g on M. Then the maps l_g and α_g are diffeomorphisms of G/G_x and M respectively, and

$$\alpha_g \circ \bar{\alpha}_x \circ l_{g^{-1}} = \bar{\alpha}_x.$$

By taking the tangent map of both sides at $\pi(g)$ and applying the chain rule we may now conclude that $\bar{\alpha}_x$ is immersive at $\pi(g)$. □

The action of G on M is called *transitive* if it has only one orbit, namely the full manifold M. In this case the G-manifold M is said to be a *homogeneous space* for G. The following result asserts that all homogeneous spaces for G are of the form G/H with H a closed subgroup of G.

Proposition 15.5 *Let the smooth action of G on M be transitive, and let $x \in M$. Then the map $\alpha_x : G \to M,\ g \mapsto gx$ induces a diffeomorphism $G/G_x \simeq M$.*

Proof: The map $\bar{\alpha}_x : G/G_x \to M$ is a smooth immersion and a bijection. By Corollary 16.6 (see intermezzo on the Baire theorem) it must be a submersion at some point of G/G_x. By homogeneity it must be a submersion everywhere. Hence $\bar{\alpha}_x$ is a local diffeomorphism. Since $\bar{\alpha}_x$ is a bijection, we conclude that it is a diffeomorphism. $\square$

Example 15.6 Let $n \geq 0$. The special orthogonal group $\mathrm{SO}(n+1)$ acts smoothly and naturally on $\mathbb{R}^{n+1}$. Let e_1 be the first standard basis vector in $\mathbb{R}^{n+1}$. Then the orbit $\mathrm{SO}(n+1)e_1$ equals the n-dimensional unit sphere $S = S^n$ in $\mathbb{R}^{n+1}$. Since S is a smooth submanifold of $\mathbb{R}^{n+1}$, it follows that the action of $\mathrm{SO}(n+1)$ on S is smooth and transitive. The stabilizer $\mathrm{SO}(n+1)_{e_1}$ equals the subgroup consisting of $(n+1) \times (n+1)$ matrices of the form

$$\begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix}, \quad \text{with} \qquad B \in \mathrm{SO}(n).$$

It follows that S^n is diffeomorphic to $\mathrm{SO}(n+1)/\mathrm{SO}(n+1)_{e_1} \simeq \mathrm{SO}(n+1)/\mathrm{SO}(n)$.

Example 15.7 Let $n \geq 0$. We recall that n-dimensional real projective space $\mathbb{P} := \mathbb{P}^n(\mathbb{R})$ is defined to be the space of 1-dimensional linear subspaces of $\mathbb{R}^{n+1}$. It has a structure of smooth manifold, characterized by the requirement that the natural map $\pi : \mathbb{R}^{n+1} \setminus \{0\} \to \mathbb{P}$, $v \mapsto \mathbb{R}v$ is a smooth submersion.

We consider the natural smooth action of $G := \mathrm{GL}(n+1, \mathbb{R})$ on $\mathbb{R}^{n+1} \setminus \{0\}$ given by $(g, x) \mapsto gx$. Then G maps fibers of π onto fibers, hence the given action induces an action $G \times \mathbb{P} \to \mathbb{P}$. Since π is a submersion, it follows by application of Lemma 12.4 that the action of G on $\mathbb{P}$ is smooth. Let $m \in \mathbb{P}$ be the line spanned by the first standard basis vector e_1 of $\mathbb{R}^{n+1}$. Then G_m equals the group of invertible $(n+1) \times (n+1)$ matrices with first column a multiple of e_1. One readily sees that the action is transitive. Therefore, the induced map $G/G_m \to \mathbb{P}$ is a diffeomorphism of manifolds.

We now consider the subgroup $K = \mathrm{O}(n+1)$ of G. One readily sees that K already acts transitively on $\mathbb{P}$. Hence the action induces a diffeomorphism from K/K_m onto $\mathbb{P}$. Here we note that $K_m = K \cap G_m$ consists of the matrices

$$\begin{pmatrix} a & 0 \\ 0 & B \end{pmatrix},$$

with $a = \pm 1$ and $B \in \mathrm{O}(n)$. Thus, $K_m \simeq \mathrm{O}(1) \times \mathrm{O}(n)$, and we see that

$$\mathbb{P}^n(\mathbb{R}) \simeq \mathrm{O}(n+1)/(\mathrm{O}(1) \times \mathrm{O}(n)).$$

16 Intermezzo: the Baire category theorem

Let X be a topological space. A subset $A \subset X$ whose closure equals X is said to be dense. Equivalently this means that $A \cap U \neq \emptyset$ for every non-empty open subset U of X.

If $\mathcal{O}_1, \mathcal{O}_2$ are two open dense subsets of X, then $\mathcal{O}_1 \cap \mathcal{O}_2$ is still open dense. Indeed, if $U \subset X$ is open non-empty, then $U \cap \mathcal{O}_1$ is open non-empty by density of $\mathcal{O}_1$. Hence, $U \cap \mathcal{O}_1 \cap \mathcal{O}_2$ is open non-empty by density of $\mathcal{O}_2$.

It follows that the intersection of finitely many open dense subsets of X is still open and dense.

Definition 16.1 The topological space X is called a Baire space if every *countable* intersection of open dense subsets is dense.

Remark 16.2 Let X be a topological space. A subset $S \subset X$ is said to be nowhere dense in X if its closure $\bar{S}$ has empty interior. We leave it to the reader to verify that X is Baire if and only if every countable union of nowhere dense subsets of X has empty interior.

We recall that a topological space X is said to be locally compact if every point p of X is contained in a compact neighborhood C. If X is assumed to be locally compact and Hausdorff, then it is known that for every point $p \in X$ and every neighborhood N of p there exists a compact neighborhood C of p contained in N.

Theorem 16.3 (Baire category theorem) *Let X be a Hausdorff topological space. Then X is a Baire space as soon as one of the following two conditions if fulfilled.*

(a) *X is locally compact.*

(b) *There exists a complete metric on X that induces the topology of X.*

Proof: Let $\{O_k \mid k \in \mathbb{N}\}$ be a countable collection of open dense subsets of X. Let $x_0 \in X$ be any point and let U_0 be an open neighborhood of x_0. In case (a) we assume that $\overline{U}_0$ is compact, in case (b) assume that $\overline{U}_0$ is contained in a ball of radius 1. It now suffices to show that U_0 has a non-empty intersection with $\cap_{n\in\mathbb{N}} O_n$.

We will show inductively that we may select a sequence of non-empty open subsets U_k of X, for $k \in \mathbb{N}$, with the property that $\overline{U}_{k+1} \subset O_k \cap U_k$ for all $k \in \mathbb{N}$. In case (b) we will show that this can be done with the additional assumption that U_k is contained in a ball of radius $1/(k+1)$.

Suppose that $U_0, \ldots, U_n$ have been selected. Since O_n is open dense, $O_n \cap U_n \neq \emptyset$. Select a point x_{n+1} of the latter set, then in either of the cases (a) and (b) we may select an open neighborhood U_{n+1} of x_{n+1} whose closure is contained in $O_n \cap U_n$. In case (b) we may select U_{n+1} with the additional property that it is contained in the ball of radius $1/(n+2)$ around x_{n+1}.

The sequence $(\overline{U}_n)$ is a descending sequence of non-empty closed subsets of the subset U_0. At the end of the proof we will show that its intersection is non-empty. Since obviously $\cap_{n\geq} U_n$ is contained in $U_0 \cap \cap_{n\in\mathbb{N}} O_n$, it then follows that the latter intersection is non-empty.

Thus, it remains to show that the intersection of the sets $\overline{U}_n$ is non-empty. In case (b) this follows from the lemma below. In case (a), the sequence $(\overline{U}_n)$ is a decreasing sequence of closed subsets of the compact set $\overline{U}_0$. Since each finite intersection contains a set U_m, it is non-empty. Hence, by compactness, the intersection is non-empty. $\square$

Lemma 16.4 *Let (X, d) be a complete metric space and let C_k be a decreasing sequence of non-empty closed subsets of X whose diameters $d(C_k)$ tend to zero. Then $\cap_{k\in\mathbb{N}} C_k$ consists of precisely one point.*

Proof: The condition about the diameter means that we may select a ball of radius r_k containing the set C_k, for $k \in \mathbb{N}$, such that $r_k \to 0$ as $k \to \infty$. For each k we may select $x_k \in C_k$. Then $d(x_m, x_n) < 2r_k$ for all $m, n \geq k$, hence (x_n) is a Cauchy sequence. By completeness of the metric, the sequence (x_n) has a limit x.

Fix $k \in \mathbb{N}$. Let $\delta > 0$, then there exists $n \geq k$ such that $x_n \in B(x; \delta)$. Hence $B(x; \delta) \cap C_k \neq \emptyset$. It follows that x belongs to the closure of C_k, hence to C_k, for every $k \in \mathbb{N}$. Hence $x \in \cap_{k\in\mathbb{N}} C_k$. If y is a second point in the intersection, then for every k, both x, y belong to C_k, hence $d(x, y) < 2r_k$. It follows that $d(x, y) = 0$, hence $x = y$. ☐

A useful application of the above is the following result.

Proposition 16.5 *Let X be a manifold of dimension n. Let $\{Y_k \mid k \in \mathbb{N}\}$ be a countable collection of submanifolds of X of dimension strictly smaller than n. Then the union $\cup_{k\in\mathbb{N}} Y_k$ has empty interior.*

Proof: Since X is locally compact, it is a Baire space. Fix $k \in \mathbb{N}$. If $y \in Y_k$, then by the definition of submanifold, there exists an open neighborhood U_y of y in Y_k such that U_y is nowhere dense in X. By the second countability assumption for manifolds, it follows that Y_k can be covered with countably many neighborhoods $U_{k,j}$ that are nowhere dense in X. Thus, the union $\cup_{k\in\mathbb{N}} Y_k$ is the countable union of the sets $U_{k,j}$. Since all of them are nowhere dense, their union has empty interior. ☐

Corollary 16.6 *Let X and Y be smooth manifolds with* $\dim X < \dim Y$. *Let $\varphi : X \to Y$ be a smooth immersion. Then $\varphi(X)$ has empty interior.*

Proof: Put $d = \dim X$. For every $x \in X$ there exists an open neighborhood U_x of x in X such that $\varphi(U_x)$ is a smooth submanifold of Y of dimension d. By the second countability assumption there exists a countable covering of X by open subsets U_k of X such that $\varphi(U_k)$ is a smooth submanifold of Y of dimension d. It follows that $\varphi(X) = \cup_{k\in\mathbb{N}} \varphi(U_k)$ has empty interior. ☐

17 Normal subgroups and ideals

If G is a Lie group and H a closed subgroup, then the coset space G/H is a smooth manifold in a natural way. If H is a normal subgroup, i.e., $gHg^{-1} = H$ for all $g \in G$, then G/H is a group as well. The following result asserts that these structures are compatible and turn G/H into a Lie group.

Proposition 17.1 *Let G be a Lie group and H a closed normal subgroup. Then G/H has a unique structure of Lie group such that the canonical map $\pi : G \to G/H$ is a homomorphism of Lie groups.*

Proof: We equip G/H with the unique manifold structure for which π is a submersion. Since H is normal, G/H has a unique group structure such that π is a group homomorphism. Let $\bar{\mu}$ denote the multiplication map of the quotient group G/H. Then the following diagram commutes

$$\begin{array}{ccc} G \times G & \xrightarrow{\mu} & G \\ \downarrow \pi\times\pi & & \downarrow \pi \\ G/H \times G/H & \xrightarrow{\bar{\mu}} & G/H \end{array}$$

Since μ and π are smooth, so is $\pi \circ \mu$. Since the left vertical map is a submersion, it follows from Lemma 12.4 that $\bar{\mu}$ is smooth. In a similar fashion it follows that the inversion map of G/H is smooth. Hence G/H is a Lie group, and π is a Lie group homomorphism.

Suppose that G/H is equipped with a second structure of Lie group such that $\pi : G \to G/H$ is a Lie group homomorphism. We shall denote G/H, equipped with this structure of Lie group, by $(G/H)'$. The identity map $I : G/H \to (G/H)'$ clearly is an injective homomorphism of groups. Since π is a submersion, it follows by application of Lemma 12.4 that I is smooth, hence a Lie group homomorphism. Since I is injective, it follows by Lemma 7.6 that I is immersive everywhere. Hence, by Lemma 17.2 below we see that I is a submersion. Thus, I is a bijective local diffeomorphism, hence a diffeomorphism. Therefore, I is an isomorphism of Lie groups, establishing the uniqueness. $\square$

Lemma 17.2 *Let $\varphi : G \to G'$ be an immersive homomorphism of Lie groups. Then φ is a submersion if and only if $\varphi(G)$ is an open subgroup of G'.*

Remark 17.3 In Proposition 17.7 we will see that the assumption that φ be immersive is superfluous.

Proof: If φ is a submersion, then $\varphi(G)$ is open in G'. Conversely, assume that $\varphi(G)$ is open in G'. Then it follows by Corollary 16.6 that $\dim G = \dim G'$. Hence $T_e\varphi : T_eG \to T_eG'$ is an injective linear map between spaces of equal dimension. Therefore, it is surjective as well. By homogeneity it follows that φ is a submersion everywhere. $\square$

Theorem 17.4 (The isomorphism theorem for Lie groups). *Let $\varphi : G \to G'$ be a homomorphism of Lie groups. Then $H := \ker \varphi$ is a closed normal subgroup of G. Moreover, the induced homomorphism $\bar{\varphi} : G/H \to G'$ is a smooth injective immersion. If φ is surjective, then $\bar{\varphi}$ is an isomorphism of Lie groups.*

Proof: The following diagram is a commutative diagram of group homomorphisms

$$\begin{array}{ccc} G & \xrightarrow{\varphi} & G' \\ \downarrow \pi & \nearrow \bar{\varphi} & \\ G/H & & \end{array}$$

Moreover, since π is a submersion, whereas φ is smooth, it follows that $\bar{\varphi}$ is smooth. Hence $\bar{\varphi}$ is an injective homomorphism of Lie groups. It follows by Lemma 7.6 that φ is an immersion. Now assume that φ is surjective. Then $\bar{\varphi}$ is surjective, and it follows by application of Lemma 17.2 that $\bar{\varphi}$ is a submersion. We conclude that $\bar{\varphi}$ is a local diffeomorphism, hence a diffeomorphism, hence an isomorphism of Lie groups. $\square$

Example 17.5 The isomorphism of Proposition 10.3 is an isomorphism of Lie groups.

Example 17.6 Let G be a Lie group. Then Ad is a Lie group homomorphism from G into $\mathrm{GL}(\mathfrak{g})$. It induces an injective Lie group homomorphism $G/\ker \mathrm{Ad} \to \mathrm{GL}(\mathfrak{g})$, realizing the image $\mathrm{Ad}(G)$ as a Lie subgroup of $\mathrm{GL}(\mathfrak{g})$. If G is connected, then $\ker \mathrm{Ad}$ is the center $Z(G)$ of G, see exercises. Consequently, $\mathrm{Ad}(G) \simeq G/Z(G)$ in this case.

Proposition 17.7 *Let $\varphi : G \to G'$ be a homomorphism of Lie groups. Then $\varphi(G)$ is an open subgroup of G' if and only if φ is submersive.*

Proof: If φ is submersive, then $\varphi(G)$ is open. Thus, it remains to prove the 'only if' statement. Let H be the kernel of φ. Then by Theorem 17.4 it follows that the induced map $\bar{\varphi} : G/H \to G'$ is an injective homomorphism of Lie groups. By application of Lemma 17.2 it follows that $\bar{\varphi}$ is a submersion. Since $\bar{\varphi} = \varphi \circ \pi$, whereas π is surjective, we now deduce that φ is a submersion everywhere. $\square$

We end this section with a discussion of the Lie algebra of a quotient of a Lie group by a closed normal subgroup.

Definition 17.8 Let $\mathfrak{l}$ be a Lie algebra. An *ideal* of $\mathfrak{l}$ is by definition a linear subspace $\mathfrak{a}$ of $\mathfrak{l}$ such that $[\mathfrak{l}, \mathfrak{a}] \subset \mathfrak{a}$, i.e. $[X, Y] \in \mathfrak{a}$ for all $X \in \mathfrak{l}, Y \in \mathfrak{a}$.

Remark 17.9 Note that an ideal is always a Lie subalgebra.

Lemma 17.10 (a) *Let $\mathfrak{l}$ be a Lie algebra, $\mathfrak{a} \subset \mathfrak{l}$ an ideal. Then the quotient (linear) space $\mathfrak{l}/\mathfrak{a}$ has a unique structure of Lie algebra such that the canonical projection $\pi : \mathfrak{l} \to \mathfrak{l}/\mathfrak{a}$ is a homomorphism of Lie algebras.*

(b) *Let $\varphi : \mathfrak{l} \to \mathfrak{l}'$ be a homomorphism of Lie algebras, with kernel $\mathfrak{a}$. Then $\mathfrak{a}$ is an ideal in $\mathfrak{l}$ and φ factors through an injective homomorphism of Lie algebras $\bar{\varphi} : \mathfrak{l}/\mathfrak{a} \to \mathfrak{l}'$.*

Proof: Left as an exercise for the reader. $\square$

Lemma 17.11 *Let G be a Lie group and let $\mathfrak{h}$ be a subalgebra of its Lie algebra $\mathfrak{g}$.*

(a) *$\mathfrak{h}$ is an ideal if and only if $\mathfrak{h}$ is invariant under $\mathrm{Ad}(G_e)$.*

Let H be a Lie subgroup of G with Lie algebra $\mathfrak{h}$.

(b) *If H is normal in G, then $\mathfrak{h}$ is an ideal in $\mathfrak{g}$.*

(c) *If $\mathfrak{h}$ is an ideal in $\mathfrak{g}$, then H_e is normal in G_e.*

Proof: Left to the reader. Use Lemmas 4.3 and 4.6. $\square$

Lemma 17.12 *Let H be a closed subgroup of the Lie group G. If H is normal, then its Lie algebra $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. Moreover, the tangent map at e of the canonical projection $\pi : G \to G/H$ induces an isomorphism from the quotient Lie algebra $\mathfrak{g}/\mathfrak{h}$ onto the Lie algebra of the Lie group G/H.*

Proof: Let $\mathfrak{l} = T_{\pi(e)}(G/H)$ be equipped with the Lie algebra structure induced by the Lie group structure of G/H. The tangent map π_* of the canonical projection $\pi : G \to G/H$ is a Lie algebra homomorphism from $\mathfrak{g}$ onto $\mathfrak{l}$. On the other hand, its kernel is $\mathfrak{h}$. Hence, by Lemma 17.10, π_* factors through a Lie algebra isomorphism from $\mathfrak{g}/\mathfrak{h}$ onto $\mathfrak{l}$. $\square$

Remark 17.13 Accordingly, if G is a Lie group and H a closed normal subgroup, then we shall identify the Lie algebra of G/H with $\mathfrak{g}/\mathfrak{h}$ via the isomorphism described above. In this fashion, π_* becomes the canonical projection $\mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$.

Corollary 17.14 *Let $\varphi : G \to G'$ be a homomorphism of Lie groups, with kernel H.*

(a) *The induced map $\bar{\varphi} : G/H \to G'$ is a homomorphism of Lie groups.*

(b) *Put $\varphi_* = T_e\varphi$. Then $\ker \varphi_*$ equals the Lie algebra $\mathfrak{h}$ of H.*

(c) *The tangent map $\bar{\varphi}_* = T_{\bar{e}}(\bar{\varphi})$ is the linear map $\mathfrak{g}/\mathfrak{h} \to \mathfrak{g}'$ induced by φ_*.*

(d) *If φ is surjective, then $\bar{\varphi}$ and $\bar{\varphi}_*$ are isomorphisms.*

Proof: Assertion (a) follows by application of Theorem 17.4. Let $\pi : G \to G/H$ be the canonical projection. Then π is a homomorphism and a smooth submersion. By the preceding remark, its tangent map π_* is identified with the natural projection $\mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$. From $\bar{\varphi} \circ \pi = \varphi$ it follows by differentiation at e and application of the chain rule that $\bar{\varphi}_* \circ \pi_* = \varphi_*$. Since $\bar{\varphi}$ is a smooth immersion by Theorem 17.4, it follows that $\ker \varphi_* = \ker \pi_* = \mathfrak{h}$. Hence, (b). From $\bar{\varphi}_* \circ \pi_* = \varphi_*$ we also deduce (c). If φ is surjective, then $\bar{\varphi}$ is an isomorphism of Lie groups by Theorem 17.4. Hence, $\bar{\varphi}_*$ is an isomorphism of Lie algebras. $\square$

18 Detour: actions of discrete groups

Let H be a group (without additional structure) acting on a topological space M from the right by continuous transformations. Equivalently, this means that the action map $M \times H \to M$ is continuous relative to the discrete topology on H (the topology for which all subsets of H are open).

The action of H on M is said to be *properly discontinuous* if for each $m \in M$ there exists an open neighborhood U of m such that $Uh \cap U = \emptyset$ for all $h \in H \setminus \{e\}$. This condition amounts to saying that the action of H is locally of trivial PFB type relative to the discrete topology on H. A third equivalent way of phrasing the condition is that the action of H is free and that the canonical projection $\pi : M \to M/H$ is a covering map.

Now assume that M is locally compact and Hausdorff. Then in the setting of an action by diffeomorphisms on a smooth manifold M the following result may be viewed as a consequence of Theorem 13.5. We will give a direct proof to cover the topological setting.

Lemma 18.1 *Let M be a locally compact Hausdorff space equipped with a right H-action by continuous transformations. Then the following assertions are equivalent.*

(a) *The action of H on M is continuous, proper and free for the discrete topology on H.*

(b) *The action of H on M is properly discontinuous and the associated quotient space M/H is Hausdorff.*

Proof: Assume (a), and fix $m \in M$. Then there exists a compact neighborhood N of m. The set $H_{N,N}$ of $h \in H$ with $Nh \cap N \neq \emptyset$ is compact in H, hence finite. Put $C = H_{N,N} \setminus \{e\}$. For every $h \in C$ we may select an open neighborhood $U_h \ni m$ such that $U_h h \cap U_h = \emptyset$ (observe that $mh \neq m$ by freeness and use continuity of the action). Let U be the intersection of the finite collection of open sets U_h $(h \in C)$ with the interior of N. Then U is open and $Uh \cap U = \emptyset$ for all $h \in H \setminus \{e\}$. It follows that the action of H is properly discontinuous. By the same argument as in the proof of Theorem 13.5 it follows that CH is closed for every compact subset $C \subset M$. By Lemma 11.8 we conclude that G/H is Hausdorff.

Next assume (b), and let H be equipped with the discrete topology. We will first show that the action map $\alpha : M \times H \to M$ is continuous. Let $U \subset M$ be open. Then for each $h \in H$ the set Uh^{-1} is open in M. Hence $Uh^{-1} \times \{h\}$ is open in $M \times H$. The preimage $\alpha^{-1}(U)$ equals the union of these sets for $h \in H$, hence is open.

Now suppose that C_1, C_2 are compact subsets of M. We will show that the set

$$H_{C_1,C_2} = \{h \in H \mid C_1 h \cap C_2 \neq \emptyset\}$$

is finite hence compact; from this (a) will follow.

For every $m \in M$ there exists an open neighborhood U_m of m such that $U_m \cap U_m h = \emptyset$ for $h \in H, h \neq e$. It follows that $U_m h_1 \cap U_m h_2 = \emptyset$ for distinct $h_1, h_2 \in H$. Let N_m be a compact neighborhood of m contained in U_m. By compactness there exists a finite collection F of points from C_1 such that $C_1 \subset \cup_{m \in F} U_m$. It follows that H_{C_1,C_2} is contained in the finite union $\cup_{m \in F} H_{N_m,C_2}$. Therefore, it suffices to show that the set $H_m := H_{N_m,C_2}$ is finite.

It follows from Lemma 11.8 that $N_m H$ is closed. The complement $U_0 := M \setminus N_m H$ is open in M. The sets U_0 and $U_m h,\ h \in H_m$, form an open cover of C_2. Hence, there exists a finite subset $S \subset H_m$ such that C_2 is contained in the union of U_0 and $\cup_{h \in S} N_m h$. Let $h \in H_m$. Then $N_m h \cap C_2$ is non-empty; let c be one of its points. As $c \notin U_0$ there exists a $h' \in S$ such that $c \in U_m h'$. From $c \in N_m h \cap U_m h' \subset U_m h \cap U_m h'$ it follows that the intersection $U_m h \cap U_m h'$ is non-empty. Hence $h = h'$ and we see that $h \in S$. We conclude that H_m is contained in S hence is finite. $\square$

19 Densities and integration

If V is an n-dimensional real linear space, then a *density* on V is a map $\omega : V^n \to \mathbb{C}$ transforming according to the rule:

$$T^*\omega := \omega \circ T^n = |\det T|\, \omega \qquad (T \in \mathrm{End}(V)).$$

In these notes the (complex linear) space of densities on V is denoted by $\mathcal{D}V$. A density $\omega \in \mathcal{D}V$ is called *positive* if it is non-zero and has values in $[0, \infty[$. The set of such densities is denoted by $\mathcal{D}_+V$. It is obviously non-empty.

Example 19.1 If ω is an element of $\wedge^n V^*$, the space of alternating multilinear maps $V^n \to \mathbb{R}$, then $|\omega|$ is a positive density on V.

If φ is a linear isomorphism from V onto a real linear space W, then the map $\varphi^* : \omega \mapsto \omega \circ \varphi^n$ is a linear isomorphism $\mathcal{D}W \to \mathcal{D}V$ of the associated spaces of densities. Indeed, if $\omega \in \mathcal{D}W$, $T \in \mathrm{End}(V)$, then

$$\begin{aligned}
[\varphi^*\omega] \circ T^n = \omega \circ [\varphi \circ T]^n &= (\omega \circ [\varphi \circ T \circ \varphi^{-1}]^n) \circ \varphi^n \\
&= |\det(\varphi \circ T \circ \varphi^{-1})|\, \varphi^*\omega = |\det T|\, \varphi^*\omega.
\end{aligned}$$

Note that φ^* maps $\mathcal{D}_+W$ onto $\mathcal{D}_+V$.

The space $\mathcal{D}V$ is one dimensional; in fact, if $v_1, \ldots, v_n$ is a basis of V then the map $\omega \mapsto \omega(v_1, \ldots, v_n)$ is a linear isomorphism from $\mathcal{D}V$ onto $\mathbb{C}$, mapping $\mathcal{D}_+V$ onto $]0, \infty[$.

If X is a smooth manifold, then by T_xX we denote the tangent space of X at a point x. By a well known procedure we may define the *bundle $\mathcal{D}TX$ of densities* on X; it is a complex line bundle with fiber $(\mathcal{D}TX)_x \simeq \mathcal{D}(T_xX)$. The space of continuous sections of $\mathcal{D}TX$ is denoted by $\Gamma(\mathcal{D}TX)$; this space is called the space of *continuous densities* on X. The space of smooth densities on X is denoted by $\Gamma^\infty(\mathcal{D}TX)$. We also have the fiber bundle $\mathcal{D}_+TX$ of positive densities on X. Its fiber above x equals $\mathcal{D}_+T_xX$. Its continuous sections are called the *positive continuous densities* on X.

If φ is a diffeomorphism of X onto a manifold Y, then we define the (pull-back) map $\varphi^* : \Gamma(\mathcal{D}TY) \to \Gamma(\mathcal{D}TX)$ by

$$(\varphi^*\omega)(x) = D\varphi(x)^*\omega(\varphi(x)).$$

Note that φ^* maps positive densities to positive densities.

Let $e_1, \ldots, e_n$ be the standard basis of $\mathbb{R}^n$. The density $\lambda \in \mathcal{D}\mathbb{R}^n$ given by $\lambda(e_1, \ldots, e_n) = 1$ is called the standard density on $\mathbb{R}^n$. Let $U \subset \mathbb{R}^n$ be an open subset. Then by triviality of the tangent bundle $TU \simeq U \times \mathbb{R}^n$, the map $f \mapsto f\lambda$ defines a linear isomorphism from $C^\infty(U)$ onto $\Gamma^\infty(\mathcal{D}TU)$. If f belongs to the space $C_c(U)$ of compactly supported continuous functions $U \to \mathbb{C}$, we define the integral

$$\int_U f\lambda := \int_{\mathbb{R}^n} f(x)\, dx,$$

where dx denotes normalized Lebesgue measure. If φ is a diffeomorphism from U onto a second open subset $V \subset \mathbb{R}^n$, then, for $g \in C_c(V)$, we have $\varphi^*(g\lambda)(x) = g(\varphi(x))\, |\det D\varphi(x)|\, \lambda(\varphi(x))$. Thus, by the substitution of variables theorem:

$$\int_U \varphi^*\omega = \int_V \omega \qquad (\omega \in \Gamma_c(\mathcal{D}TV)). \tag{17}$$

Let now (Ω, χ) be a coordinate chart of X. If ω is a continuous density on X with compact support $\operatorname{supp} \omega \subset \Omega$, then we define

$$\int_X \omega := \int_{\chi(\Omega)} (\chi^{-1})^* \omega.$$

This definition is unambiguous, because if (Ω', χ') is a second chart such that $\operatorname{supp} \omega \subset \Omega'$, then

$$\int_{\chi'(\Omega')} (\chi')^{-1*} \omega = \int_{\chi(\Omega)} (\chi' \circ \chi^{-1})^* (\chi')^{-1*} \omega = \int_{\chi(\Omega)} (\chi^{-1})^* \omega$$

by the substitution of variables theorem.

We can now define the *integral of a compactly supported density* on the manifold X as follows. Let $\{\Omega_\alpha \mid \alpha \in \mathcal{A}\}$ be an open cover of the manifold X with coordinate neighborhoods. Then there exists a partition of unity $\{\psi_\alpha \mid \alpha \in \mathcal{A}\}$ subordinate to this cover. We recall that the ψ_α are functions in $C_c^\infty(X)$ with $0 \leq \psi_\alpha \leq 1$. Moreover, the collection of supports $\{\operatorname{supp} \psi_\alpha\}$ is locally finite and $\sum_{\alpha \in \mathcal{A}} \psi_\alpha = 1$ on X (note that the sum is finite at every point of X, by the local finiteness of the collection of supports). Let $\omega \in \Gamma_c(\mathcal{D}TX)$ be a continuous density on X with compact support. Then we define

$$\int_X \omega = \sum_{\alpha \in \mathcal{A}} \int_{\Omega_\alpha} \psi_\alpha \omega.$$

Just as in the theory of integration of differential forms one shows that this definition is independent of the particular choice of partition of unity. Note that integration of forms is oriented, whereas the present integration of densities is non-oriented.

We note that $\omega \mapsto \int_X \omega$ is a linear map $\Gamma_c \mathcal{D}TX \to \mathbb{C}$. Moreover, the following lemma is an easy consequence of the definitions (reduction to charts et cetera).

Lemma 19.2 *Let ω be a positive density on X. Then for every $f \in C_c(X)$ with $f \geq 0$ everywhere we have $\int_X f\omega \geq 0$. Moreover, $\int_X f\omega = 0 \Rightarrow f = 0$.*

Also, by a straightforward reduction to charts we can prove the following *substitution of variables theorem*.

Proposition 19.3 *Let $\varphi : X \to Y$ be a diffeomorphism of C^∞-manifolds. Then for every $\omega \in \Gamma_c(\mathcal{D}TY)$ we have:*

$$\int_X \varphi^* \omega = \int_Y \omega.$$

We now turn to the situation that G is a Lie group acting smoothly from the left on a smooth manifold M. If $g \in G$, we write l_g for the diffeomorphism $M \to M,\ m \mapsto gm$.

Definition 19.4 A density $\omega \in \Gamma(\mathcal{D}TM)$ is said to be G-*invariant* if $l_g^* \omega = \omega$ for all $g \in G$. The space of G-invariant continuous densities on M is denoted by $\Gamma(\mathcal{D}TM)^G$.

The following result will be very important for applications.

Lemma 19.5 *Let ω be a G-invariant continuous density on M. Then for every $f \in C_c(M)$ and all $g \in G$ we have:*

$$\int_M l_g^*(f)\,\omega = \int_M f\,\omega. \tag{18}$$

Here $l_g^* f := f \circ l_g$.

Proof: We note that by invariance of ω we have $l_g^*(f)\omega = l_g^*(f)l_g^*(\omega) = l_g^*(f\omega)$. Now observe that l_g is a diffeomorphism of M and apply the substitution of variables theorem (Proposition 19.3) with $\varphi = l_g$. $\square$

Lemma 19.6 *Let G be a Lie group and let $\Gamma(\mathcal{D}TG)^G$ denote the space of left invariant continuous densities on G.*

(a) *The evaluation map $\epsilon : \omega \mapsto \omega(e)$ defines a linear isomorphism $\Gamma(\mathcal{D}TG)^G \xrightarrow{\simeq} \mathcal{D}\mathfrak{g}$.*

(b) *A density $\omega \in \Gamma(\mathcal{D}TG)^G$ is positive if and only if $\omega(e)$ is positive.*

Proof: The map ϵ is linear. If ω is a left invariant continuous density on G, then $\omega(g) = ((l_g^{-1})^*\omega)(g) = T_e(l_g)^{-1*}\omega(e)$ for all $g \in G$. Hence, ϵ has trivial kernel. On the other hand, if ω_0 is a density in $\mathcal{D}\mathfrak{g}$ then the formula

$$\omega(g) := T_e(l_g)^{-1*}\omega_0 \tag{19}$$

defines a continuous density on G whose value at e is ω_0. By application of the chain rule for tangent maps it follows that this density is left G-invariant. Thus, (a) follows. Assertion (b) follows from (19). $\square$

The following result is an immediate consequence of the above lemma.

Corollary 19.7 *Every Lie group G has a left (resp. right) invariant positive density. Two such densities differ by a positive factor.*

If ω is a density on G, then the map $C_c(G) \to \mathbb{R}$, $f \mapsto I(f) = \int_G f\omega$ is continuous linear, hence a *Radon measure* on G. For this reason we shall often write dx for an invariant density on G, and $\int_G f(x)\, dx$ for the associated invariant integral of a function $f \in C_c(G)$. Note that in the example of $G = \mathbb{R}^n$ with addition, dx is a (complex) multiple of Lebesgue measure. Positivity then means that the multiple is positive, and invariance corresponds with translation invariance of the Lebesge measure.

We now recollect some of the above results in the present notation. Let dx be a left invariant positive density on G. (Analogous statements will be valid for right invariant positive densities.)

Proposition 19.8 *The map $f \mapsto I(f) = \int_G f(x)\, dx$ is a complex linear functional on $C_c(G)$. It satisfies the following, for every $f \in C_c(G)$.*

(a) *If f is real then so is $I(f)$; if $f \geq 0$ then $I(f) \geq 0$.*

(b) *If $f \geq 0$ and $I(f) = 0$ then $f = 0$.*

(c) *For every $y \in G$:*

$$\int_G f(yx)\, dx = \int_G f(x)\, dx. \tag{20}$$

Proof: Assertion (a) follows from the positivity of ω. Assertion (b) is immediate from Lemma 19.2. Finally (c) is a reformulation of Lemma 19.5. $\square$

Remark 19.9 One can show that up to a positive factor the linear functional I is uniquely determined by the requirement $I \neq 0$ and the properties (a) and (c). In particular property (b) is a consequence. For details we refer the reader to the book by Bröcker and tom Dieck.

It follows from the proposition that the Radon measure associated with a left invariant density is left invariant, non-trivial and positive.

In the literature a left G-invariant positive Radon measure on G is called a left *Haar measure* of G. The above statement about the uniqueness of I is referred to as 'uniqueness of the Haar measure.' More generally a left (resp. right) Haar measure exists (and is unique up to a positive factor) for any locally compact topological group. Of course one cannot use the present differential geometric method of proof to establish the existence and uniqueness result in that generality.

Lemma 19.10 *Let G be a compact Lie group. Then there exists a unique left invariant density dx on G with*

$$\int_G dx = 1.$$

This density is positive.

Proof: Fix a positive density λ on G. Then it follows from assertions (a) and (b) of Proposition 19.8 for $f = 1$, that $\int_G \lambda$ equals a positive constant $c > 0$. The densitity $dx = c^{-1}\lambda$ satifies the above. This proves existence. If ω is a density with the same property, then $\omega = C dx$ for a constant $C \in C$. Integration over G shows that $C = 1$. This establishes uniqueness. $\square$

Remark 19.11 The density of the above lemma is called the normalized left invariant density of G. The associated Haar measure is called *normalized Haar measure.*

The following result expresses how left invariant densities behave under right translation.

Lemma 19.12 *Let dx be a left invariant density on a Lie group G. Then for every $g \in G$,*

$$r_g^*(dx) = |\mathrm{detAd}(g)|^{-1}\, dx.$$

Proof: Without loss of generality we may assume that dx is non-zero. For $g, h \in G$ we have that $l_h r_g = r_g l_h$, hence $r_g^* l_h^* = l_h^* r_g^*$ and we see that $l_h^*(r_g^*(dx)) = r_g^*(l_h^* dx) = r_g^*(dx)$. It follows that $r_g^*(dx)$ is a left invariant positive density. This implies that $r_g^*(dx) = c\, dx$ for a non-zero constant c. Applying $l_{g^{-1}}^*$ to both sides of this equation we find $\mathcal{C}_{g^{-1}}^*\, dx = c\, dx$. Evaluating both sides of the latter identity in e we obtain

$$c\, dx(e) = T_e(\mathcal{C}_{g^{-1}})^* dx(e) = \mathrm{Ad}(g^{-1})^* dx(e) = |\det \mathrm{Ad}(g^{-1})|\, dx(e).$$

It follows that $c = |\det \mathrm{Ad}(g)|^{-1}$. $\square$

A Lie group G with $|\det \mathrm{Ad}(g)| = 1$ for all $g \in G$ is said to be *unimodular*. The following result is an immediate consequence of the lemma.

Corollary 19.13 *Let G be a unimodular Lie group. Then every left invariant density is also right invariant.*

Lemma 19.14 *Let G be a compact Lie group. Then G is unimodular.*

Remark 19.15 It follows that the normalized Haar measure of a compact Lie group is bi-invariant.

Proof: The map $x \mapsto |\det \mathrm{Ad}(x)|$ is a continuous group homomorphism from G into the group $(\mathbb{R}_+, \cdot)$ of positive real numbers equipped with multiplication. Its image H is a compact subgroup of $(\mathbb{R}_+, \cdot)$. Now apply the lemma below to conclude that $H = \{1\}$. $\square$

Lemma 19.16 *The only compact subgroup of* $(\mathbb{R}_+, \cdot)$ *is* $\{1\}$.

Proof: Let Γ be a compact subgroup of $(\mathbb{R}_+, \cdot)$. Then $1 \in \Gamma$. Assume Γ contains an element $\gamma > 0$ different from 1. Since Γ contains both γ and γ^{-1} we may as well assume that $\gamma > 1$. The sequence $(\gamma^n)_{n \geq 1}$ belongs to Γ and is unbounded from above, contradicting the compactness of Γ. It follows that $\Gamma = \{1\}$. $\square$

We shall now investigate the existence of invariant densities on homogeneous spaces for G. According to Proposition 15.5 such a space is of the form $X = G/H$, with H a closed subgroup of G. Here G acts on X by left translation. For $g \in G$ we write $l_g : X \to X, xH \mapsto gxH$.

The tangent map at e of the canonical projection $\pi : G \to G/H$ induces a linear isomorphism $\mathfrak{g}/\mathfrak{h} \simeq T_{eH}(G/H)$ by which we identify. If $h \in H$, then $\mathcal{C}_h : G \to G$, $g \mapsto hgh^{-1}$ leaves H invariant. Differentiation at e gives that $\mathrm{Ad}(h)$ leaves the subspace $\mathfrak{h}$ of $\mathfrak{g}$ invariant, hence induces a linear automorphism $A(h)$ of the quotient space $\mathfrak{g}/\mathfrak{h}$. The following lemma will be useful in the sequel.

Lemma 19.17 *Let $h \in H$. Then the tangent map of $l_h : G/H \to G/H, xH \mapsto hxH$ at e is given by*

$$T_{eH}(l_h) = A(h).$$

Proof: Let $h \in H$. Recall that $\mathcal{C}_h : G \to G, x \mapsto hxh^{-1}$ has tangent map $\mathrm{Ad}(h)$ at e. We note that $\pi \circ \mathcal{C}_h = l_h \circ \pi$. Differentiating at e and applying the chain rule we find $T_e\pi \circ \mathrm{Ad}(h) = T_{eH}(l_h) \circ T_e\pi$. It follows from this that $T_{eH}(l_h)$ is the endomorphism of $\mathfrak{g}/\mathfrak{h}$ induced by $\mathrm{Ad}(h)$. $\square$

The fiber of the bundle $\mathcal{D}T(G/H)$ over eH is identified with $\mathcal{D}(\mathfrak{g}/\mathfrak{h})$. Thus $A(h)^*$ is an automorphism of the associated space of densities $\mathcal{D}(\mathfrak{g}/\mathfrak{h})$. Note that for $\omega \in \mathcal{D}(\mathfrak{g}/\mathfrak{h})$ we have:

$$A(h)^*\omega = |\det A(h)|\,\omega = \frac{|\det \mathrm{Ad}(h)|_{\mathfrak{g}}|}{|\det \mathrm{Ad}(h)|_{\mathfrak{h}}|}\,\omega. \tag{21}$$

We write $\mathcal{D}(\mathfrak{g}/\mathfrak{h})^H$ for the linear space of densities ω on $\mathfrak{g}/\mathfrak{h}$ satisfying $A(h)^*\omega = \omega$. Such densities are called H-invariant. Since $\mathcal{D}(\mathfrak{g}/\mathfrak{h})$ is one dimensional, the space of H-invariant densities is either 0 or 1 dimensional. In view of (21) the latter is the case if and only if $|\det\mathrm{Ad}(h)|_{\mathfrak{g}}| = |\det\mathrm{Ad}(h)|_{\mathfrak{h}}|$ for all $h \in H$.

Lemma 19.18

(a) *The evaluation map $\epsilon : \omega \mapsto \omega(eH)$ defines a bijection from $\Gamma(\mathcal{D}T(G/H))^G$ onto $\mathcal{D}(\mathfrak{g}/\mathfrak{h})^H$. This bijection maps positive densities onto positive densities.*

(b) *The space of G-invariant densities on G/H is at most one dimensional. It is one dimensional if and only if*

$$|\det \mathrm{Ad}(h)|_{\mathfrak{g}}| = |\det \mathrm{Ad}(h)|_{\mathfrak{h}}| \qquad (h \in H).$$

Proof: Clearly ϵ is a linear map. Assume that ω is a G-invariant density on G/H. Then for $g \in G$ we have: $T_{eH}(l_g)^*\omega(gH) = l_g^*(\omega)(eH) = \epsilon(\omega)$, hence

$$\omega(gH) = (T_{eH}(l_g)^{-1})^*\epsilon(\omega) = A(g)^{*-1}\epsilon(\omega).$$

This shows that the map ϵ has a trivial kernel, hence is injective, and that its image is contained in $\mathcal{D}(\mathfrak{g}/\mathfrak{h})^H$. To establish its surjectivity, let $\omega_0 \in \mathcal{D}(\mathfrak{g}/\mathfrak{h})^H$. Then for all $h \in H$ we have

$$(T_{eH}(l_h)^{-1})^*\omega_0 = A(h)^{*-1}\omega_0 = \omega_0,$$

hence we may define a density on G/H by

$$\omega(gH) = (T_{eH}(l_g)^{-1})^*\omega_0.$$

Note that the right hand side of this equation stays the same if g is replaced by $gh,\ h \in H$. Hence the definition is unambiguous. One readily verifies that ω thus defined is smooth, G-invariant, and has image ω_0 under ϵ. This proves (a); the statement about positivity is obvious from the above.

From (a) it follows that the dimension of $\Gamma(\mathcal{D}T(G/H))^G$ equals $\dim\mathcal{D}(\mathfrak{g}/\mathfrak{h})$; hence, it is at most one. The final assertion now follows from what was said in the preceding text. $\square$

Corollary 19.19 *Let G be a Lie group, H a compact subgroup. Then G/H has a G-invariant positive density. Two such densities differ by a positive factor.*

Proof: For $h \in H$, we put

$$\Delta(h) = \frac{|\mathrm{detAd}(h)|_{\mathfrak{g}}|}{|\mathrm{detAd}(h)|_{\mathfrak{h}}|}.$$

Clearly Δ is a Lie group homomorphism from H to the group R^+ consisting of the positive real numbers, equipped with multiplication. Thus, $\Delta(H)$ is a compact subgroup of $\mathbb{R}^+$. In view of Lemma 19.16 this implies that $\Delta(H) = \{1\}$. The result follows. $\square$

Example 19.20 As $S^n \simeq \mathrm{SO}(n+1)/\mathrm{SO}(n)$, see Example 15.6, it follows that S^n has a unique $\mathrm{SO}(n+1)$-invariant density of total volume 1. Similarly, $\mathbb{P}^n(\mathbb{R})$ has a unique $\mathrm{SO}(n+1)$-invariant density of total volume 1; see Example 15.7. Real projective space is non-orientable, so it does not have a volume form, i.e., a nowhere vanishing exterior differential form of top degree. This problem of possible non-orientability of homogeneous spaces has been our motivation in using densities rather than forms to define invariant integration.

20 Representations

In this section G will always be a Lie group.

In the following we will give some of the basic definitions of representation theory with V a complete *locally convex space* over $\mathbb{C}$. Every *Banach space* is an example of such a space. Natural spaces of importance for analysis, like $C(M)$, $C_c(M)$, $C^\infty(M)$, $C_c^\infty(M)$, with M a smooth manifold, and also the spaces $\mathcal{D}'(M)$ and $\mathcal{E}'(M)$ of distributions and compactly supported distributions, respectively, are complete locally convex, but in general not Banach. Of course *Hilbert spaces* are Banach spaces; thus, they are covered as well.

Definition 20.1 Let V be a locally convex space. A *continuous representation* $\pi = (\pi, V)$ of G in V is a continuous left action $\pi : G \times V \to V$, such that $\pi(x) : v \mapsto \pi(x)v = \pi(x, v)$ is a linear endomorphism of V, for every $x \in G$. The representation is called *finite dimensional* if $\dim V < \infty$.

Remark 20.2 If G is just a group, and V just a linear space, one defines a representation of G in V similarly, but without the requirement of continuity.

Example 20.3

(a) Let $G \times X \to X, (g, x) \mapsto gx$ be a left action of G on a set X, and let $\mathcal{F}(X)$ denote the space of functions $X \to \mathbb{C}$. Then the action naturally induces the representation L of G on $\mathcal{F}(X)$ given by

$$L_g\varphi(x) = \varphi(g^{-1}x),$$

for $\varphi \in \mathcal{F}(X)$, $g \in G$ and $x \in X$.

(b) Let L be the action of G on $\mathcal{F}(G)$ induced by the left action $G \times G \to G$, $(g, x) \mapsto gx$. This is called the *left regular representation* of G. It is given by the formula $L_g\varphi(x) = \varphi(g^{-1}x)$, for $x, g \in G$.

Similarly, the right multiplication of G on itself induces the *right regular representation* of G on $\mathcal{F}(G)$ given by

$$R_g\varphi(x) = \varphi(xg),$$

for $\varphi \in \mathcal{F}(G)$, $g, x \in G$. These representations leave the subspace $C(G) \subset \mathcal{F}(G)$ invariant. Similarly, if dx is a left or right invariant Haar measure on G, then the associated space $L^2(G)$ of square integrable functions is invariant under both L and R. One can show that the restrictions of L and R to $L^2(G)$ are continuous, see Proposition 20.10.

(c) The natural action of $\mathrm{SU}(2)$ on $\mathbb{C}^2$ induces a representation π of $\mathrm{SU}(2)$ on $\mathcal{F}(\mathbb{C}^2)$ given by

$$\pi(x)\varphi(z) = \varphi(x^{-1}z) = \varphi(\bar{\alpha}z_1 + \bar{\beta}z_2, -\beta z_1 + \alpha z_2),$$

for $\varphi \in \mathcal{F}(\mathbb{C}^2)$, $z \in \mathbb{C}^2$ and

$$x = \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix} \in \mathrm{SU}(2),$$

i.e., $\alpha, \beta \in \mathbb{C}$, and $|\alpha|^2 + |\beta|^2 = 1$.

Lemma 20.4 *Let (π, V) be a finite dimensional representation of G. If π is continuous, then π is smooth.*

Proof: By finite dimensionality of V, the group $\mathrm{GL}(V)$ is a Lie group. The map $\pi : x \mapsto \pi(x)$ is a homomorphism from G to $\mathrm{GL}(V)$. The hypothesis that the representation is continuous means that the map $(x, v) \mapsto \pi(x)v$ is continuous. By finite dimensionality of V this implies that $\pi : G \to \mathrm{GL}(V)$ is continuous. By Corollary 9.3 it follows that $\pi : G \to \mathrm{GL}(V)$ is smooth. This in turn implies that $(g, v) \mapsto \pi(g)v$ is smooth $G \times V \to V$. $\square$

In the setting of the above lemma, the tangent map of $\pi : G \to \mathrm{GL}(V)$ at e is a Lie algebra homomorphism $\pi_* : \mathfrak{g} \to \mathrm{End}(V)$, where the latter space is equipped with the commutator bracket. This motivates the following definition.

Definition 20.5 By a *representation* of $\mathfrak{l}$ in V we mean a Lie algebra homomorphism $\rho : \mathfrak{l} \to \mathrm{End}(V)$, i.e., ρ is a linear map such that for all $X, Y \in \mathfrak{l}$ we have:

$$\rho([X, Y]) = \rho(X)\rho(Y) - \rho(Y)\rho(X).$$

A representation of $\mathfrak{l}$ in V is also called a structure of $\mathfrak{l}$-*module* on V. Accordingly, ρ is often suppressed in the notation, by writing $\rho(X)v = Xv$, for $X \in \mathfrak{l}, v \in V$. With this notation, the above rule becomes

$$[X, Y]v = XYv - YXv \qquad (X, Y \in \mathfrak{l},\ v \in V).$$

Remark 20.6 Similarly, a complete locally convex space V, equipped with a continuous representation of a Lie group G, will sometimes be called a *G-module*

Example 20.7 $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$ is a continuous representation of G in $\mathfrak{g}$. The associated infinitesimal representation of $\mathfrak{g}$ in $\mathrm{End}(\mathfrak{g})$ is given by $(X, Y) \mapsto (\mathrm{ad} X)Y = [X, Y]$.

Proposition 20.8 *Let π be a representation of G in a Banach space V. Then the following conditions are equivalent:*

(a) $\pi : G \times V \to V$ *is continuous.*

(b) *For every $x \in G$ the map $\pi(x)$ is continuous, and for every $v \in V$ the map $G \to V, x \mapsto \pi(x)v$ is continuous at e.*

Proof: That (b) follows from (a) is obvious. We will establish the converse implication by application of the *Banach-Steinhaus* (or uniform boundedness) theorem.

Assume (b). Fix $x_0 \in G$. If $v \in V$ then $\pi(x)v = \pi(x_0)\pi(x_0^{-1}x)v$; using (b) we see that $x \mapsto \pi(x)v$ is continuous at x_0.

Now fix $v_0 \in V$. Select a compact neighborhood N of x_0 in G. Then $\{\pi(x) \mid x \in N\}$ is a collection of continuous linear maps $V \to V$. Moreover, for every $v \in V$, the map $x \mapsto \|\pi(x)v\|$ is continuous, hence bounded on N. By the uniform boundedness theorem it follows that the collection of operator norms $\|\pi(x)\|$, for $x \in N$ is bounded, say by a constant $C > 0$. It follows that for $x \in N, v \in V$ we have

$$\begin{aligned}\|\pi(x)v - \pi(x_0)v_0\| &\leq \|\pi(x)v - \pi(x)v_0\| + \|\pi(x)v_0 - \pi(x_0)v_0\| \\ &\leq C\|v - v_0\| + \|\pi(x)v_0 - \pi(x_0)v_0\|.\end{aligned}$$

The second term on the right-hand side tends to 0 if $x \to x_0$, by (b). Hence $(x, v) \mapsto \pi(x)v$ is continuous in (x_0, v_0). $\square$

Remark 20.9 The above proof is based on the principle of uniform boundedness, and readily generalizes to the category of complete locally convex spaces for which this principle holds, the so called barrelled spaces.

The following result is in particular of interest if $X = G$ and dx a left invariant positive density on X.

Proposition 20.10 *Let X be a manifold equipped with a continuous left G-action. Let dx be a G-invariant positive continuous density on X. Then the natural representation L of G in $L^2(X)$ is continuous.*

Proof: In view of the previous proposition it suffices to show that for every $\varphi \in L^2(X)$ the map $\Phi : x \mapsto L_x\varphi,\ G \to L^2(X)$ is continuous at e. Thus we must estimate the L^2-norm of the function $L_x\varphi - \varphi$ as $x \to e$. Let $\epsilon > 0$. Then there exists a $\psi \in C_c(X)$ such that $\|\varphi - \psi\|_2 < \frac{1}{3}\epsilon$.

Let $g \in C_c(G)$ be a non-negative function such that $g = 1$ on an open neigbourhood of suppψ. Then for x sufficiently close to e we have $g = 1$ on supp$L_x\psi$. Thus for such x we have:

$$\begin{aligned}\|L_x\varphi - \varphi\|_2 &\leq \frac{2}{3}\epsilon + \|L_x\psi - \psi\|_2 \\ &= \frac{2}{3}\epsilon + \|(L_x\psi - \psi)g\|_2 \\ &\leq \frac{2}{3}\epsilon + \|L_x\psi - \psi\|_\infty\|g\|_2.\end{aligned}$$

Fix a compact neighborhood N of suppψ. For x sufficiently close to e one has supp$L_x\psi \subset N$. By uniform continuity of ψ on N, it now follows that $\|L_x\psi - \psi\|_\infty\|g\|_2 < \frac{\epsilon}{3}$ for x sufficiently close to e. □

Definition 20.11 Let π be a representation of G in a (complex) linear space V. By an *invariant subspace* we mean a linear subspace $W \subset V$ such that $\pi(x)W \subset W$ for every $x \in G$.

A continuous representation π of G in a complete locally convex space V is called *irreducible*, if 0 and V are the only closed invariant subspaces of V.

Remark 20.12 Note that for a finite dimensional representation (π, V) an invariant subspace is automatically closed. Thus, such a representation is irreducible if the only invariant subspaces are 0 and V.

Definition 20.13 By a *unitary representation* of G we will always mean a continuous representation π of G in a (complex) Hilbert space $\mathcal{H}$, such that $\pi(x)$ is unitary for every $x \in G$.

Remark 20.14 Let V be a complex linear space. Then by a *sesquilinear form* on V we mean a map $\beta : V \times V \to \mathbb{C}$ which is linear in the first variable, and conjugate linear in the second, i.e., $\beta(v, \lambda w + w') = \bar{\lambda}\beta(v, w) + \beta(v, w')$ for all $v, w, w' \in V$, $\lambda \in \mathbb{C}$.

A *Hermitian inner product* on V is a sesquilinear form $\langle \cdot\,, \cdot \rangle$ that is conjugate symmetric, i.e. $\langle v\,, w\rangle = \overline{\langle w\,, v\rangle}$, and positive definite, i.e., $\langle v\,, v\rangle \geq 0$ and $\langle v\,, v\rangle = 0 \Rightarrow v = 0$ for all $v \in V$.

Finally, we recall that a *complex Hilbert space* is a complex linear space $\mathcal{H}$ equipped with a Hermitian inner product $\langle \cdot\,, \cdot \rangle$, whose associated norm is complete.

Remark 20.15 According to the above definition, a continuous representation of G in $\mathcal{H}$ is unitary if and only if

$$\langle \pi(x)v\,, w\rangle = \langle v\,, \pi(x^{-1})w\rangle \qquad (v, w \in \mathcal{H},\ x \in G).$$

Definition 20.16 A continuous finite dimensional representation (π, V) of G will be called *unitarizable* if there exists a Hermitian inner product on V for which π is unitary.

Proposition 20.17 *Let G be compact, and suppose that (π, V) is a continuous finite dimensional representation of G. Then π is unitarizable.*

Proof: Let dx denote right Haar measure on G, and fix any positive definite Hermitian inner product $\langle \cdot\, ,\, \cdot \rangle_1$ on V. Then we define a new Hermitian pairing on V by

$$\langle v\, ,\, w\rangle = \int_G \langle \pi(x)v\, ,\, \pi(x)w\rangle_1\, dx \qquad (v, w \in V).$$

Notice that the integrand $\iota_{v,w}(x) = \langle \pi(x)v\, ,\, \pi(x)w\rangle_1$ in the above equation is a continuous function of x. We claim that the pairing thus defined is positive definite. Indeed, if $v \in V$ then the function $\iota_{v,v}$ is continuous and positive on G. Hence $\langle v\, ,\, v\rangle = \int_G \iota_{v,v}(x)\, dx \geq 0$ by positivity of the measure. Also, if $\langle v\, ,\, v\rangle = 0$, then $\iota_{v,v} \equiv 0$ by Lemma 19.2, and hence $\langle v\, ,\, v\rangle = \iota_{v,v}(e) = 0$, and positive definiteness follows.

Finally we claim that π is unitary for the inner product thus defined. Indeed this follows from the invariance of the measure. If $y \in G$, and $v, w \in V$, then

$$\langle \pi(y)v\, ,\, \pi(y)w\rangle = \int_G \iota_{v,w}(xy)dx = \int_G \iota_{v,w}(x)dx = \langle v\, ,\, w\rangle.$$

$\square$

Lemma 20.18 *Let $(\pi, \mathcal{H})$ be a unitary representation of G. If $\mathcal{H}_1$ is an invariant subspace for π, then its orthocomplement $\mathcal{H}_2 = \mathcal{H}_1^\perp$ is a closed invariant subspace for π. If $\mathcal{H}_1$ is closed, then we have the direct sum $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ of closed invariant subspaces.*

Proof: Let $v \in \mathcal{H}_2$ and let $x \in G$. We will show that $\pi(x)v \in \mathcal{H}_2$. If $w \in \mathcal{H}_1$, then $\pi(x^{-1})w$ belongs to $\mathcal{H}_1$ as well, so that $\langle \pi(x)v\, ,\, w\rangle = \langle v\, ,\, \pi(x^{-1})w\rangle = 0$. It follows that $\pi(x)v \in \mathcal{H}_1^\perp$.
$\square$

Corollary 20.19 *Let (π, V) be a continuous finite dimensional representation of G. If π is unitarizable, then it decomposes as a finite direct sum of irreducibles; i.e., there exists a direct sum decomposition $V = \oplus_{1\leq j\leq n} V_j$ of V into invariant subspaces such that for every j the representation π_j defined by $\pi_j(x) = \pi(x)|_{V_j}$ is irreducible.*

Proof: Fix an inner product for which π is unitary, and apply the above lemma repeatedly. $\square$

Corollary 20.20 *Let (π, V) be a continuous finite dimensional representation of a compact Lie group. Then every invariant subspace of V has a complementary invariant subspace. Moreover, π admits a decomposition as a finite direct sum of irreducible representations.*

Proof: By Proposition 20.17 π is unitarizable. Now apply Lemma 20.18 and Corollary 20.19.

Definition 20.21 Let (π, V) be a finite dimensional continuous representation of G. Then by a *matrix coefficient* of π we mean any function $m : G \to \mathbb{C}$ of the form

$$m(x) = m_{v,\eta}(x) := \langle \pi(x)v\, ,\, \eta\rangle$$

with $v \in V$ and $\eta \in V^*$.

Remark 20.22 Note that the map $x \mapsto \pi(x)$ is smooth, so that every matrix coefficient belongs to $C^\infty(G)$.

If $\langle \cdot\,,\, \cdot \rangle$ is a Hermitian inner product on V, then a matrix coefficient of π may also be characterized as a function of the form

$$m = m_{v,w} : x \mapsto \langle \pi(x)v\,,\, w \rangle,$$

with $v, w \in V$, since $w \mapsto \langle \cdot\,,\, w \rangle$ is a conjugate linear bijection from V onto V^*.

Let now (π, V) be a finite dimensional unitary representation of G, and fix an orthonormal basis $u_1, \ldots, u_n$ of V. Then for every $x \in G$ we define the matrix $M(x) = M_{\underline{u}}(x)$ by

$$M(x)_{ij} = m_{u_i, u_j}(x).$$

This is just the matrix of $\pi(x)$ with respect to the basis $\underline{u}$. Note that it is unitary. Note also that $M(xy) = M(x)M(y)$. Thus M is a continuous group homomorphism from G to the group $U(n)$ of unitary $n \times n$ matrices.

Definition 20.23 If (π_j, V_j), for $j = 1, 2$, are continuous representations of G in complete locally convex spaces, then a continuous linear map $T : V_1 \to V_2$ is said to be *equivariant*, or *intertwining* if the following diagram commutes for every $x \in G$:

$$\begin{array}{ccccc} & V_1 & \xrightarrow{T} & V_2 & \\ \pi_1(x) & \uparrow & & \uparrow & \pi_2(x) \\ & V_1 & \xrightarrow{T} & V_2 & \end{array}$$

The representations π_1 and π_2 are said to be *equivalent* if there exists a topological linear isomorphism T from V_1 onto V_2 which is equivariant.

If the above representations are finite dimensional, then one does not need to require T to be continuous, since every linear map $V_1 \to V_2$ has this property. In the case of finite dimensional representations we shall write $\mathrm{Hom}_G(V_1, V_2)$ for the linear space of interwining linear maps $V_1 \to V_2$ and $\mathrm{End}_G(V_1)$ for the space of intertwining linear endomorphisms of V_1.

If V is a complex linear space, we write $\mathrm{End}(V)$ for the space of linear maps from V to itself, and $\mathrm{GL}(V)$ for the group of invertible elements in $\mathrm{End}(V)$. If π is a representation of G in V, then we may define a representation $\tilde{\pi}$ of G in $\mathrm{End}(V)$ by

$$\tilde{\pi}(g)A = \pi(g)A\pi(g)^{-1}.$$

Note that if π is finite dimensional and continuous, then so is $\tilde{\pi}$. Note also that the space

$$\mathrm{End}(V)^G = \{A \in \mathrm{End}(V) \mid \tilde{\pi}(g)A = A\}$$

of G-invariants in V is just the space $\mathrm{End}_G(V)$ of G-equivariant linear maps $V \to V$.

Exercise 20.24 Let (π_j, V_j), for $j = 1, 2$, be two finite dimensional representations of G. Show that π_1 and π_2 are equivalent if and only if there exist choices of bases for V_1 and V_2, such that for the associated matrices one has:

$$\mathrm{mat}\pi_1(x) = \mathrm{mat}\pi_2(x).$$

Example 20.25 We recall that SU(2) is the group of matrices of the form

$$g = \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix}$$

with $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$. The group SU(2) acts on $\mathbb{C}^2$ in a natural way, and we have the associated representation π on the space $P(\mathbb{C})$ of polynomial functions $p : \mathbb{C}^2 \to \mathbb{C}$. It is given by the formula

$$\pi(g)p(z) = p(g^{-1}z) = p(\bar{\alpha}z_1 + \bar{\beta}z_2, -\beta z_1 + \alpha z_2)$$

The subspace $P_n = P_n(\mathbb{C}^2)$ of homogeneous polynomials of degree n is an invariant subspace for π. We write π_n for the restriction of π to P_n.

We will now discuss a result that will allow us to show that the representations π_n of the above example are irreducible. We first need the following lemma from linear algebra.

Lemma 20.26 *Let V be a finite dimensional complex linear space, and let $A, B \in \mathrm{End}(V)$ be such that $AB = BA$. Then A leaves $\ker B, \mathrm{im} B$ and all the eigenspaces of B invariant.*

Proof: Elementary, and left to the reader. □

From now on all representations of G are assumed to be continuous.

Lemma 20.27 (Schur's lemma) *Let (π, V) be a finite dimensional representation of G. Then the following holds.*

(a) *If π is irreducible then $\mathrm{End}_G(V) = \mathbb{C}\,\mathrm{I}_V$.*

(b) *Conversely, if π is unitarizable and $\mathrm{End}_G(V) = \mathbb{C}\,\mathrm{I}_V$, then π is irreducible.*

Proof: '(a)' Suppose that π is irreducible, and let $A \in \mathrm{End}(V)^G$. Let $\lambda \in \mathbb{C}$ be an eigenvalue of A, and let $E_\lambda = \ker(A - \lambda\mathrm{I})$ be the associated eigenspace. Note that for non-triviality of this eigenspace we need V to be complex. For every $x \in G$ we have that $\pi(x)$ commutes with A, hence leaves E_λ invariant. In view of the irreducibility of π it now follows that $E_\lambda = V$, hence $A = \lambda\mathrm{I}$.

'(b)' By unitarizability of π, there exists a positive definite inner product $\langle\cdot\,,\,\cdot\rangle$ for which π is unitary.

Let $0 \neq W \subset V$ be a G-invariant subspace. For the proof that π is irreducible it suffices to show that we must have $W = V$. Let P be the orthogonal projection $V \to W$. Since W and $W^\perp$ are both G-invariant, we have, for $g \in G$, that $\pi(g)P = \pi(g) = P\pi(g)$ on W, and $\pi(g)P = 0 = P\pi(g)$ on $W^\perp$. Hence $P \in \mathrm{End}_G(V)$, and it follows that $P = \lambda\mathrm{I}$ for some $\lambda \in \mathbb{C}$. Now $P \neq 0$, hence $\lambda \neq 0$. Also, $P^2 = P$, hence $\lambda^2 = \lambda$, and we see that $\lambda = 1$. Therefore $P = I$, and $W = V$. □

We will now apply the above lemma to prove the following.

Proposition 20.28 *The representations* $(\pi_n, P_n(\mathbb{C}^2))$ *of* $\mathrm{SU}(2)$, *for* $n \in \mathbb{N}$, *are irreducible.*

For the proof we will need compactness of SU(2). In fact we have the following more general result.

Exercise 20.29 For $n \geq 1$, let $\mathrm{M}(n, \mathbb{R})$ and $\mathrm{M}(n, \mathbb{C})$ denote the linear spaces of $n \times n$ matrices with entries in $\mathbb{R}$ and $\mathbb{C}$ respectively. Show that $\mathrm{SU}(n)$ is a closed and bounded subset of $\mathrm{M}(n, \mathbb{C})$. Show that $\mathrm{SO}(n) = \mathrm{SU}(n) \cap \mathrm{M}(n, \mathbb{R})$. Finally show that the Lie groups $\mathrm{SO}(n)$ and $\mathrm{SU}(n)$ are compact.

Proof of Proposition 20.28: Let $n \geq 0$ be fixed, and put $\pi = \pi_n$ and $V = P_n(\mathbb{C}^2)$. Then π_n is unitarizable, since SU(2) is compact. Suppose that $A \in \mathrm{End}(V)$ is equivariant. Then in view of Lemma 20.27 (b) it suffices to show that A is a scalar.

For $0 \leq k \leq n$ we define the polynomial $p_k \in V$ by

$$p_k(z) = z_1^{n-k} z_2^k.$$

Then $\{p_k \mid 0 \leq k \leq n\}$ is a basis for V. For $\varphi \in \mathbb{R}$ we put

$$t_\varphi = \begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix}, \qquad r_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

Then

$$T = \{t_\varphi \mid \varphi \in \mathbb{R}\} \quad \text{and} \quad R = \{r_\varphi \mid \varphi \in \mathbb{R}\}$$

are (closed) subgroups of SU(2). One readily verifies that for $0 \leq k \leq n$ and $\varphi \in \mathbb{R}$ we have:

$$\pi(t_\varphi) p_k = e^{i(2k-n)\varphi} p_k.$$

Thus every p_k is a joint eigenvector for T. Fix a φ such that the numbers $e^{i(2k-n)\varphi}$ are mutually different. Then for every $0 \leq k \leq n$ the space $\mathbb{C}p_k$ is eigenspace for $\pi(t_\varphi)$ with eigenvalue $e^{i(2k-n)\varphi}$. Since A and $\pi(t_\varphi)$ commute it follows that A leaves all the spaces $\mathbb{C}p_k$ invariant. Hence there exist $\lambda_k \in \mathbb{C}$ such that

$$Ap_k = \lambda_k p_k, \quad 0 \leq k \leq n.$$

Let E_0 be the eigenspace of A with eigenvalue λ_0. We will show that $E_0 = V$, thereby completing the proof. The space E_0 is SU(2)-invariant, and contains p_0. Hence it contains $\pi(r_\varphi)p_0$ for every $\varphi \in \mathbb{R}$. By a straightforward computation one sees that

$$\pi(r_\varphi)p_0(z_1, z_2) = (\cos\varphi\, z_1 + \sin\varphi\, z_2)^n = \sum_{k=0}^{n} \binom{n}{k} \cos^{n-k}\varphi\, \sin^k\varphi\ p_k.$$

From this one sees by application of A and using the intertwining property, that

$$\sum_{k=0}^{n} \binom{n}{k} \cos^{n-k}\varphi\, \sin^k\varphi\ (\lambda_0 - \lambda_k)\, p_k = 0,$$

for all $\varphi \in \mathbb{R}$. By linear independence of the p_k, it follows that $\lambda_k = \lambda_0$, for every $0 \leq k \leq n$. Hence $E_0 = V$. $\square$

We end this section with two useful consequences of Schur's lemma.

Lemma 20.30 *Let (π, V), (π', V') be two irreducible finite dimensional representations of G. If π and π' are not equivalent, then every intertwining linear map $T : V \to V'$ is trivial.*

Proof: Let T be intertwining, and non-trivial. Then $\ker T \subset V$ is a proper G-invariant subspace. Hence $\ker T = 0$, and it follows that T is injective. Therefore its image $\mathrm{im} T$ is a non-trivial G-invariant subspace of V'. It follows that $\mathrm{im} T = V'$, hence T is a bijection, contradicting the inequivalence. $\square$

If (π, V) is a representation for a group G, then a sesquilinear form β on V is called equivariant if $\beta(\pi(g)v, \pi(g)w) = \beta(v, w)$ for all $v, w \in V$, $g \in G$.

Lemma 20.31 *Let (π, V) be an irreducible finite dimensional unitary representation of a locally compact group G. Then the equivariant sesquilinear forms on V are precisely the maps $\beta : V \times V \to \mathbb{C}$ of the form $\beta = \lambda\langle \cdot\, , \cdot \rangle$, $\lambda \in \mathbb{C}$. Here $\langle \cdot\, , \cdot \rangle$ denotes the (equivariant) inner product of the Hilbert space V.*

Proof: Let $\beta : V \times V \to \mathbb{C}$ be sesquilinear. Then for every $w \in V$ the map $v \mapsto \beta(v, w)$ is a linear functional on V. Hence there exists a unique $A(w) \in V$ such that $\beta(v, w) = \langle v, A(w)\, , .\rangle$ One readily verifies that $A : V \to V$ is a linear map. Moreover, the equivariance of β and $\langle \cdot\, , \cdot \rangle$ imply that A is equivariant. Since π is irreducible it follows by Schur's lemma that $A = \lambda \mathrm{I}$ for some $\lambda \in \mathbb{C}$, whence the result. $\square$

21 Schur orthogonality

Assumption *In the rest of these notes every finite dimensional representation of a Lie group will be assumed to be continuous, unless specified otherwise.*

In this section G will be a compact Lie group, unless stated otherwise. Let dx be the unique left invariant density on G with $\int_G dx = 1$; for its existence, see Lemma 19.10. Then dx is positive. By Remark 19.15, the density dx is right invariant as well.

If π is a finite dimensional irreducible unitary representation of G we write

$$C(G)_\pi \tag{22}$$

for the linear span of the space of matrix coefficients of π. Notice that the space $C(G)_\pi$ does not depend on the chosen (unitary) inner product on V. Thus, by Proposition 20.17 we may define $C(G)_\pi$ for any irreducible finite dimensional (continous) representation π of G.

There is a nice way to express sums of matrix coefficients of a finite dimensional unitary representation (π, V) of G by means of the trace of a linear map. Let $v, w \in V$. Then we shall write $L_{v,w}$ for the linear map $V \to V$ given by

$$L_{v,w}(u) = \langle u\, ,\ w \rangle v.$$

One readily sees that

$$\operatorname{tr}(L_{v,w}) = \langle v\,,\,w\rangle, \qquad v, w \in V. \tag{23}$$

Indeed both sides of the above equation are sesquilinear forms in (v, w), so it suffices to check the equation for v, w members of an orthonormal basis, which is easily done.

It follows from the above equation that

$$m_{v,w}(x) = \operatorname{tr}(\pi(x)L_{v,w}).$$

Hence every sum m of matrix coefficients is of the form $m(x) = \operatorname{tr}(\pi(x)A)$, with $A \in \operatorname{End}(V)$. Conversely if $\{e_k \mid 1 \leq k \leq n\}$ is an orthonormal basis for V, then one readily sees that any endomorphism $A \in \operatorname{End}(V)$ may be expressed as

$$A = \sum_{1\leq i,j\leq n} \langle Ae_j\,,\,e_i\rangle L_{e_i,e_j}.$$

Using this one may express every function of the form $x \mapsto \operatorname{tr}(\pi(x)A)$ as a sum of matrix coefficients.

We now define the linear map $T_\pi : \operatorname{End}(V) \to C(G)$ by

$$T_\pi(A)(x) = \operatorname{tr}(\pi(x)A), \qquad x \in G,$$

for every $A \in \operatorname{End}(V)$. Let π be irreducible, then it follows from the above discussion that T_π maps V onto $C(G)_\pi$. Define the representation $\pi \otimes \pi^*$ of $G \times G$ on $\operatorname{End}(V)$ by

$$[\pi \otimes \pi^*](x, y)A = \pi(x)A\pi(y)^{-1},$$

for $A \in \operatorname{End}(V)$ and $x, y \in G$.

We define the representation $R \times L$ of $G \times G$ on $C(G)$ by

$$(R \times L)(x, y) := R_x \circ L_y = L_y \circ R_x.$$

Lemma 21.1 *Let (π, V) be a finite dimensional irreducible representation of G. Then $C(G)_\pi$ is invariant under $R \times L$. The map $T_\pi : V \to C(G)_\pi$ is surjective, and intertwines the representations $\pi \otimes \pi^*$ and $R \times L$ of $G \times G$.*

Proof: We first prove the equivariance of $T_\pi : \operatorname{End}(V) \to C(G)$. Let $A \in \operatorname{End}(V)$ and $x, y \in G$, then for all $g \in G$,

$$T_\pi([\pi \otimes \pi^*](x, y)A)(g) = \operatorname{tr}(\pi(g)\pi(x)A\pi(y^{-1})) = \operatorname{tr}(\pi(y^{-1}gx)A) = R_xL_y(T_\pi(A))(g).$$

Note that it follows from this equivariance that the image of T_π is $R \times L$-invariant. In an earlier discussion we showed already that $\operatorname{im}(T_\pi) = C(G)_\pi$. $\square$

Corollary 21.2 *If π and π' are equivalent finite dimensional irreducible representations of G, then $C(G)_\pi = C(G)_{\pi'}$.*

Proof: Let V, V' be the associated representation spaces. Then by equivalence there exists a linear isomorphism $T : V \to V'$ such that $T \circ \pi(x) = \pi'(x) \circ T$ for all $x \in G$. Hence for $A \in \mathrm{End}(V)$ and $x \in G$,

$$T_{\pi'}(TAT^{-1})(x) = \mathrm{tr}(\pi'(x)TAT^{-1}) = \mathrm{tr}(T^{-1}\pi'(x)TA) = \mathrm{tr}(\pi(x)A) = T_\pi(A)(x).$$

Now use that T_π and $T_{\pi'}$ have images $C(G)_\pi$ and $C(G)_{\pi'}$, respectively, by Lemma 21.1. $\square$

We now have the following.

Theorem 21.3 (Schur orthogonality). *Let (π, V) and (π', V') be two irreducible finite dimensional representations of G. Then the following holds.*

(a) *If π and π' are not equivalent, then $C(G)_\pi \perp C(G)_{\pi'}$ (with respect to the Hilbert structure of $L^2(G)$).*

(b) *Let V be equipped with an inner product for which π is unitary. If $v, w, v', w' \in V$, then the L^2-inner product of the matrix coefficients $m_{v,w}$ and $m_{v',w'}$ is given by:*

$$\int_G m_{v,w}(x)\,\overline{m_{v',w'}(x)}\;dx = \dim(\pi)^{-1}\;\langle v\,,\,v'\rangle\overline{\langle w\,,\,w'\rangle} \tag{24}$$

Remark 21.4 The relations (24) are known as the *Schur orthogonality relations*. Of course the assumption that dx is normalized is a necessary assumption for (24) to hold.

Proof: For $w \in V$ and $w' \in V'$ we define the linear map $L_{w',w} : V \to V'$ by $L_{w',w}u = \langle u\,,\,w\rangle w'$. Consider the following linear map $V \to V'$, defined by averaging,

$$I_{w',w} = \int_G \pi'(x)^{-1} \circ L_{w',w} \circ \pi(x)\;dx.$$

One readily verifies that

$$\langle I_{w',w}v\,,\,v'\rangle = \langle m_{v,w}\,,\,m_{v',w'}\rangle_{L^2}. \tag{25}$$

Moreover, by right invariance of the measure dx it readily follows that $I_{w',w}$ is an intertwining map from (V,π) to (V',π').

(a): If π and π' are inequivalent then the intertwining map $I_{w',w}$ is trivial by Lemma 20.30. Now apply (25) to prove (a).

(b): Now assume $V = V'$. Then for all $w, w' \in V$ we have $I_{w',w} \in \mathrm{End}_G(V)$, hence $I_{w',w}$ is a scalar. It follows that there exists a sesquilinear form β on V such that

$$I_{w',w} = \beta(w', w)\,\mathrm{I}_V.$$

Applying the trace to both sides of the above equation we find that $\mathrm{tr}(I_{w',w}) = d_\pi\beta(w', w)$. Here we have abbreviated $d_\pi = \dim(\pi)$. On the other hand, since tr is linear,

$$\mathrm{tr}(I_{w',w}) = \int_G \mathrm{tr}(\pi(x)^{-1}L_{w',w}\pi(x))\;dx = \int_G \mathrm{tr}(L_{w',w})\;dx = \mathrm{tr}(L_{w',w}) = \langle w'\,,\,w\rangle.$$

Hence

$$I_{w',w} = \beta(w', w)\,\mathrm{I}_V = d_\pi^{-1}\langle w'\,,\,w\rangle\,\mathrm{I}_V.$$

Now apply (25) to prove (b). $\square$

Another way to formulate the orthogonality relations is the following (V is assumed to be equipped with an inner product for which π is unitary). If $A \in \mathrm{End}(V)$, let A^* denote the Hermitian adjoint of A. Then one readily verifies that

$$(A, B) \mapsto \langle A\,,\ B\rangle := \mathrm{tr} B^* A$$

defines a Hermitian inner product on $\mathrm{End}(V)$. Moreover, the representation $\pi \otimes \pi^*$ is readily seen to be unitary for this inner product.

Corollary 21.5 *The map* $^\circ T_\pi := \sqrt{d_\pi}\, T_\pi$ *is a unitary* G*-equivariant isomorphism from* $\mathrm{End}(V)$ *onto* $C(G)_\pi$.

Proof: We begin by establishing a few properties of the endomorphisms $L_{v,w}$, for $v, w \in V$. From the definition one readily sees that, for $v', w' \in V$ the adjoint of $L_{v',w'}$ is given by

$$L^*_{v',w'} = L_{w',v'}.$$

Moreover, one also readily checks that

$$L_{w',v'} \circ L_{v,w} = \langle v\,,\ v'\rangle L_{w',w}.$$

From these two properties it follows in turn that

$$\langle L_{v,w}\,,\ L_{v',w'}\rangle = \mathrm{tr}(L_{w',v'} \circ L_{v,w}) = \langle v\,,\ v'\rangle\overline{\langle w\,,\ w'\rangle}. \tag{26}$$

Finally, we recall that

$$T_\pi(L_{v,w})(x) = \mathrm{tr}(\pi(x)L_{v,w}) = m_{v,w}(x) \qquad (x \in G),$$

hence

$$\langle {}^\circ T_\pi(L_{v,w})\,,\ {}^\circ T_\pi(L_{v',w'})\rangle_{L^2} = d_\pi \langle m_{v,w}\,,\ m_{v',w'}\rangle_{L^2}. \tag{27}$$

From (26) and (27) we see that the Schur orthogonality relations may be reformulated as

$$\langle {}^\circ T_\pi(L_{v,w})\,,\ {}^\circ T_\pi(L_{v',w'})\rangle_{L^2} = \langle L_{v,w}\,,\ L_{v',w'}\rangle, \tag{28}$$

for all $v, v', w, w' \in V$. The maps $L_{v,w}$, for $v, w \in V$, span the space $\mathrm{End}(V_\pi)$. Hence the Schur orthogonality relations are equivalent to the assertion that $^\circ T_\pi$ is an isometry from $\mathrm{End}(V)$ into $C(G)_\pi$. We proved already that $^\circ T_\pi$ is surjective onto $C(G)_\pi$; hence $^\circ T_\pi$ is a unitary isomorphism. The equivariance of $^\circ T_\pi$ has been established before. $\square$

Definition 21.6 Let (V, π) be a finite dimensional representation of G. The function $\chi_\pi : G \to \mathbb{C}$ defined by

$$\chi_\pi(x) = \mathrm{tr}\pi(x), \qquad (x \in G),$$

is called the *character* of π.

Remark 21.7 Since the representation π is continuous, it is also smooth, hence $\chi_\pi \in C^\infty(G)$. Note that χ_π is a sum of matrix coefficients of π. Thus, if G is compact and π irreducible, then $\chi_\pi \in C(G)_\pi$.

Lemma 21.8 *Let (π, V) be an irreducible finite dimensional representation of G. Then χ_π is the unique conjugation invariant function in $C(G)_\pi$ with $\chi_\pi(e) = d_\pi$. Its L^2-norm relative to the normalized Haar measure is $\|\chi_\pi\|_2 = 1$.*

Proof: We equip V with an inner product for which π is unitary, and define the associated inner product on $\mathrm{End}(V)$ as above. Let $\varphi \in C(G)_\pi$. Then $\varphi = {}^\circ T_\pi(A)$ for a unique $A \in \mathrm{End}(V)$. By equivariance of T_π, the function φ is conjugation invariant if and only if A is G-intertwining, which in turn is equivalent to $A = c\mathrm{I}_V$ for a constant $c \in \mathbb{C}$. We observe that $c = \varphi(e)/d_\pi^{3/2}$. This implies that there exists a unique conjugation invariant function φ with $\varphi(e) = d_\pi$. For this function we have $c = 1/\sqrt{d_\pi}$ and

$$\varphi(x) = {}^\circ T_\pi(c\mathrm{I}_V)(x) = \sqrt{d_\pi}\mathrm{tr}(\pi(x)c\mathrm{I}_V) = \mathrm{tr}\pi(x) = \chi_\pi(x).$$

The assertion about the L^2-norm follows from

$$\|\chi_\pi\|_2 = \mathrm{tr}[(cI)^* cI] = c^2\mathrm{tr}(I) = c^2 d_\pi = 1.$$

□

22 Characters

In this section we assume that G is a Lie group. We shall discuss properties of characters of finite dimensional representations of G.

If V is a finite dimensional complex linear space, we write $\mathrm{End}(V)$ for the space of complex linear maps from V to itself, and $\det = \det_V$ and $\mathrm{tr} = \mathrm{tr}_V$ for the complex determinant and trace functions $\mathrm{End}(V) \to \mathbb{C}$.

Lemma 22.1 *Let $T : V \to W$ be a linear isomorphism of finite dimensional linear spaces. Then for every linear map $A : V \to V$,*

$$\det_W(T \circ A \circ T^{-1}) = \det_V A \quad \text{and} \quad \mathrm{tr}_W(T \circ A \circ T^{-1}) = \mathrm{tr}_V A.$$

Let V be a finite dimensional linear space. Then for all $A, B \in \mathrm{End}(V)$,

$$\mathrm{tr}(A \circ B) = \mathrm{tr}(B \circ A).$$

Proof: Exercise for the reader. □

The character χ_π of a finite dimensional representation (π, V_π) of G is defined as in Definition 21.6.

Lemma 22.2 *Let π, ρ be finite dimensional representations of G. If π and ρ are equivalent, their characters are equal: $\chi_\pi = \chi_\rho$.*

Proof: Let $T : V_\pi \to V_\rho$ be an equivariant linear isomorphism. Then $\rho(x) = T \circ \pi(x) \circ T^{-1}$ for every $x \in G$. The result now follows by application of Lemma 22.1. $\square$

Lemma 22.3 *Let (π, V) be a finite dimensional representation of G. Then, for all $x, y \in G$,*

$$\chi_\pi(xyx^{-1}) = \chi_\pi(y).$$

Proof: Exercise for the reader. $\square$

Definition 22.4 Let π be a representation of G in a finite dimensional complex linear space V. We define the *contragredient* or *dual* of π to be the representation $\pi^\vee$ of G in the dual linear space V^* given by

$$\pi^\vee(x) = \pi(x^{-1})^* : v^* \mapsto v^* \circ \pi(x^{-1}), \qquad (x \in G).$$

Lemma 22.5 *Let (π, V) be a finite dimensional representation of G.*

(a) *If π is continuous, then $\pi^\vee$ is continuous as well.*

(b) *The character of $\pi^\vee$ is given by*

$$\chi_{\pi^\vee}(x) = \chi_\pi(x^{-1}) \qquad (x \in G).$$

Proof: Let $v_1, \ldots, v_n$ be a basis for V and let $v^1, \ldots, v^n$ be the dual basis for V^*, i.e., $v^i(v_j) = \delta_{ij}$. Then, for $x \in G$, the matrix of $\pi^\vee(x)$ with respect to the basis $v_1, \ldots, v_n$ is given by

$$\pi^\vee(x)_{ij} = \langle \pi(x^{-1})^* v^j \,,\, v_i \rangle = \langle v^j \,,\, \pi(x^{-1}) v_i \rangle = \pi(x^{-1})_{ji}.$$

If π is continuous, then its matrix coefficients are continuous functions. Therefore, so are the matrix coefficients of $\pi^\vee$, and (a) follows. Assertion (b) follows from the above identity as well. $\square$

Characters of unitarizable representations have the following special property.

Lemma 22.6 *Let π be a finite dimensional representation of G. If π is unitarizable, then*

$$\chi_\pi(x^{-1}) = \overline{\chi_\pi(x)}, \qquad (x \in G).$$

Proof: Exercise for the reader. $\square$

If (π_1, V_1) and (π_2, V_2) are two continuous representations of G, then we define the *direct sum* representation $\pi = \pi_1 \oplus \pi_2$ in the direct sum $V = V_1 \oplus V_2$ by

$$\pi(x)(v_1, v_2) = (\pi_1(x)v_1, \pi_2(x)v_2) \qquad (v_1 \in V_1,\ v_2 \in V_2,\ x \in G).$$

Lemma 22.7 *Let π_1, π_2 be finite dimensional representations of G. Then*

$$\chi_{\pi_1 \oplus \pi_2} = \chi_{\pi_1} + \chi_{\pi_2}.$$

Proof: Exercise for the reader. □

If (π_1, V_1) and (π_2, V_2) are two finite dimensional representations of G, we define their *tensor product* $\pi_1 \otimes \pi_2$ to be the representation of G in the tensor product space $V_1 \otimes V_2$ given by $(\pi_1 \otimes \pi_2)(x) = \pi_1(x) \otimes \pi_2(x)$. Thus, for $x \in G$, the linear endomorphism $(\pi_1 \otimes \pi_2)(x)$ of $V_1 \otimes V_2$ is determined by

$$(\pi_1 \otimes \pi_2)(x)(v_1 \otimes v_2) = \pi_1(x)v_1 \ \otimes\ \pi_2(x)v_2,$$

for all $v_1 \in V_1,\ v_2 \in V_2$.

Lemma 22.8 *Let (π_1, V_1) and (π_2, V_2) be finite dimensional representations of G. Then the character of their tensor product $\pi_1 \otimes \pi_2$ is given by*

$$\chi_{\pi_1 \otimes \pi_2} = \chi_{\pi_1}\, \chi_{\pi_2}.$$

Proof: Exercise for the reader. Establish, more generally, an identity of the form $\mathrm{tr}(A \otimes B) = \mathrm{tr}(A)\mathrm{tr}(B)$, by choosing suitable bases. □

Exercise 22.9 Recall the definition, for $n \in \mathbb{N}$, of the representation π_n of SU(2) in the finite dimensional space $P_n(\mathbb{C}^2)$ of homogeneous polynomial functions $\mathbb{C}^2 \to \mathbb{C}$ of degree n. Show that the character χ_n of π_n is completely determined by its restriction to $T = \{t_\varphi \mid \varphi \in \mathbb{R}\}$. Hint: use that every matrix in SU(2) is conjugate to a matrix of T.

Show that:

$$\chi_n(t_\varphi) = \frac{\sin(n+1)\varphi}{\sin\varphi},$$

for $\varphi \in \mathbb{R}$. Here t_φ denotes the diagonal matrix with entries $e^{i\varphi}$ and $e^{-i\varphi}$.

Assumption: *In the rest of this section we assume that the Lie group G is compact. We denote by $\langle \cdot\ ,\ \cdot \rangle$ the L^2-inner product with respect to the normalized Haar measure dx on G.*

Lemma 22.10 *Let π, ρ be finite dimensional irreducible representations of G.*

(a) *If $\pi \sim \rho$ then $\langle \chi_\pi\ ,\ \chi_\rho \rangle = 1$.*

(b) *If $\pi \not\sim \rho$ then $\langle \chi_\pi\ ,\ \chi_\rho \rangle = 0$.*

Proof: This follows easily from Theorem 21.3. □

Let π be a finite dimensional representation of the compact group G. Then π is unitarizable, and therefore equivalent to a direct sum $\oplus_{i=1}^n \pi_i$ of irreducible representations. It follows that $\chi_\pi = \sum_{i=1}^n \chi_{\pi_i}$. Using the lemma above we see that for every irreducible representation δ of G we have

$$\#\{i \mid \pi_i \sim \delta\} = \langle \chi_\pi\,,\, \chi_\delta \rangle. \tag{29}$$

In particular this number is independent of the particular decomposition of π into irreducibles. For obvious reasons the number (29) is called the *multiplicity* of δ in π. We shall also denote it by $m(\delta, \pi)$.

Let $\widehat{G}$ denote the set of equivalence classes of finite dimensional irreducible representations of G. Then by abuse of language we shall write $\delta \in \widehat{G}$ to indicate that δ is a representative for an element of $\widehat{G}$. (A better notation would perhaps be $[\delta] \in \widehat{G}$.) If $\delta \in \widehat{G}$ and $m \in \mathbb{N}$, then we write $m\delta$ for (the equivalence class of) the direct sum of m copies of δ.

We have proved the folllowing lemma.

Lemma 22.11 *Let π be a finite dimensional representation of the compact group G. Then*

$$\pi \sim \bigoplus_{\delta \in \widehat{G}} m(\delta, \pi)\delta,$$

where $m(\delta, \pi) = \langle \chi_\pi\,,\, \chi_\delta \rangle \in \mathbb{N}$, for every $\delta \in \widehat{G}$. Any decomposition of π into irreducibles is equivalent to the above one.

Exercise 22.12 This exercise is meant to illustrate that a decomposition of a representation into irreducibles is not unique. Let π_1, π_2 be irreducible representations in V_1, V_2 respectively. Assume that π_1, π_2 are equivalent, and let $T : V_1 \to V_2$ be an intertwining isomorphism.

Equip $V = V_1 \oplus V_2$ with the direct sum representation π, and show that $W_1 = \{(v, Tv) \mid v \in V_1\}$ is an invariant subspace of V. Show that the restriction of π to W_1 is irreducible, and equivalent to π_1. Find a complementary invariant subspace W_2 and show that the restriction of π to this space is also equivalent to π_1.

The following result expresses that the character is a powerful invariant.

Corollary 22.13 *Let π, ρ be two finite dimensional continuous representations of G. Then*

$$\pi \sim \rho \iff \chi_\pi = \chi_\rho.$$

Proof: The implication '$\Rightarrow$' follows from Lemma 22.2. For the converse implication, assume that $\chi_\pi = \chi_\rho$. Then for every $\delta \in \widehat{G}$ we have $m(\delta, \pi) = \langle \chi_\pi\,,\, \chi_\delta \rangle = \langle \chi_\rho\,,\, \chi_\delta \rangle = m(\delta, \rho)$. Now use the previous lemma. □

Corollary 22.14 *Let π be a finite dimensional representation of G. Then π is irreducible if and only if its character χ_π has L^2-norm one.*

Proof: By Schur orthogonality, the characters χ_δ, for $\delta \in \widehat{G}$ form an orthonormal set in $L^2(G)$. It follows that $\|\chi_\pi\|^2 = \sum_\delta m(\delta, \pi)^2$. The result now easily follows. □

23 The Peter-Weyl theorem

In this section we assume that G is a compact Lie group. We denote by $\widehat{G}$ the set of (equivalence classes of) irreducible continuous finite dimensional representations of G.

Definition 23.1 We define the space $\mathcal{R}(G)$ of *representative functions* to be the space of functions $f : G \to \mathbb{C}$ that may be written as a finite sum of functions $f_\delta \in C(G)_\delta$, for $\delta \in \widehat{G}$.

Note that the space $\mathcal{R}(G)$ is contained in $C^\infty(G)$. Moreover, it is invariant under both the left- and the right regular representations of G.

Exercise 23.2 Show that $\mathcal{R}(G)$ is the linear span of the set of all matrix coefficients of finite dimensional continuous representations of G. Hint: consider the decomposition of finite dimensional representations into irreducibles.

Proposition 23.3 *The space of representative functions decomposes according to the algebraic direct sum*

$$\mathcal{R}(G) = \bigoplus_{\delta \in \widehat{G}} C(G)_\delta.$$

The summands are mutually orthogonal with respect to the L^2-inner product. Every summand $C(G)_\delta$ is invariant under the representation $R \times L$ of $G \times G$. Moreover, the restriction of $R \times L$ to that summand is an irreducible representation of $G \times G$.

Proof: The orthogonality of the summands follows from Schur orthogonality. It follows that the above sum is direct.

The map $T_\delta : \mathrm{End}(V_\delta) \to C(G)_\delta$ is bijective and intertwines $\delta \otimes \delta^*$ with $R \times L$. Hence it suffices to show that $\delta \otimes \delta^*$ is an irreducible representation of $G \times G$.

By a straightforward computation one checks that

$$\chi_{\delta\otimes\delta^*}(x, y) = \chi_\delta(x)\overline{\chi_\delta(y)},$$

for $(x, y) \in G \times G$. If dx and dy are normalized right Haar measure on G, then the product measure $dx\, dy$ is the normalized right Haar measure on $G \times G$. Moreover, by Fubini's theorem,

$$\begin{aligned} \|\chi_{\delta\otimes\delta^*}\|^2_{L^2(G\times G)} &= \int_{G\times G} |\chi_\delta(x)|^2 |\chi_\delta(y)|^2 \, dx\, dy \\ &= \int_G \left(\int_G |\chi_\delta(x)|^2 |\chi_\delta(y)|^2 \, dx \right) dy \\ &= \|\chi_\delta\|^2_{L^2(G)} \|\chi_\delta\|^2_{L^2(G)} = 1, \end{aligned}$$

since δ is an irreducible representation of G. It follows from Corollary 22.14 that $\delta \otimes \delta^*$ is an irreducible representation of $G \times G$. $\square$

The proof of the following result is based on the spectral theorem for compact self-adjoint operators in a Hilbert space. It will be given in the next section.

Proposition 23.4 *The space $\mathcal{R}(G)$ is dense in $L^2(G)$.*

Let $\mathcal{H}_\alpha$ be a collection of Hilbert spaces, indexed by a set $\mathcal{A}$. Then the algebraic direct sum

$$\bigoplus_{\alpha\in\mathcal{A}} \mathcal{H}_\alpha$$

is a pre-Hilbert space when equipped with the direct sum inner product: $\langle \sum_\alpha v_\alpha \, , \, \sum_\alpha w_\alpha \rangle = \sum_\alpha \langle v_\alpha \, , \, w_\alpha \rangle$. Its completion is called the Hilbert direct sum of the spaces $\mathcal{H}_\alpha$, and denoted by

$$\widehat{\bigoplus_{\alpha\in\mathcal{A}}} \mathcal{H}_\alpha. \tag{30}$$

This completion may be realized as the space of sequences $v = (v_\alpha)_{\alpha\in\mathcal{A}}$ with $v_\alpha \in \mathcal{H}_\alpha$ and

$$\|v\|^2 = \sum_{\alpha\in\mathcal{A}} \|v_\alpha\|^2 < \infty.$$

Its inner product is given by

$$\langle v \, , \, w \rangle = \sum_{\alpha\in\mathcal{A}} \langle v_\alpha \, , \, w_\alpha \rangle.$$

If π_α is a unitary representation of G in $\mathcal{H}_\alpha$, for every $\alpha \in \mathcal{A}$, then the direct sum of the π_α extends to a unitary representation of G in (30). We call this representation the Hilbert sum of the π_α.

Theorem 23.5 (The Peter-Weyl Theorem). *The space $L^2(G)$ decomposes as the Hilbert sum*

$$L^2(G) = \widehat{\bigoplus_{\delta\in\widehat{G}}} \; C(G)_\delta,$$

each of the summands being an irreducible invariant subspace for the representation $R \times L$ of $G \times G$.

Proof: This follows from Propositions 23.3 and 23.4. $\square$

Exercise 23.6 Fix, for every (equivalence class of an) ireducible unitary representation (δ, V_δ) an orthonormal basis $e_{\delta,1}, \ldots, e_{\delta,\dim(\delta)}$. Denote the matrix coefficient associated to $e_{\delta,i}$ and $e_{\delta,j}$ by $m_{\delta,ij}$. Use Schur orthogonality and the Peter-Weyl theorem to show that the functions

$$\sqrt{\dim(\delta)}\, m_{\delta,ij} \qquad \delta \in \widehat{G}, \;\; 1 \le i, j \le \dim(\delta)$$

constitute a complete orthonormal system for $L^2(G)$.

24 Appendix: compact self-adjoint operators

Definition 24.1 Let V, W be Banach spaces. A linear map $T : V \to W$ is said to be *compact* if the image $T(B)$ of the unit ball $B = B(0; 1) \subset V$ has compact closure in W.

A compact operator $T : V \to W$ is obviously bounded. The set of compact operators forms a linear subspace of the space $L(V, W)$ of bounded linear operators $V \to W$. The latter space is a Banach space for the operator norm.

Lemma 24.2 *Let V, W be Banach spaces, and let $L(V, W)$ be the Banach space of bounded linear operators $V \to W$, equipped with the operator norm. Then the subspace of compact linear operators $V \to W$ is closed in $L(V, W)$.*

Proof: See a standard textbook on functional analysis. □

Remark 24.3 A linear map $T : V \to W$ is said to be of finite rank if its image $T(V)$ is finite dimensional. Clearly an operator of finite rank is compact. Thus, if T_j is a sequence of operators in $L(V, W)$ all of which are of finite rank, and if $T_j \to T$ with respect to the operator norm, then it follows from the above result that T is compact.

We recall that a bounded linear operator T from a complex Hilbert space $\mathcal{H}$ to itself is said to self-adjoint if $T^* = T$, or, equivalently, if $\langle Tv\,,\, w\rangle = \langle v\,,\, Tw\rangle$ for all $v, w \in \mathcal{H}$.

We now recall the important *spectral theorem* for compact self-adjoint operators in Hilbert space. It will play a crucial role in the proof of the Peter-Weyl theorem in the next section. For a proof of the spectral theorem, we refer to a standard text book on functional analysis.

Theorem 24.4 *Let T be a compact self-adjoint operator in the (complex) Hilbert space $\mathcal{H}$. Then there exists a discrete subset $\Lambda \subset \mathbb{R} \setminus \{0\}$ such that the following hold.*

(a) *For every $\lambda \in \Lambda$ the associated eigenspace $\mathcal{H}_\lambda$ of T in $\mathcal{H}$ is finite dimensional.*

(b) *If $\lambda, \mu \in \Lambda, \lambda \neq \mu$ then $\mathcal{H}_\lambda \perp \mathcal{H}_\mu$.*

(c) *For every $\lambda \in \Lambda$, let P_λ denote the orthogonal projection $\mathcal{H} \to \mathcal{H}_\lambda$. Then*

$$T = \sum_{\lambda \in \Lambda} \lambda\, P_\lambda,$$

the convergence being absolute with respect to the operator norm.

(d) *The set Λ is bounded in $\mathbb{R}$ and has 0 as its only limit point.*

We will end this section by describing a nice class of compact self-adjoint operators in $L^2(G)$, for G a compact Lie group. First we examine the space of compactly supported continuous functions on product space.

Let X, Y be locally compact topological Hausdorff spaces. If $\varphi \in C(X)$, and $\psi \in C(Y)$, then we write $\varphi \otimes \psi$ for the continuous function on $X \times Y$ defined by:

$$\varphi \otimes \psi : (x, y) \mapsto \varphi(x)\psi(y).$$

The linear span of such functions in $C(X \times Y)$ is denoted by $C(X) \otimes C(Y)$. If $\varphi \in C_c(X)$ and $\psi \in C_c(Y)$ then $\varphi \otimes \psi$ is compactly supported. Hence the span $C_c(X) \otimes C_c(Y)$ of such functions is a subspace of $C_c(X \times Y)$.

Proposition 24.5 *Let X, Y be locally compact Hausdorff spaces. Then for every open subset $U \subset X \times Y$ with compact closure, every $\Phi \in C_c(U)$ and every $\epsilon > 0$, there exists a function $\varphi \in C_c(X) \otimes C_c(Y)$ with* $\text{supp}\varphi \subset U$ *and* $\sup_{z \in U} |\Phi(z) - \varphi(z)| < \epsilon$. *In particular, the space $C_c(X) \otimes C_c(Y)$ is dense in $C_c(X \times Y)$.*

Proof: Using C_c-partitions of unity for X and Y, we see that we may reduce to the case that $U = U_X \times U_Y$, with U_X and U_Y open neighborhoods with compact closures in X and Y respectively.

Fix $\Phi \in C_c(X \times Y)$, with $K = \text{supp}\Phi \subset U$. Then, by compactness, $K \subset K_X \times K_Y$ for compact subsets $K_X \subset U_X$ and $K_Y \subset U_Y$. Let $\epsilon > 0$. Then by compactness there exists a finite open covering $\{V_j\}$ of K_X such that for every j and all $x_1, x_2 \in V_j$, $y \in K_Y$ one has

$$\Phi(x_1, y) - \Phi(x_2, y) < \epsilon.$$

Without loss of generality we may assume that $V_j \subset U_X$ for all j. Select a partition of unity $\{\varphi_j\}$ which is subordinate to the covering $\{V_j\}$, and fix for every j a point $\xi_j \in V_j$. Let $x \in K_X$, $y \in K_Y$. If j is such that $x \in V_j$, then $|\Phi(x_j, y) - \Phi(x, y)| < \epsilon$. It follows from this that

$$\begin{aligned}
|[\sum_j \varphi_j(x)\Phi(x_j, y)] - \Phi(x, y)| &= |\sum_j [\varphi_j(x)\Phi(x_j, y) - \varphi_j(x)\Phi(x, y)]| \\
&\leq \sum_j \varphi_j(x)|\Phi(x_j, y) - \Phi(x, y)| \\
&< \sum_j \varphi_j(x)\epsilon = \epsilon.
\end{aligned}$$

Hence, if we put $\psi_j(y) = \Phi(x_j, y)$, then

$$\| \sum_j \varphi_j \otimes \psi_j - \Phi\|_\infty < \epsilon.$$

Moreover, $\text{supp}\varphi_j \otimes \psi_j \subset U_X \times U_Y \subset U_X \times K_Y \subset U$. $\square$

Let now G be a Lie group. We fix a left invariant density dx on G and equip $G \times G$ with the left invariant product of dx with itself. This *product density*, denoted $dxdy$, is determined by the formula

$$\int_{G \times G} f(x, y)\, dxdy = \int_G \left(\int_G f(x, y) dx \right) dy = \int_G \left(\int_G f(x, y) dy \right) dx,$$

for $f \in C_c(X \times Y)$.

If $K \in C_c(G \times G)$, then we define the linear operator $T_K : C_c(G) \to C_c(G)$ by

$$T_K(\varphi)(x) = \int_G K(x, y)\varphi(y)dy.$$

For obvious reasons this is called an *integral operator* with *kernel* K.

Lemma 24.6 *Let $K \in C_c(G \otimes G)$. Then the operator T_K extends uniquely to a bounded linear endomorphism of $L^2(G)$ with operator norm $\|T_K\|_{\rm op} \leq \|K\|_2$. Moreover, this extension is compact.*

Proof: Let $\varphi \in C_c(G)$. Then

$$\begin{aligned}
\langle T_K(\varphi)\, ,\, \psi\rangle &= \int_G T_K(\varphi)(x)\overline{\psi(x)}\; dx \\
&= \int_G \left(\int_G K(x, y)\varphi(y)\; dy\right)\overline{\psi(x)}\; dx \\
&= \langle K\, ,\, \varphi \otimes \bar{\psi}\rangle_{L^2(G\times G)} \\
&\leq \|K\|_{L^2(G\times G)}\|\varphi \otimes \bar{\psi}\|_{L^2(G\times G)} = \|K\|_2\|\varphi\|_2\|\psi\|_2.
\end{aligned}$$

Hence $\|T_K\varphi\|_2 \leq \|K\|_2\|\varphi\|_2$. This implies the first assertion, since $C_c(G)$ is dense in $L^2(G)$.

For the second assertion, note that by Proposition 24.5 there exists a sequence K_j in $C_c(G)\otimes C_c(G)$ which converges to K with respect to the L^2-norm on $G \times G$. It follows that

$$\|T_{K_j} - T_K\|_{\rm op} \leq \|K_j - K\|_2 \to 0.$$

Every operator T_{K_j} has a finite dimensional image hence is compact. The subspace of compact endomorphisms of $L^2(G)$ is closed for the operator norm, by Lemma 24.2. Therefore, T_K is compact. □

Let G be a Lie group, equipped with a left invariant density dx. If (π, V) is a continuous finite dimensional representation of G, then for $f \in C_c(G)$ we define the linear operator $\pi(f) : V \to V$ by

$$\pi(f)v = \int_G f(x)\pi(x)v\; dx.$$

Referring to integration with values in a Banach space, this definition actually makes sense if π is a continuous representation in a Banach space; it is readily seen that then $\pi(f)$ is a continuous linear operator. In particular, the definition may be applied to the regular representations L and R of G in $L^2(G)$. Thus, for $f \in C_c(G)$ and $\varphi \in L^2(G)$,

$$[R(f)\varphi](x) = \int_G f(y)\varphi(xy)\; dy = \int_G f(x^{-1}y)\varphi(y)\; dy \qquad (x \in G). \tag{31}$$

Of course, this formula can also be used as the defining formula, without reference to Banach-valued integration.

Corollary 24.7 *Assume that G is compact, and let $f \in C(G)$. Then the operator $R(f) : L^2(G) \to L^2(G)$ is compact.*

Proof: If $\varphi \in C(G)$, then from (31) we see that $R(f) = T_K$, with $K(x, y) = f(x^{-1}y)$. The result now follows by application of Lemma 24.6. $\square$

Remark 24.8 Note that for this argument it is crucial that G is compact. For if not, and $f \in C_c(G)$, then the associated integral kernel K need not be compactly supported.

The following lemmas will in particular be needed for the right regular representation R.

Lemma 24.9 *Let $(\pi, \mathcal{H})$ be a unitary representation of G in a Hilbert space. Let $f \in C_c(G)$, then*

$$\pi(f)^* = \pi(f^*),$$

where $f^*(x) = \overline{f(x^{-1})}$.

Proof: Straightforward and left to the reader. $\square$

Lemma 24.10 *Let π be a continuous representation of G in a Banach space V. If $f \in C_c(G)$ is conjugation invariant, then $\pi(f)$ is intertwining.*

Proof: Straightforward and left to the reader. $\square$

Corollary 24.11 *Assume that G is compact, and let $f \in C(G)$ be such that $f^* = f$. Then $R(f)$ (and $L(f)$ as well) is a compact self-adjoint operator. If, in addition, f is conjugation invariant then $R(f)$ is G-equivariant.*

Proof: This follows by combining Corollary 24.11 and Lemmas 24.9 and 24.10. $\square$

25 Proof of the Peter-Weyl Theorem

In the beginning of this section we assume that G is any Lie group. At a later stage we will restrict our attention to compact G. We assume that G is equipped with a positive left invariant density dx.

Lemma 25.1 *Let $\varphi \in C_c(G)$. Then $R(\varphi)$ maps $L^2(G)$ into $C(G)$.*

Proof: Let $x_0 \in G$ and let $\epsilon > 0$. Since φ has compact support $C := \mathrm{supp}\varphi$, it follows by the principle of uniform continuity that there exists a compact neighborhood U of e in G such that $|\varphi(u) - \varphi(v)| < \epsilon(2\|1_C\|_2 + 1)^{-1}$ for all $u, v \in G$ with $vu^{-1} \in U$.

Let now $f \in L^2(G)$. For $x, y \in G$ with $x \in x_0U$ we have $(x_0^{-1}y)(x^{-1}y)^{-1} = x_0^{-1}x \in U$, hence

$$\begin{aligned} |R(\varphi)f(x) - R(\varphi)f(x_0)| &= |\int_G [\varphi(x^{-1}y) - \varphi(x_0^{-1}y)]f(y)\, dy| \\ &\leq \int_{xC\cup x_0C} \epsilon\,|f(y)|\, dy = \epsilon \int_G 1_{xC\cup x_0C}|f(y)|\, dy \\ &\leq \epsilon\,\|1_{xC\cup yC}\|_2\,\|f\|_2 \leq 2\epsilon\,\|1_C\|_2\,\|f\|_2 \leq \epsilon\|f\|_2. \end{aligned}$$

From this we deduce that $R(\varphi)f$ is continuous in x_0. $\square$

Lemma 25.2 *Let $f \in L^2(G)$ and let $\epsilon > 0$. There exists an open neighborhood U of e in G such that for all $x \in U$ we have $\|R_x f - f\|_2 < \epsilon$. Moreover, if U is any neighborhood with this property and if $\varphi \in C_c(U)$ satisfies $\varphi \geq 0$ and $\int_G \varphi(x)\, dx = 1$, then*

$$\|R(\varphi)f - f\|_2 < \epsilon. \tag{32}$$

Proof: The first assertion follows from the continuity of the map $x \mapsto R_x f$, see Proposition 20.10. Let U, φ be as stated. Then, for all $x \in G$,

$$R(\varphi)f(x) - f(x) = \int_G \varphi(y)[f(xy) - f(x)]\, dy.$$

Hence, for every $g \in L^2(G)$ we have

$$\begin{aligned} |\langle R(\varphi)f - f\,,\, g\rangle| &\leq \int_G\int_G \varphi(y)\,|f(xy) - f(x)|\, g(x)\, dy\, dx \\ &= \int_G\int_G \varphi(y)\,|f(xy) - f(x)|\, g(x)\, dx\, dy \\ &\leq \int_G \varphi(y)\,\|R_y f - f\|_2\|g\|_2\, dy \\ &\leq \epsilon\,\|g\|_2. \end{aligned}$$

From this the estimate (32) follows. $\square$

From now on we assume that the group G is compact.

Lemma 25.3 *Let V be a finite dimensional right G-invariant subspace of $L^2(G)$. Then $V \subset \mathcal{R}(G)$.*

Proof: Decomposing V into a direct sum of irreducible subspaces, we see that we may reduce the case that V is irreducible. We claim that V consists of continuous functions. For this we observe that $C(G) \cap V$ is an invariant subspace. Hence it suffices to show that V contains a non-trivial continuous function. Fix $f \in V \setminus \{0\}$ and fix $0 < \epsilon < 1/2\|f\|_2$. Choose U and φ as in Lemma 25.2. Then $\|R(\varphi)f\| > 1/2\epsilon$, hence $R(\varphi)f \neq 0$. From Lemma 25.1 it follows that $R(\varphi)f \in C(G)$. Moreover, since V is right invariant, it follows that $R(\varphi)f \in V$. This establishes the claim that $V \subset C(G)$.

Choose an orthonormal basis (ψ_i) of V. Then for $f \in V$ we have

$$R_x f = \sum_i \langle R_x f\,,\,\psi_i\rangle \psi_i,$$

hence by evaluation in e,

$$f(x) = \sum_i \langle R_x f\,,\,\psi_i\rangle \psi_i(e).$$

By definition of $\mathcal{R}(G)$ it now follows that $f \in \mathcal{R}(G)$. □

Lemma 25.4 *Let U be an open neighborhood of e in G. Then there exists a $\varphi \in C_c(U)$ such that:*

(a) $\varphi \geq 0$ *and* $\int_G \varphi(x)\,dx = 1$;

(b) $\varphi^* = \varphi$;

(c) φ *is conjugation invariant.*

Proof: From the continuity of the map $x \mapsto x^{-1}$ one sees that there exists a compact neighborhood V of e such that $V \subset U$ and $V^{-1} \subset U$. For every $x \in G$ there exist an open neighborhood N_x of x and a compact neighborhood V_x of e in V such that $zyz^{-1} \in V$ for all $z \in N_x$, $y \in V_x$. By compactness of G finitely many of the N_x cover G. Let Ω be the intersection of the corresponding V_x. Then Ω is a compact neighborhood of e and for all $x \in G$ and $y \in \Omega$ we have $xyx^{-1} \in V$.

Now select $\psi_0 \in C_c(\Omega)$ such that $\psi_0 \geq 0$ and $\int_G \psi_0(x)\,dx = 1$. Define

$$\psi(x) = \int_G \psi_0(yxy^{-1})\,dy.$$

Since $(x, y) \mapsto \psi(yxy^{-1})$ is a continuous function, it follows that ψ is a continuous function. Clearly $\psi \geq 0$. Moreover, by interchanging the order of integration, and using the fact that dx is bi-invariant and normalized, we deduce that $\int_G \psi(x)\,dx = 1$. If $\psi(x) \neq 0$, then $yxy^{-1} \in \mathrm{supp}\psi_0$ for some $y \in G$, hence $x \in \cup_{y\in G} y^{-1}\Omega y \subset V$. It follows that $\mathrm{supp}\psi \subset V$. One now readily verifies that the function $\varphi = \frac{1}{2}(\psi + \psi^*)$ satisfies all our requirements. □

Corollary 25.5 *Let $f \in L^2(G)$, $f \neq 0$. Then there exists a left and right G-equivariant bounded linear operator $T : L^2(G) \to L^2(G)$ with:*

(a) $Tf \neq 0$.

(b) *T is self-adjoint and compact;*

(c) *T maps every right G-invariant closed subspace of $L^2(G)$ into itself.*

Proof: Let $\epsilon = \frac{1}{2}\|f\|_2$, and fix an open neighborhood U of e in G that satisfies the assertion of Lemma 25.2 Let $\varphi \in C_c(U)$ be as in Lemma 25.4, and define $T = R(\varphi)$. Then $\|Tf - f\| < \epsilon$, hence (a). Moreover, every closed right invariant subspace V of $L^2(G)$ equipped with the restriction of R is a continuous representation in a Banach space, hence invariant under $T = R(\varphi)$. This implies (c).

The operator T is left G-equivariant, since L and R commute. It is right G-equivariant because φ is conjugation invariant, cf. Lemma 24.10. Finally (b) follows from Corollary 24.11. $\square$

Proof of Propostion 23.4. The space $\mathcal{R}(G)$ is left and right G-invariant, and by unitarity so is its orthocomplement V. Suppose that V contains a non-trivial element f. Let T be as in Corollary 25.5. Then $T|_V : V \to V$ is a non-trivial compact self-adjoint operator which is both left and right G-equivariant. By the spectral theorem for compact self-adjoint operators, Theorem 24.4, there exists a $\lambda \in \mathbb{R}$, $\lambda \neq 0$, such that the eigenspace $V_\lambda = \ker(T - \lambda \mathrm{I}_V)$ is non-trivial. By compactness of T the eigenspace V_λ is finite dimensional, and by equivariance of T it is both left and right G-invariant. By Lemma 25.3 it now follows that $V_\lambda \subset \mathcal{R}(G)$, contradiction. Therefore, V must be trivial. $\square$

26 Class functions

By a *class function* on a compact Lie group G we mean a function $f : G \to \mathbb{C}$ that is conjugation invariant, i.e., $L_x R_x f = f$ for all $x \in G$. The name class function comes from the fact that a conjugation invariant function is constant on the conjugacy classes, hence may be viewed as a function on the set of conjugacy classes.

The space $C(G, \text{class})$ of continuous class functions is a closed subspace of $C(G)$ (with respect to the sup norm). Its closure in $L^2(G)$ equals $L^2(G, \text{class})$, the space of square integrable class functions on G.

If $\delta \in \widehat{G}$, we denote the orthogonal projection from $L^2(G)$ onto the finite dimensional subspace $C(G)_\delta$ by

$$P_\delta : L^2(G) \to C(G)_\delta$$

Note that P_δ is equivariant for both the representations R and L of G. In particular, this implies that P_δ maps $C(G, \text{class})$ into its intersection with $C(G)_\delta$. Hence, by Lemma 21.8

$$P_\delta(C(G, \text{class})) = C(G)_\delta \cap C(G, \text{class}) = \mathbb{C}\chi_\delta.$$

It follows from this that the space $\mathcal{R}(G, \text{class}) = C(G, \text{class}) \cap \mathcal{R}(G)$ of representative class functions is the linear span of the characters χ_δ, $\delta \in \widehat{G}$.

Lemma 26.1 *Let G be a compact Lie group. Then the characters χ_δ, $\delta \in \widehat{G}$, form a complete orthonormal system for $L^2(G, \text{class})$.*

Proof: By Schur orthogonality, the characters form an orthonormal system. To establish its completeness, let $f \in L^2(G, \text{class})$ and assume that $f \perp \chi_\delta$ for all $\delta \in \widehat{G}$.

From $P_\delta f \in C(G)_\delta = \mathbb{C}\chi_\delta$, we see that

$$P_\delta f = \langle P_\delta f\,,\, \chi_\delta \rangle\, \chi_\delta = \langle f\,,\, \chi_\delta \rangle\, \chi_\delta = 0.$$

Hence $f \perp \mathcal{R}(G)$. By the Peter-Weyl theorem, the latter implies that $f = 0$. $\square$

Corollary 26.2 *Let $f \in L^2(G, \text{class})$. Then*

$$f = \sum_{\delta \in \widehat{G}} \langle f\,,\, \chi_\delta \rangle\, \chi_\delta,$$

with convergence in the L^2-norm.

27 Abelian groups and Fourier series

In this section we consider the special case that the compact Lie group G is commutative. If, in addition, G is connected, then $G \simeq \mathbb{R}^n/\mathbb{Z}^n$ for some $n \in \mathbb{N}$, and we will see that the Peter-Weyl theorem specializes to the theory of Fourier series.

By a *multiplicative character* of G we mean a continuous (hence smooth) group homomorphism $\xi : G \to \mathbb{C}^*$, where $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ is equipped with complex multiplication. By the lemma below, if ξ is a multiplicative character, then $|\xi(x)| = 1$, $x \in G$.

Lemma 27.1 *Let H be a compact subgroup of $\mathbb{C}^*$. Then $H \subset \mathbb{T}$.*

Proof: By compactness, there exists a constant $r > 0$ such that $r^{-1} < |z| < r$ for all $z \in H$. Let $w \in H$, then applying the estimate to $z = w^n$ we obtain that $r^{-1/n} \leq |w| \leq r^{1/n}$. Taking the limit for $n \to \infty$ we see that $|w| = 1$. $\square$

Lemma 27.2 *Let G be a commutative compact Lie group. If (δ, V_δ) is a finite dimensional irreducible representation of G, then $\dim V_\delta = 1$. Moreover, $\delta(x) = \chi_\delta(x) I_{V_\delta}$. The map $\delta \mapsto \chi_\delta$ induces a bijection from $\widehat{G}$ onto the set of multiplicative characters of G.*

Proof: If $x \in G$, then $\delta(y)\delta(x) = \delta(yx) = \delta(xy) = \delta(x)\delta(y)$ for all $y \in G$, hence $\delta(x)$ is equivariant, and it follows that

$$\delta(x) = \xi(x) I, \tag{33}$$

for some $\xi(x) \in \mathbb{C}$, by Schur's lemma. It follows from this that every linear subspace of V_δ is invariant. By irreducibility of δ this implies that the dimension of V_δ must be one. From the fact that δ is a representation it follows immediately that $x \mapsto \xi(x)$ is a character. Applying the trace to (33) we see that $\xi = \chi_\delta$, the character of δ. Thus $\delta \mapsto \chi_\delta$ induces a map from the space $\widehat{G}$ of equivalence classes of finite dimensional irreducible representations to the set of multiplicative characters of G. This map is injective by Corollary 22.13. If ξ is a multiplicative character then (33) defines an irreducible representation δ of G in $\mathbb{C}$, and $\xi = \chi_\delta$. Therefore the map $\delta \to \chi_\delta$ is surjective onto the set of multiplicative characters. $\square$

Corollary 27.3 *Assume that G is a commutative compact Lie group. Then the set of multiplicative characters $\chi_\delta,\ \delta \in \widehat{G}$, is a complete orthonormal system for $L^2(G)$.*

Proof: This follows immediately from the previous lemma combined with the theorem of Peter and Weyl (Theorem 23.5). $\square$

In the present setting we define the Fourier transform $\hat{f} : \widehat{G} \to \mathbb{C}$ of a function $f \in L^2(G)$ by

$$\hat{f}(\delta) = \langle f \,,\, \chi_\delta \rangle.$$

Let $\widehat{G}$ be equipped with the counting measure. Then the associated L^2-space is $l^2(\widehat{G})$, the space of functions $\varphi : \widehat{G} \to \mathbb{C}$ such that $\sum_{\delta \in \widehat{G}} |\varphi(\delta)|^2 < \infty$, equipped with the inner product:

$$\langle \varphi \,,\, \psi \rangle := \sum_{\delta \in \widehat{G}} \varphi(\delta)\, \overline{\psi(\delta)}.$$

Corollary 27.4 (The Plancherel theorem). *Let G be a commutative compact Lie group. Then the Fourier transform $f \mapsto \hat{f}$ is an isometry from $L^2(G)$ onto $l^2(\widehat{G})$. Moreover, if $f \in L^2(G)$, then*

$$f = \sum_{\delta \in \widehat{G}} \hat{f}(\delta)\, \chi_\delta,$$

with convergence in the L^2-sense.

Proof: Exercise for the reader. $\square$

If in addition it is assumed that the group G is connected, then $G \simeq (\mathbb{R}/\mathbb{Z})^n$ for some $n \in \mathbb{N}$, see Theorem 6.1. The purpose of the following exercise is to accordingly view the classical theory of Fourier series as a special case of the Peter-Weyl theory.

Exercise 27.5 Let $G = \mathbb{R}^n/2\pi\mathbb{Z}^n$. If $m \in \mathbb{Z}^n$, show that

$$\chi_m : x \mapsto e^{i(m \cdot x)}$$

defines a multiplicative character of G. (Here $m \cdot x = m_1x_1 + \cdots + m_nx_n$.) Show that every multiplicative character is of this form. Thus $\widehat{G} \simeq \mathbb{Z}^n$. Accordingly for $f \in L^2(G)$ we view the Fourier transform $\hat{f}$ as a map $\mathbb{Z}^n \to \mathbb{C}$.

Show that the normalized Haar integral of G is given by

$$I(f) = \frac{1}{(2\pi)^n} \int_0^{2\pi} \dots \int_0^{2\pi} f(x_1, \dots, x_n)\, dx_1 \dots dx_n.$$

Show that for $f \in L^2(G)$, $m \in \mathbb{Z}^n$ we have:

$$\hat{f}(m) = \frac{1}{(2\pi)^n} \int_0^{2\pi} \cdots \int_0^{2\pi} f(x_1, \dots, x_n)\, e^{-i(m_1 x_1 + \cdots + m_n x_n)}\, dx_1 \dots dx_n.$$

Moreover, show that we have the inversion formula

$$f(x) = \sum_{m \in \mathbb{Z}^n} \hat{f}(m)\, e^{i(m \cdot x)} \qquad (x \in \mathbb{R}^n/2\pi\mathbb{Z}^n)$$

in the L^2-sense.

28 The group SU(2)

In this section we assume that G is the compact Lie group $\mathrm{SU}(2)$.

Recall the definition of the representation π_n of $\mathrm{SU}(2)$ in the space $V_n = P_n(\mathbb{C}^2)$ of homogeneous polynomials of degree n from Section 20. In Proposition 20.28 it was shown that π_n is irreducible. Moreover, the associated character is determined by the formula:

$$\chi_n(t_\varphi) = \frac{\sin(n+1)\varphi}{\sin\varphi} \qquad (\varphi \in \mathbb{R}) \tag{34}$$

(see Exercise 22.9). The purpose of this section is to prove the following result:

Proposition 28.1 *Every finite dimensional irreducible representation of* $\mathrm{SU}(2)$ *is equivalent to* π_n*, for some* $n \in \mathbb{N}$.

We recall that every element of $\mathrm{SU}(2)$ is conjugate to an element of $T = \{t_\varphi \mid \varphi \in \mathbb{R}\}$. Therefore a class function f on $\mathrm{SU}(2)$ is completely determined by its restriction $f|_T$ to T. For every $\varphi \in \mathbb{R}$ the diagonal matrices t_φ and $t_\varphi^{-1} = t_{-\varphi}$ are conjugate. Therefore, the restriction $f|_T$ is invariant under the substitution $t \mapsto t^{-1}$. Thus, if $C(T)_{\mathrm{ev}}$ denotes the space of continuous functions $g : T \to \mathbb{C}$ satisfying $g(t^{-1}) = g(t)$ for all $t \in T$, then restriction to T defines an injective linear map $r : C(G, \mathrm{class}) \to C(T)_{\mathrm{ev}}$.

Lemma 28.2 *The map* $r : C(G, \mathrm{class}) \to C(T)_{\mathrm{ev}}$ *is bijective. Moreover,* r *is isometric, i.e., it preserves the sup-norms.*

Proof: That r is isometric follows from the observation that the set of values of a function $f \in C(G, \mathrm{class})$ is equal to the set of values of its restriction $r(f)$. Thus it remains to establish the surjectivity of r. Let $g \in C(T)_{\mathrm{ev}}$. Then $g(t_\varphi) = \tilde{g}(e^{i\varphi})$ for a unique continuous function $\tilde{g} : \mathbb{T} \to \mathbb{C}$ satisfying $\tilde{g}(z) = \tilde{g}(z^{-1})$. Now $\tilde{g}(z) = G(\mathrm{Re} z)$ for a unique continuous function $G : [-1, 1] \to \mathbb{C}$. It follows that $g(t_\varphi) = G(\cos\varphi)$, for $\varphi \in \mathbb{R}$.

An element $x \in \mathrm{SU}(2)$ has two eigenvalues $z(x)$ and $z(x)^{-1}$, with $|z(x)| = 1$. Clearly $x \mapsto \mathrm{Re} z(x)$ is a well defined continous function on $\mathrm{SU}(2)$.

Define $f(x) = G(\mathrm{Re} z(x))$. Then f is a well defined continuous class function. Moreover, $f(t_\varphi) = G(\mathrm{Re} e^{i\varphi}) = g(t_\varphi)$, hence $r(f) = g$. $\square$

Corollary 28.3 *The linear span of the characters χ_n, for $n \in \mathbb{N}$, is dense in $C(G, \text{class})$.*

Proof: By Lemma 28.2 it suffices to show that the linear span S of the functions $\chi_n|_T$ is dense in $C(T)_{\text{ev}}$. From formula (34) we see that $\chi_n(t_\varphi) = \sum_{k=0}^{n} e^{i(n-2k)\varphi}$. Hence S equals the linear span of the functions $\gamma_n : t_\varphi \mapsto e^{in\varphi} + e^{-in\varphi} = 2\cos n\varphi$, $(n \in \mathbb{N})$. The latter span is dense in $C(T)_{\text{ev}}$, by the classical theory of Fourier series. $\square$

Corollary 28.4 *Let $f \in C(G, \text{class})$. If $f \perp \chi_n$ for all $n \in \mathbb{N}$, then $f = 0$.*

Remark 28.5 Once we know that the π_n exhaust $\widehat{G}$ this follows from the Peter-Weyl theorem.

Proof: We first note that, for $g \in C(G)$,

$$\|g\|_{L^2}^2 = \int_G |g(x)|^2 \, dx \leq \|g\|_\infty^2,$$

where $\|\cdot\|_\infty$ denotes the sup norm. Using this estimate we see that the linear span of the characters χ_n is dense in $C(G, \text{class})$ with respect to the L^2-norm. Thus, if $f \in C(G, \text{class})$ is perpendicular to all χ_n, then it follows that $f \perp C(G, \text{class})$. In particular, $\|f\|_2^2 = \langle f, f \rangle = 0$, which implies that $f = 0$. $\square$

Corollary 28.6 *Every finite dimensional irreducible representation of* $\mathrm{SU}(2)$ *is equivalent to one of the π_n, $n \in \mathbb{N}$.*

Proof: Suppose not. Then there exists a $\delta \in \widehat{G}$ such that δ is not equivalent to π_n, for every $n \in \mathbb{N}$. Hence the class function χ_δ is perpendicular to χ_n for every $n \in \mathbb{N}$. This implies that $\chi_\delta = 0$. This is impossible, since $\chi_\delta(e) = \dim(\delta) \geq 1$. $\square$

From the fact that every element of SU(2) is conjugate to an element of T one might expect that there should exist a Jacobian $J : T \to [0, \infty[$ such that for every continuous class function f on SU(2) we have

$$\int_{\mathrm{SU}(2)} f(x) \, dx = \int_0^{2\pi} f(t_\varphi) \, J(t_\varphi) \, d\varphi.$$

It is indeed possible to compute this Jacobian by a substitution of variables. However, we shall obtain the above integration formula by other means.

Lemma 28.7 *For every continuous class function $f : \mathrm{SU}(2) \to \mathbb{C}$ we have:*

$$\int_{\mathrm{SU}(2)} f(x) \, dx = \int_0^{2\pi} f(t_\varphi) \, \frac{\sin^2 \varphi}{\pi} \, d\varphi. \tag{35}$$

Proof: Consider the linear map L which assigns to $f \in C(G, \text{class})$ the expression on the left-hand side minus the expression on the right-hand side of the above equation. Then we must show that L is zero.

Obviously the linear functional $L : C(G, \text{class}) \to \mathbb{C}$ is continuous with respect to the sup norm. Hence by density of the span of the characters it suffices to show that $L(\chi_n) = 0$ for every $n \in \mathbb{N}$. The function χ_0 is identically one; therefore left- and right-hand side of (35) both equal 1 if one substitutes $f = \chi_0$. Hence $L(\chi_0) = 0$. On the other hand, if $n \geq 1$, and $f = \chi_n$, then the left hand side of (35) equals $\langle \chi_n \,,\, \chi_0 \rangle = 0$. The right hand side of (35) also equals 0, hence $L(\chi_n) = 0$ for all n. □

Corollary 28.8 *Let $f \in C(G)$. Then*

$$\int_G f(x)\; dx = \int_0^{2\pi} \int_G f(x t_\varphi x^{-1})\, dx\, \frac{\sin^2 \varphi}{\pi}\; d\varphi$$

Remark 28.9 The interpretation of the above formula is that the integration over $G = \mathrm{SU}(2)$ may be split into an integration over conjugacy classes, followed by an integration over the circle group T.

Proof: Put

$$F(y) = \int_G f(x y x^{-1})\; dx.$$

Then by bi-invariance of the Haar measure, F is a continuous class function. Hence by the previous result

$$\int_G F(y)\; dy = \int_0^{2\pi} \int_G f(x t_\varphi x^{-1})\, dx\, \frac{\sin^2 \varphi}{\pi}\; d\varphi.$$

On the other hand,

$$\begin{aligned} \int_G F(y)\; dy &= \int_G \int_G f(x y x^{-1})\, dx\, dy \\ &= \int_G \int_G f(x y x^{-1})\, dy\, dx, \end{aligned}$$

by Fubini's theorem. By bi-invariance of the Haar measure, the inner integral is independent of x. Therefore,

$$\begin{aligned} \int_G F(y)\; dy &= \int_G \int_G f(y)\, dy\, dx \\ &= \int_G f(y)\, dy. \end{aligned}$$

This completes the proof. □

We end this section with a description of all irreducible representations of SO(3). From Section 10 we recall that there exists a surjective Lie group homomorphism $\varphi : \mathrm{SU}(2) \to \mathrm{SO}(3)$ with kernel $\ker \varphi = \{-I, I\}$. Accordingly, $\mathrm{SO}(3) \simeq \mathrm{SU}(2)/\{\pm I\}$ (Thm. 17.4).

Proposition 28.10 *For $k \in \mathbb{N}$ the representation π_{2k} of* SU(2) *factors through a representation $\bar{\pi}_{2k}$ of* $\mathrm{SO}(3) \simeq \mathrm{SU}(2)/\{\pm I\}$. *The representations $\bar{\pi}_{2k}$ are mutually inequivalent and exhaust* $\widehat{\mathrm{SO}(3)}$.

Proof: One readily verifies that $\pi_{2k}(x) = I$ for $x \in \{\pm I\}$. Hence π_{2k} factors through a representation $\bar{\pi}_{2k}$ of SO(3). Every invariant subspace of the representation space V_{2k} of π_{2k} is $\pi_{2k}(\mathrm{SU}(2))$ invariant if and only if it is $\bar{\pi}_{2k}(\mathrm{SO}(3))$ invariant. A non-trivial SO(3)-equivariant map $V_{2k} \to V_{2l}$ would also be SU(2)-equivariant. Hence the $\bar{\pi}_{2k}$ are mutually inequivalent. Finally, to see that the representations $\bar{\pi}_{2k}$ exhaust $\widehat{\mathrm{SO}(3)}$, assume that (π, V) is an irreducible representation of SO(3). Then $\varphi^*\pi := \pi \circ \varphi$ is an irreducible representation of SU(2), hence equivalent to some π_n, $n \in \mathbb{N}$. From $\varphi^*\pi = I$ on $\ker \varphi$ it follows that $\pi_n = I$ on $\{\pm I\}$, hence n is even. $\square$

29 Lie algebra representations

Let V be a finite dimensional complex linear space. If π is a continuous representation of G in V, then π is a (smooth) Lie group homomorphism $G \to \mathrm{GL}(V)$, in view of Corollary 9.3. Accordingly, the tangent map $\pi_* : \mathfrak{g} \to \mathrm{End}(V)$ at e is a homomorphism of Lie algebras. Thus, π_* is a representation of $\mathfrak{g}$ in V. In other words, we see that a finite dimensional G-module V automatically is a $\mathfrak{g}$-module (see Remark 20.6 and the text preceding the remark for the terminology used here).

By the chain rule one readily sees that

$$\pi_*(X)v = \left.\frac{d}{dt}\right|_{t=0} \pi(\exp tX)v, \tag{36}$$

for $v \in V$ and $X \in \mathfrak{g}$. On the other hand, it follows from Lemma 4.16 that for all $X \in \mathfrak{g}$ we have:

$$\pi(\exp X) = e^{\pi_*(X)}. \tag{37}$$

When G is connected this equation allows us to compare the G- and the $\mathfrak{g}$-module structures on V. When there is no chance of confusion, we will omit the star in the notation of the representation of $\mathfrak{g}$ in V.

Lemma 29.1 *Assume that G is connected, and let V, V' be two finite dimensional G-modules.*

(a) *Let W be a linear subspace of V. Then W is G-invariant if and only if W is $\mathfrak{g}$-invariant.*

(b) *The G-module V is irreducible if and only V is irreducible as a $\mathfrak{g}$-module.*

(c) *Let $T : V \to V'$ be a linear map. Then T is G-equivariant if and only if T is $\mathfrak{g}$-equivariant.*

(d) *V and V' are isomorphic as G-modules if and only if they are isomorphic as $\mathfrak{g}$-modules.*

Proof: Write π and π' for the representations of G in V and V' respectively. As agreed, we denote the associated representations of $\mathfrak{g}$ in V and V' by the same symbols, i.e., we omit the stars in the notation.

(a): If W is $\mathfrak{g}$-invariant, then it follows from (37) that W is invariant under the group G_e which is generated by $\exp \mathfrak{g}$. But $G_e = G$, since G is connected. The converse implication is proved by differentiating $\pi(\exp(tX))$ at $t = 0$.

(b): This is now an immediate consequence of (a).

(c): Suppose that T is $\mathfrak{g}$-equivariant. Then for all $X \in \mathfrak{g}$ we have: $\pi'(X) \circ T = T \circ \pi(X)$, hence $\pi'(X)^n \circ T = T \circ \pi(X)^n$ for all $n \in \mathbb{N}$, and since T is continuous linear it follows that

$$e^{\pi'(X)} \circ T = T \circ e^{\pi(X)}.$$

From this it follows that $\pi'(x) \circ T = T \circ \pi(x)$ for all $x \in \exp \mathfrak{g}$, and hence for $x \in G_e = G$. The reverse implication follows by a straightforward differentiation argument as in part (a) of this proof.

(d): This follows immediately from (c). $\square$

Lemma 29.2 *Let G be a connected compact Lie group, and let π be a representation of G in a finite dimensional Hilbert space V. Then π is unitary if and only if*

$$\pi(X)^* = -\pi(X) \tag{38}$$

for all $X \in \mathfrak{g}$.

Proof: We recall that $\pi : G \to \mathrm{GL}(V)$ is a Lie group homomorphism. Hence for all $X \in \mathfrak{g}$, $t \in \mathbb{R}$ we have:

$$\pi(\exp tX) = e^{t\pi(X)}.$$

If π is unitary, then $\pi(\exp tX)^* = \pi(\exp(-tX))$, hence

$$e^{t\pi(X)^*} = e^{-t\pi(X)}. \tag{39}$$

Differentiating this relation at $t = 0$ we find (38). Conversely, if (38) holds, then (39) holds for all X, t and it follows that $\pi(x)$ is unitary for $x \in \exp \mathfrak{g}$. This implies that $\pi(x)$ is unitary for $x \in G_e = G$. $\square$

It will turn out to be convenient to extend representations of $\mathfrak{g}$ to its *complexification* $\mathfrak{g}_{\mathbb{C}}$. If E is a real linear space, its complexification $E_{\mathbb{C}}$ is defined as the real linear space $E \otimes_{\mathbb{R}} \mathbb{C}$, equipped with the complex scalar multiplication $\lambda(v \otimes z) = v \otimes \lambda z$. We embed E as a real linear subspace of $E_{\mathbb{C}}$ by the map $v \mapsto v \otimes 1$. Then $E_{\mathbb{C}} = E \oplus iE$ as a real linear space. In terms of this decomposition, the complex scalar multiplication is given in the obvious fashion. If $\mathfrak{g}$ is a real Lie algebra, then its complexification $\mathfrak{g}_{\mathbb{C}}$ is equipped with the complex bilinear extension of the Lie bracket. Thus, $\mathfrak{g}_{\mathbb{C}}$ is a complex Lie algebra.

Any representation ρ of $\mathfrak{g}$ in a complex vector space V has a unique extension to a (complex) representation of $\mathfrak{g}_{\mathbb{C}}$ in V; this extension, denoted $\rho_{\mathbb{C}}$, is given by

$$\rho_{\mathbb{C}}(X + iY) = \rho(X) + i\rho(Y),$$

for $X, Y \in \mathfrak{g}$.

Lemma 29.3 *Let V, V' be $\mathfrak{g}$-modules, and let $W \subset V$ a (complex) linear subspace, and $T : V \to V'$ a (complex) linear map.*

(a) *The space W is $\mathfrak{g}$-invariant if and only if it is $\mathfrak{g}_{\mathbb{C}}$-invariant.*

(b) *V is irreducible as a $\mathfrak{g}$-module if and only if it is so as a $\mathfrak{g}_{\mathbb{C}}$-module.*

(c) *T is $\mathfrak{g}$-equivariant if and only if it is $\mathfrak{g}_{\mathbb{C}}$-equivariant.*

(d) *V and V' are isomorphic as $\mathfrak{g}$-modules if and only if they are isomorphic as $\mathfrak{g}_{\mathbb{C}}$-modules.*

Proof: Left to the reader. $\square$

Example 29.4 The Lie algebra $\mathfrak{su}(2)$ of SU(2) consists of complex 2×2 matrices $A \in \mathrm{M}(2,\mathbb{C})$, satisfying $\mathrm{tr}A = 0$ and $A^* = -A$. It follows from this that $i\mathfrak{su}(2)$ is the real linear subspace of $\mathrm{M}(2,\mathbb{C})$ consisting of matrices A with $\mathrm{tr}A = 0$ and $A^* = A$. In particular, we see that $\mathfrak{su}(2) \cap i\mathfrak{su}(2) = \{0\}$. Therefore, the embedding $\mathfrak{su}(2) \hookrightarrow \mathrm{M}(2,\mathbb{C})$ extends to a complex linear embedding

$$j : \mathfrak{su}(2)_{\mathbb{C}} \hookrightarrow \mathrm{M}(2,\mathbb{C}).$$

Clearly, the image of j is contained in the Lie algebra of $\mathrm{SL}(2,\mathbb{C})$, which is given by

$$\mathfrak{sl}(2,\mathbb{C}) = \{A \in \mathrm{M}(2,\mathbb{C}) \mid \mathrm{tr}A = 0\}.$$

On the other hand, if $A \in \mathfrak{sl}(2,\mathbb{C})$, then $\frac{1}{2}(A - A^*)$ belongs to $\mathfrak{su}(2)$ and $\frac{1}{2}(A + A^*)$ belongs to $i\mathfrak{su}(2)$; summing these elements, we see that $A \in j(\mathfrak{su}(2)_{\mathbb{C}})$. Therefore, j is an isomorphism from $\mathfrak{su}(2)_{\mathbb{C}}$ onto $\mathfrak{sl}(2,\mathbb{C})$, via which we shall identify from now on.

30 Representations of sl(2,C)

It follows from the discussion in the previous section that the SU(2)-module $P_n(\mathbb{C}^2)$, for $n \in \mathbb{N}$, carries a natural structure of $\mathfrak{sl}(2,\mathbb{C})$-module. The associated representation of $\mathfrak{sl}(2,\mathbb{C})$ in $P_n(\mathbb{C}^2)$ equals $(\pi_{n*})_{\mathbb{C}}$, the complexification of π_{n*}. We shall now compute this structure in terms of the basis $p_0, \ldots, p_n$ of $P_n(\mathbb{C}^2)$ given by

$$p_j(z) = z_1^j z_2^{n-j}, \qquad (z \in \mathbb{C}^2).$$

Let $p \in P_n(\mathbb{C}^2)$. Then we recall that, for $x \in \mathrm{SU}(2)$, $[\pi_n(x)p](z) = p(x^{-1}z)$, $z \in \mathbb{C}^2$. It follows from this that, for $\xi \in \mathfrak{su}(2)$,

$$[\pi_{n*}(\xi)p](z) = \frac{d}{dt} p(e^{-t\xi}z)\bigg|_{t=0},$$

hence, by the chain rule

$$[\pi_{n*}(\xi)p](z) = \frac{\partial p}{\partial z_1}(z)(-\xi z)_1 + \frac{\partial p}{\partial z_2}(z)(-\xi z)_2.$$

The expression on the right-hand side is complex linear in ξ; hence it also gives $\xi p = (\pi_{n*})_{\mathbb{C}}(\xi)p$ for $\xi \in \mathfrak{sl}(2,\mathbb{C})$. Thus, we obtain, for $\xi \in \mathfrak{sl}(2,\mathbb{c})$ and $p \in P_n(\mathbb{C}^2)$,

$$\xi p = -[(\xi z)_1 \frac{\partial}{\partial z_1} + (\xi z)_2 \frac{\partial}{\partial z_2}]p. \tag{40}$$

We shall now compute the action of the basis H, X, Y of $\mathfrak{sl}(2,\mathbb{C})$ given by

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

By a straightforward computation we see that

$$[H, X] = 2X, \quad [H, Y] = -2Y, \quad [X, Y] = H. \tag{41}$$

Definition 30.1 Let $\mathfrak{l}$ be a Lie algebra. By a *standard* $\mathfrak{sl}(2)$*-triple* in $\mathfrak{l}$ we mean a collection of linear independent elements $H, X, Y \in \mathfrak{l}$ satisfying the relations (41).

Remark 30.2 Let $\mathfrak{l}$ be a complex Lie algebra. Then the complex linear span of an $\mathfrak{sl}(2)$-triple in $\mathfrak{l}$ is a Lie subalgebra isomorphic to $\mathfrak{sl}(2,\mathbb{C})$.

Substituting H, X and Y for ξ in (40), we obtain, for $p \in P_n(\mathbb{C}^2)$,

$$H\ p = [-z_1 \frac{\partial}{\partial z_1} + z_2 \frac{\partial}{\partial z_2}]p, \quad X\ p = -z_2 \frac{\partial}{\partial z_1} p, \quad Y\ p = -z_1 \frac{\partial}{\partial z_2} p. \tag{42}$$

By a straightforward computation we now see that the action of the triple H, X, Y on the basis element p_j is given by

$$Hp_j = (n-2j)p_j, \quad Xp_j = -jp_{j-1}, \quad Yp_j = (j-n)p_{j+1}.$$

For the matrices of the action of H, X, Y on $P_n(\mathbb{C}^2)$ relative to the basis $p_0, \ldots, p_n$ we thus find

$$\mathrm{mat}(H) = \begin{pmatrix} n & 0 & \ldots & 0 \\ 0 & n-2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & \ldots & -n \end{pmatrix},$$

and

$$\mathrm{mat}(X) = \begin{pmatrix} 0 & -1 & 0 & \ldots & 0 \\ 0 & 0 & -2 & \ldots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ & & & 0 & -n \\ 0 & \ldots & & \ldots & 0 \end{pmatrix} \qquad \mathrm{mat}(Y) = \begin{pmatrix} 0 & \ldots & & \ldots & 0 \\ -n & 0 & & & \vdots \\ 0 & 1-n & 0 & & \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & -1 & 0 \end{pmatrix}.$$

These matrices will guide us through the proof of the following theorem.

Theorem 30.3 *Every irreducible finite dimensional* $\mathfrak{sl}(2, \mathbb{C})$*-module is isomorphic to* $P_n(\mathbb{C}^2)$*, for some* $n \in \mathbb{N}$.

Remark 30.4 From the above theorem we deduce again, using Lemmas 29.1 and 29.3, that every irreducible continuous finite dimensional representation of SU(2) is equivalent to π_n, for some $n \in \mathbb{N}$.

The proof of the the above theorem will be given in the rest of this section. Let V be an irreducible finite dimensional $\mathfrak{sl}(2, \mathbb{C})$-module.

Given $\lambda \in \mathbb{C}$, we shall write $V_\lambda := \ker(H - \lambda I)$. This space is non-trivial if and only if λ is an eigenvalue for the action of H on V.

Lemma 30.5 *Let* $\lambda \in \mathbb{C}$*. Then*

$$XV_\lambda \subset V_{\lambda+2}, \quad YV_\lambda \subset V_{\lambda-2}.$$

Proof: Let $v \in V_\lambda$. Then $HXv = XHv + [H, X]v = \lambda Xv + 2Xv = (\lambda + 2)Xv$, hence $Xv \in V_{\lambda+2}$. This proves the first inclusion. The second inclusion is proved in a similar manner. □

By a *primitive vector* of V we mean a vector $v \in V \setminus \{0\}$ with the property that $Xv = 0$. The idea behind this definition is to get hold of the analogue of $p_0 \in P_n(\mathbb{C}^2)$.

Lemma 30.6 V *contains a primitive vector that is an eigenvector for* H.

Proof: Let λ be an eigenvalue of the action of H on V. Fix an eigenvector $w \in V_\lambda$, $w \neq 0$ and consider the sequence of vectors $w_k, k \geq 0$, defined by $w_0 = w$ and $w_{k+1} = Xw_k$. Then $w_k \in V_{\lambda+2k}$. If all vectors w_k were non-zero, then they would be eigenvectors for different eigenvalues of H, hence they would be linear independent, contradicting the finite dimensionality of V. It follows that there exists a largest k such that $w_k \neq 0$. The vector w_k is primitive. □

In the following we assume that $v \in V$ is a fixed primitive vector that is an eigenvector for H. The associated eigenvalue is denoted by λ. We now consider the vectors v_k defined by $v_0 = v$ and $v_{k+1} = Yv_k$. By a similar reasoning as in the above proof it follows that there exists a largest number n such that $v_n \neq 0$.

Lemma 30.7

(a) *The vectors* $v_k = Y^k v$, $0 \leq k \leq n$*, form a basis for* V.

(b) *The eigenvalue* λ *equals* $n = \dim V - 1$.

(c) *For every* $0 \leq k \leq n$,

$$Hv_k = (\lambda - 2k)v_k, \qquad Xv_k = k(\lambda - k + 1)v_{k-1}.$$

(d) *The primitive vectors in V are the non-zero multiples of v_0.*

Proof: We first prove (c) for all $k \in \mathbb{N}$ (but note that $v_k = 0$ for $k > n$). It follows from repeated application of Lemma 30.5 that $v_k \in V_{\lambda-2k}$, hence $Hv_k = (\lambda - 2k)v_k$. We prove the second assertion of (c) by induction. Since $v_0 = v$ is primitive, the second assertion of (c) holds for $k = 0$. Let now $k > 0$ and assume that the assertion has been established for strictly smaller values of k. Then

$$\begin{aligned}
Xv_k &= XYv_{k-1} \\
&= YXv_{k-1} + [X, Y]v_{k-1} \\
&= YXv_{k-1} + Hv_{k-1} \\
&= (k-1)(\lambda - (k-2))Yv_{k-2} + (\lambda - 2(k-1))v_{k-1} \\
&= k(\lambda - k + 1)v_{k-1}
\end{aligned}$$

and (c) follows.

Let W be the linear span of the vectors v_k, for $0 \leq k \leq n$. Then by definition of the vectors v_k, $Yv_k = v_{k+1}$. Therefore, Y leaves W invariant. By (c), H and X leave W invariant as well. It follows that W is a non-trivial invariant subspace of V, hence $V = W$ by irreducibility. The vectors v_k, for $0 \leq k \leq n$, must be linear independent since they are eigenvectors for H for distinct eigenvalues; hence (a).

Finally, we have established the second assertion of (c) for all $k \geq 0$, in particular for $k = n + 1$. Now $v_{n+1} = 0$, hence $0 = (n+1)(\lambda - n)v_n$ and since $v_n \neq 0$ it follows that $\lambda = n$. This establishes (b).

It follows from (a) and (c) that the only primitive vectors in V are non-zero multiples of v_0.
□

Corollary 30.8 *Let V and V' be two irreducible finite dimensional $\mathfrak{sl}(2, \mathbb{C})$-modules. Then $V \simeq V'$ if and only if* $\dim V = \dim V'$. *Moreover, if v and v' are primitive vectors of V and V', respectively, then there is a unique isomorphism $T : V \to V'$ mapping v onto v'.*

Proof: Clearly if $V \simeq V'$ then V and V' have equal dimension. Conversely, assume that $\dim V = \dim V' = n$ and that v and v' are primitive vectors of V and V' respectively. Then by the above lemma, the vectors $v_k = Y^k v$, $0 \leq k \leq n$ form a basis of V. Similarly the vectors $v'_k = Y^k v'$, $0 \leq k \leq n$ form a basis of V'. Any intertwining operator $T : V \to V'$ that maps v onto v' must map the basis v_k onto the basis v'_k, hence is uniquely determined. Let $T : V \to V'$ be the linear map determined by $Tv_k = v'_k$, for $0 \leq k \leq n$. Then T is a linear bijection. Moreover, by the above lemma we see that T intertwines the actions of H, X, Y on V and V'. It follows that T is equivariant, hence $V \simeq V'$. □

Completion of the proof of Theorem 30.3: The space $P_n(\mathbb{C}^2)$ is an irreducible $\mathfrak{sl}(2, \mathbb{C})$-module, of dimension $n + 1$. Hence if V is an irreducible $\mathfrak{sl}(2, \mathbb{C})$-module of dimension $m \geq 1$, then $V \simeq P_n(\mathbb{C}^2)$, with $n = m - 1$. □

31 Roots and weights

Let $\mathfrak{t}$ be a finite dimensional commutative real Lie algebra, and let (ρ, V) be a finite dimensional representation of $\mathfrak{t}$ in V.

Let $\mathfrak{t}_{\mathbb{C}}^*$ denote the space of complex linear functionals on $\mathfrak{t}_{\mathbb{C}}$. Note that $\mathfrak{t}^*$, the space of real linear functionals on $\mathfrak{t}$ may be identified with the space of $\lambda \in \mathfrak{t}_{\mathbb{C}}^*$ that are real valued on $\mathfrak{t}$. Thus, $\mathfrak{t}^*$ is viewed as a real linear subspace of $\mathfrak{t}_{\mathbb{C}}^*$. Accordingly $i\,\mathfrak{t}^*$ equals the space of $\lambda \in \mathfrak{t}_{\mathbb{C}}^*$ such that $\lambda|_{\mathfrak{t}}$ has values in $i\mathbb{R}$.

If $\lambda \in \mathfrak{t}_{\mathbb{C}}^*$, then we define the following subspace of V :

$$V_\lambda = \bigcap_{H \in \mathfrak{t}} \ker(\rho(H) - \lambda(H)I). \tag{43}$$

In other words, V_λ consists of the space of $v \in V$ such that $\rho(H)v = \lambda(H)v$ for all $H \in \mathfrak{t}$. If $V_\lambda \neq 0$, then λ is called a *weight* of $\mathfrak{t}$ in V, and V_λ is called the associated *weight space.* The set of weights of $\mathfrak{t}$ in V is denoted by $\Lambda(\rho)$.

Lemma 31.1 *Let $T \in \mathrm{End}(V)$ be a ρ-intertwining linear endomorphism, then T leaves V_λ invariant, for every $\lambda \in \Lambda(\rho)$.*

Proof: Let $\lambda \in \Lambda(\rho)$. The endomorphism T commutes with $\rho(H)$ hence leaves the eigenspace $\ker(\rho(H) - \lambda(H))$ invariant, for every $H \in \mathfrak{t}$. Hence T leaves the intersection V_λ of all these spaces invariant. $\square$

Lemma 31.2 *The set $\Lambda(\rho)$ is a non-empty finite subset of $\mathfrak{t}_{\mathbb{C}}^*$. Assume that $\rho(X)$ is diagonalizable for every $X \in \mathfrak{t}$. Then*

$$V = \bigoplus_{\lambda \in \Lambda(\rho)} V_\lambda. \tag{44}$$

Moreover, if W is a $\mathfrak{t}$-invariant subspace of V, then W is the direct sum of the spaces $W \cap V_\lambda$, for $\lambda \in \Lambda(\rho)$.

Proof: Fix a basis $X_1, \ldots, X_n$ of $\mathfrak{t}$. The endomorphism $\rho(X_1)$ has at least one eigenvalue, say λ_1, with corresponding eigenspace $E_1 \subset V$. Since $\mathfrak{t}$ is commutative, this eigenspace is invariant under the action of $\mathfrak{t}$. Proceeding by induction on $\dim \mathfrak{t}$, we obtain a sequence of non-trivial subspaces $E_n \subset E_{n-1} \subset \cdots \subset E_1$ such that X_j acts by a scalar λ_j on E_j, for each $1 \leq j \leq n$. Define $\lambda \in \mathfrak{t}_{\mathbb{C}}^*$ by $\lambda(X_j) = \lambda_j$, then $E_n \subset V_\lambda$, hence $\lambda \in \Lambda(\rho)$. This establishes the first assertion.

If $\rho(X)$ diagonalizes, for every $X \in \mathfrak{t}$, then, in particular, V admits a decomposition of eigenspaces for the endomorphism $\rho(X_1)$. Each of these eigenspaces is invariant under $\mathfrak{t}$. Therefore, by induction on $\dim \mathfrak{t}$ there exists a direct sum decomposition $V = V_1 \oplus \cdots \oplus V_N$ such that X_j acts by a scalar λ_{ij} on V_i, for all $1 \leq i \leq N$ and $1 \leq j \leq n$. Let $\lambda_i \in \mathfrak{t}_{\mathbb{C}}^*$ be defined by $\lambda_i(X_j) = \lambda_{ij}$, for $1 \leq i \leq N$. Then $\Lambda(\rho) = \{\lambda_1, \ldots, \lambda_N\}$. Moreover, one readily verifies that, for $\lambda \in \lambda(\rho)$, $V_\lambda = \oplus_{j : \lambda = \lambda_j} V_j$. Hence, (44) follows.

For the final assertion, we observe that by finite dimensionality of V the set $\Lambda(\rho)$ is finite. Hence, there exists a $X_0 \in \mathfrak{t}$ such that $(\nu - \mu)(X_0) \neq 0$ for all $\nu, \mu \in \Lambda(\rho)$ with $\nu \neq \mu$. For $\nu \in \Lambda(\rho)$, let $P_\nu : V \to V_\nu$ be the projection along the remaining summands in (44). We claim that

$$P_\nu = \prod_{\mu \in \Lambda(\rho)\setminus\{\nu\}} (\nu(X_0) - \mu(X_0))^{-1}(\rho(X_0) - \mu(X_0)).$$

Indeed this is readily checked on each of the summands V_λ of the decomposition in (44), for $\lambda \in \Lambda(\rho)$.

It follows from the above formula for P_ν that $P_\nu(W) \subset W$. Hence, $P_\nu(W) \subset W \cap V_\nu$, and the final assertion follows. $\square$

Assumption: In the rest of this section we assume that G is a compact Lie group, with Lie algebra $\mathfrak{g}$.

Definition 31.3 A *torus* in $\mathfrak{g}$ is by definition a commutative subalgebra of $\mathfrak{g}$. A torus $\mathfrak{t} \subset \mathfrak{g}$ is called *maximal* if there exists no torus of $\mathfrak{g}$ that properly contains $\mathfrak{t}$.

From now on we assume that $\mathfrak{t}$ is a fixed maximal torus in $\mathfrak{g}$.

Lemma 31.4 *The centralizer of* $\mathfrak{t}$ *in* $\mathfrak{g}$ *equals* $\mathfrak{t}$.

Proof: Since $\mathfrak{t}$ is abelian, it is contained in its centralizer. Conversely, assume that $X \in \mathfrak{g}$ centralizes $\mathfrak{t}$. Then $\mathfrak{t}' = \mathfrak{t} + \mathbb{R}X$ is a torus which contains $\mathfrak{t}$. Hence $\mathfrak{t}' = \mathfrak{t}$ by maximality, and we see that $X \in \mathfrak{t}$. $\square$

Let (π, V) be a finite dimensional representation of $\mathfrak{g}_\mathbb{C}$, the complexification of the Lie algebra $\mathfrak{g}$; i.e., π is a complex Lie algebra homomorphism from $\mathfrak{g}_\mathbb{C}$ into $\mathrm{End}(V)$ (the latter is the space of complex linear endomorphisms equipped with the commutator Lie bracket). Alternatively we will also say that V is a finite dimensional $\mathfrak{g}_\mathbb{C}$-module. We denote by $\Lambda(\pi_*) = \Lambda(\pi_*, \mathfrak{t})$ the set of weights of the representation $\rho = \pi|_\mathfrak{t}$ of $\mathfrak{t}$ in V. If $\lambda \in \mathfrak{t}_\mathbb{C}^*$, then as before, V_λ is defined as in (43), with $\pi|_\mathfrak{t}$ in place of ρ. Thus

$$V_\lambda = \{v \in V \mid \pi(H)v = \lambda(H)v \quad \text{for all} \quad H \in \mathfrak{t}\}.$$

From Lemma 31.2 we see that $\Lambda(\pi_*)$ is a non-empty finite subset of $\mathfrak{t}_\mathbb{C}^*$.

Let (π, V) be a finite dimensional continuous representation of G. Then the map $\pi : G \to \mathrm{GL}(V)$ is a homomorphism of Lie groups. Let $\pi_* = T_e\pi$. Then $\pi_* : \mathfrak{g} \to \mathrm{End}(V)$ is a Lie algebra homomorphism, or, differently said, a representation of $\mathfrak{g}$ in V. The homomorphism π_* has a unique extension to a complex Lie algebra homomorphism from $\mathfrak{g}_\mathbb{C}$ into $\mathrm{End}(V)$ (we recall that V is a complex linear space by assumption). This extension is called the *induced infinitesimal representation* of $\mathfrak{g}_\mathbb{C}$ in V.

Lemma 31.5 *Let π be a finite dimensional continuous representation of G. Then $\Lambda(\pi_*)$ is a finite subset of $i\,\mathfrak{t}^*$. Moreover,*

$$V = \bigoplus_{\lambda \in \Lambda(\pi_*)} V_\lambda.$$

If V is equipped with a G-invariant inner product, then for all $\lambda, \mu \in \Lambda(\pi_*)$ with $\lambda \neq \mu$ we have $V_\lambda \perp V_\mu$.

Proof: There exists a G-invariant inner product on V; assume such an inner product $\langle \cdot\, , \, \cdot \rangle$ to be fixed. Then π maps G into $\mathrm{U}(V)$, the associated group of unitary transformations. It follows that π_* maps $\mathfrak{g}$ into the Lie algebra $\mathfrak{u}(V)$ of $\mathrm{U}(V)$, which is the subalgebra of anti-Hermitian endomorphisms in $\mathrm{End}(V)$. It follows that for $X \in \mathfrak{g}$ the endomorphism $\pi_*(X)$ is anti-Hermitian, hence diagonalizable with imaginary eigenvalues. The proof is now completed by application of Lemma 31.2. $\square$

If $A \in \mathrm{End}(\mathfrak{g})$, then we denote by $A_\mathbb{C}$ the complex linear extension of A to $\mathfrak{g}_\mathbb{C}$. Obviously the map $A \mapsto A_\mathbb{C}$ induces a real linear embedding of $\mathrm{End}(\mathfrak{g})$ into $\mathrm{End}(\mathfrak{g}_\mathbb{C}) := \mathrm{End}_\mathbb{C}(\mathfrak{g}_\mathbb{C})$. Accordingly we shall view $\mathrm{End}(\mathfrak{g})$ as a real linear subspace of the complex linear space $\mathrm{End}(\mathfrak{g}_\mathbb{C})$ from now on. Thus, we may view Ad as a representation of G in the complexification $\mathfrak{g}_\mathbb{C}$ of $\mathfrak{g}$. The associated infinitesimal representation is the adjoint representation ad of $\mathfrak{g}_\mathbb{C}$ in $\mathfrak{g}_\mathbb{C}$. The associated collection $\Lambda(\mathrm{ad})$ of weights contains the weight 0. Indeed the associated weight space $\mathfrak{g}_{\mathbb{C}0}$ equals the centralizer of $\mathfrak{t}$ in $\mathfrak{g}_\mathbb{C}$, which in turn equals $\mathfrak{t}_\mathbb{C}$, by Lemma 31.4. Hence:

$$\mathfrak{g}_{\mathbb{C}0} = \mathfrak{t}_\mathbb{C}.$$

Definition 31.6 The weights of ad in $\mathfrak{g}_\mathbb{C}$ different from 0 are called the *roots* of $\mathfrak{t}$ in $\mathfrak{g}_\mathbb{C}$; the set of these is denoted by $R = R(\mathfrak{g}_\mathbb{C}, \mathfrak{t})$. Given $\alpha \in R$, the associated weight space $\mathfrak{g}_{\mathbb{C}\alpha}$ is called a *root space.*

It follows from the definitions that

$$\mathfrak{g}_{\mathbb{C}\alpha} = \{X \in \mathfrak{g}_\mathbb{C} \mid [H, X] = \alpha(H)X \quad \text{for all} \quad H \in \mathfrak{t}\}.$$

From Lemma 31.5 we now obtain the so called *root space decomposition* of $\mathfrak{g}_\mathbb{C}$, relative to the torus $\mathfrak{t}$.

Corollary 31.7 *The collection $R = R(\mathfrak{g}_\mathbb{C}, \mathfrak{t})$ of roots is a finite subset of $i\,\mathfrak{t}^*$. Moreover, we have the following direct sum of vector spaces:*

$$\mathfrak{g}_\mathbb{C} = \mathfrak{t}_\mathbb{C} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_{\mathbb{C}\alpha}. \tag{45}$$

Example 31.8 The Lie algebra $\mathfrak{g} = \mathfrak{su}(2)$ has complexification $\mathfrak{sl}(2, \mathbb{C})$, consisting of all complex 2×2 matrices with trace zero. Let H, X, Y be the standard basis of $\mathfrak{sl}(2, \mathbb{C})$; i.e.

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Now $\mathfrak{t} = i\mathbb{R}H$ is a maximal torus in $\mathfrak{su}(2)$. We recall that $[H, X] = 2X$, $[H, Y] = -2Y$, $[X, Y] = H$. Thus, if we define $\alpha \in \mathfrak{t}_\mathbb{C}^*$ by $\alpha(H) = 2$, then $R = R(\mathfrak{g}_\mathbb{C}, \mathfrak{t})$ equals $\{\alpha, -\alpha\}$. Moreover, $\mathfrak{g}_{\mathbb{C}\alpha} = \mathbb{C}X$ and $\mathfrak{g}_{\mathbb{C}(-\alpha)} = \mathbb{C}Y$.

We recall that, by definition, the center $\mathfrak{z} = \mathfrak{z}_\mathfrak{g}$ of $\mathfrak{g}$ is the ideal $\ker \mathrm{ad}$; i.e., it is the space of $X \in \mathfrak{g}$ that commute with all $Y \in \mathfrak{g}$.

Lemma 31.9 *The center of $\mathfrak{g}$ is contained in $\mathfrak{t}$ and equals the intersection of the root hyperplanes:*

$$\mathfrak{z}_\mathfrak{g} = \bigcap_{\alpha \in R} \ker \alpha.$$

In particular, if $\mathfrak{z}_\mathfrak{g} = 0$, then R spans the real linear space $i\mathfrak{t}^*$.

Proof: The center of $\mathfrak{g}$ centralizes $\mathfrak{t}$ in particular, hence is contained in $\mathfrak{t}$, by Lemma 31.4. Let $H \in \mathfrak{t}$ and assume that H centralizes $\mathfrak{g}$; then H centralizes $\mathfrak{g}_\mathbb{C}$, hence every root space of $\mathfrak{g}_\mathbb{C}$. This implies that $\alpha(H) = 0$ for all $\alpha \in R$. Conversely, if $H \in \mathfrak{t}$ is in the intersection of all the root hyperplanes, then H centralizes $\mathfrak{t}_\mathbb{C}$ and every root space $\mathfrak{g}_{\mathbb{C}\alpha}$. By the root space decomposition it then follows that $H \in \mathfrak{z}$. This establishes the characterization of the center.

If $\mathfrak{z} = 0$, then the root hyperplanes $\ker \alpha$ ($\alpha \in R$) have a zero intersection in $\mathfrak{t}$. This implies that the set $R \subset i\mathfrak{t}^*$ spans the real linear space $i\mathfrak{t}^*$. $\square$

Lemma 31.10 *Let (π, V) be a finite dimensional representation of $\mathfrak{g}_\mathbb{C}$. Then for all $\lambda \in \Lambda(\pi)$ and all $\alpha \in R \cup \{0\}$ we have:*

$$\pi(\mathfrak{g}_{\mathbb{C}\alpha})V_\lambda \subset V_{\lambda+\alpha}.$$

In particular, if $\lambda + \alpha \notin \Lambda(\pi)$, then $\pi(\mathfrak{g}_{\mathbb{C}\alpha})$ anihilates V_λ.

Proof: Let $X \in \mathfrak{g}_{\mathbb{C}\alpha}$ and $v \in V_\lambda$. Then, for $H \in \mathfrak{t}$,

$$\begin{aligned} \pi(H)\pi(X)v &= \pi(X)\pi(H)v + [\pi(H), \pi(X)]v \\ &= \lambda(H)\pi(X)v + \pi([H, X])v = [\lambda(H) + \alpha(H)]\pi(X)v. \end{aligned}$$

Hence $\pi(X)v \in V_{\lambda+\alpha}$. If $\lambda + \alpha$ is not a weight of π, then $V_{\lambda+\alpha} = 0$ and it follows that $\pi(X)v = 0$. $\square$

Corollary 31.11 *If $\alpha, \beta \in R \cup \{0\}$, then*

$$[\mathfrak{g}_{\mathbb{C}\alpha}, \mathfrak{g}_{\mathbb{C}\beta}] \subset \mathfrak{g}_{\mathbb{C}(\alpha+\beta)}.$$

In particular, if $\alpha + \beta \notin R \cup \{0\}$, then $\mathfrak{g}_{\mathbb{C}\alpha}$ and $\mathfrak{g}_{\mathbb{C}\beta}$ commute.

Proof: This follows from the previous lemma applied to the adjoint representation. $\square$

We shall write $\mathbb{Z}R$ for the $\mathbb{Z}$-linear span of R, i.e., the $\mathbb{Z}$-module of elements of the form $\sum_{\alpha \in R} n_\alpha \alpha$, with $n_\alpha \in \mathbb{Z}$.

In the following corollary we do not assume that π comes from a representation of G.

Corollary 31.12 *Let (π, V) be a finite dimensional representation of $\mathfrak{g}_{\mathbb{C}}$. Then*

$$W := \bigoplus_{\lambda \in \Lambda(\pi)} V_\lambda \tag{46}$$

is a non-trivial $\mathfrak{g}_{\mathbb{C}}$-submodule. If π is irreducible, then $W = V$. Moreover, if $\lambda, \mu \in \Lambda(\pi)$, then $\lambda - \mu \in \mathbb{Z}R$.

Proof: By Lemma 31.2 the set $\Lambda(\pi)$ is non-empty and finite, and therefore W is a non-trivial subspace of V. From Lemma 31.10 we see that W is $\mathfrak{g}_{\mathbb{C}}$-invariant. If π is irreducible, then $W = V$. To establish the last assertion we define an equivalence relation on $\Lambda(\pi)$ by $\lambda \sim \mu \iff \lambda - \mu \in \mathbb{Z}R$. If S is a class for $\sim$, then $V_S = \oplus_{\lambda \in S} V_\lambda$ is a non-trivial $\mathfrak{g}_{\mathbb{C}}$-invariant subspace of V, by Lemma 31.10. Hence $V_S = V$ and it follows that $S = \Lambda(\pi)$. $\square$

Remark 31.13 If $\mathfrak{g}$ has trivial center, then the above result actually holds for every finite dimensional V-module. To see that a condition like this is necessary, consider $\mathfrak{g} = \mathbb{R}$, the Lie algebra of the circle. Define a representation of $\mathfrak{g}$ in $V = \mathbb{C}^2$ by

$$\pi(x) = \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix}.$$

Then $\Lambda(\pi) = \{0\}$, but $V_0 = \mathbb{C} \times \{0\}$ is not all of V.

Note that this does not contradict the conclusion of Lemma 31.5, since π is not associated with a continuous representation of the circle group in $\mathbb{C}^2$.

Lemma 31.14 *Let $\mathfrak{t}$ be a maximal torus in $\mathfrak{g}$, and R the associated collection of roots. If $\alpha \in R$ then $-\alpha \in R$.*

Proof: Let τ be the conjugation of $\mathfrak{g}_{\mathbb{C}}$ with respect to the real form $\mathfrak{g}$. That is: $\tau(X + iY) = X - iY$ for all $X, Y \in \mathfrak{g}$. One readily checks that τ is an automorphism of $\mathfrak{g}_{\mathbb{C}}$, considered as a real Lie algebra (by forgetting the complex linear structure). Let $\alpha \in R$, and let $X \in \mathfrak{g}_{\mathbb{C}\alpha}$. Then for every $H \in \mathfrak{t}$,

$$[H, \tau(X)] = \tau[H, X] = \tau(\alpha(H)X) = \overline{\alpha(H)}\tau(X) = -\alpha(H)\tau(X).$$

For the latter equation we used that α has imaginary values on $\mathfrak{t}$. It follows that $-\alpha \in R$ and that τ maps $\mathfrak{g}_{\mathbb{C}\alpha}$ into $\mathfrak{g}_{\mathbb{C}-\alpha}$ (in fact is a bijection between these root spaces; why?). $\square$

We recall that we identify $i\mathfrak{t}^*$ with the real linear subspace of $\mathfrak{t}_{\mathbb{C}}^*$ consisting of λ such that $\lambda|\mathfrak{t}$ has values in $i\mathbb{R}$; the latter condition is equivalent to saying that $\lambda|_{i\mathfrak{t}}$ is real valued. One readily verifies that the restriction map $\lambda \mapsto \lambda|_{i\mathfrak{t}}$ defines a real linear isomorphism from $i\mathfrak{t}^*$ onto the real linear dual $(i\mathfrak{t})^*$. In the following we shall use this isomorphism to identify $i\mathfrak{t}^*$ with $(i\mathfrak{t})^*$. Now R is a finite subset of $(i\mathfrak{t})^* \setminus \{0\}$. Hence the complement of the hyperplanes $\ker \alpha \subset i\mathfrak{t}$, for $\alpha \in R$ is a finite union of connected components, which are all convex. These components are called the *Weyl chambers* associated with R. Let $\mathcal{C}$ be a fixed chamber. By definition every root is either positive or negative on $\mathcal{C}$. We define the *system of positive roots* $R^+ := R^+(\mathcal{C})$ associated with $\mathcal{C}$ by

$$R^+ = \{\alpha \in R \mid \alpha > 0 \quad \text{on} \quad \mathcal{C}\}.$$

By what we said above, for every $\alpha \in R$, we have that either α or $-\alpha$ belongs to R^+, but not both. It follows that

$$R = R^+ \cup (-R^+) \quad \text{(disjoint union).} \tag{47}$$

We write $\mathbb{N}R^+$ for the subset of $\mathbb{Z}R$ consisting of the elements that can be written as a sum of the form $\sum_{\alpha\in R^+} n_\alpha \alpha$, with $n_\alpha \in \mathbb{N}$.

Lemma 31.15 $\mathbb{N}R^+ \cap (-\mathbb{N}R^+) = 0$.

Proof: Let $\mu \in \mathbb{N}R^+$. Then $\mu \geq 0$ on $\mathcal{C}$, the chamber corresponding to R^+. If also $-\mu \in \mathbb{N}R^+$, then $\mu \leq 0$ on $\mathcal{C}$ as well. Hence $\mu = 0$ on $\mathcal{C}$. Since $\mathcal{C}$ is a non-empty open subset of $i\mathfrak{t}^*$, this implies that $\mu = 0$. $\square$

Lemma 31.16 *The spaces*

$$\mathfrak{g}_{\mathbb{C}}^+ := \sum_{\alpha\in R^+} \mathfrak{g}_{\mathbb{C}\alpha}, \qquad \mathfrak{g}_{\mathbb{C}}^- := \sum_{\beta\in -R^+} \mathfrak{g}_{\mathbb{C}\beta}$$

are $\mathrm{ad}(\mathfrak{t})$*-stable subalgebras of* $\mathfrak{g}_{\mathbb{C}}$*. Moreover,*

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{g}_{\mathbb{C}}^+ \oplus \mathfrak{t}_{\mathbb{C}} \oplus \mathfrak{g}_{\mathbb{C}}^-.$$

Proof: Let $\alpha, \beta \in R^+$ and assume that $[\mathfrak{g}_{\mathbb{C}\alpha}, \mathfrak{g}_{\mathbb{C}\beta}] \neq 0$. Then $\alpha + \beta \in R \cup \{0\}$, and $\alpha + \beta > 0$ on $\mathcal{C}$. This implies that $\alpha + \beta \in R^+$, hence $\mathfrak{g}_{\mathbb{C}(\alpha+\beta)} \subset \mathfrak{g}_{\mathbb{C}}^+$. It follows that $\mathfrak{g}_{\mathbb{C}}^+$ is a subalgebra. For similar reasons $\mathfrak{g}_{\mathbb{C}}^-$ is a subalgebra. Both subalgebras are $\mathrm{ad}(\mathfrak{t})$ stable, since root spaces are. The direct sum decomposition is an immediate consequence of (45) and (47). $\square$

We are now able to define the notion of a highest weight vector for a finite dimensional $\mathfrak{g}_{\mathbb{C}}$-module, relative to the system of positive roots R^+. This is the appropriate generalization of the notion of a primitive vector for $\mathfrak{sl}(2, \mathbb{C})$.

Definition 31.17 Let V be a finite dimensional $\mathfrak{g}_{\mathbb{C}}$-module. Then a *highest weight vector* of V is by definition a non-trivial vector $v \in V$ such that

(a) $\mathfrak{t}_{\mathbb{C}} v \subset \mathbb{C}v$;

(b) $Xv = 0$ for all $X \in \mathfrak{g}_{\mathbb{C}}^+$.

Lemma 31.18 *Let V be a finite dimensional $\mathfrak{g}_{\mathbb{C}}$-module. Then V has a highest weight vector.*

Proof: We define the $\mathfrak{g}_{\mathbb{C}}$-submodule W of V as the sum of the $\mathfrak{t}_{\mathbb{C}}$-weight spaces, see Corollary 31.12.

Let $\mathcal{C}$ be the positive chamber determining R^+. Fix $X \in \mathcal{C}$. Then $\alpha(X) > 0$ for all $\alpha \in R^+$. We may select $\lambda_0 \in \Lambda(\pi)$ such that the real part of $\lambda(X)$ is maximal. Then $\lambda_0 + \alpha \notin \Lambda(\pi)$ for all $\alpha \in R^+$. By Lemma 31.10 this implies that $\pi_*(\mathfrak{g}_{\mathbb{C}\alpha})V_\lambda \subset V_{\lambda_0+\alpha} = 0$ for all $\alpha \in R^+$. Hence $\mathfrak{g}_{\mathbb{C}}^+$ annihilates V_{λ_0}. Thus, every non-zero vector of V_{λ_0} is a highest weight vector. $\square$

Definition 31.19 Let V be a finite dimensional $\mathfrak{g}_\mathbb{C}$-module. A vector $v \in V$ is called *cyclic* if it generates the $\mathfrak{g}_\mathbb{C}$-module V, i.e., V is the smallest $\mathfrak{g}_\mathbb{C}$-submodule containing v.

Obviously, if V is irreducible, then every non-trivial vector is cyclic.

Proposition 31.20 *Let V be a finite dimensional $\mathfrak{g}_\mathbb{C}$-module and let $v \in V$ be a cyclic highest weight vector.*

(a) *There exists a (unique) $\lambda \in \Lambda(V)$ such that $v \in V_\lambda$. Moreover, $V_\lambda = \mathbb{C}v$.*

(b) *The space V is equal to the span of the vectors v and $\pi(X_1)\cdots\pi(X_n)v$, with $n \in \mathbb{N}$ and $X_j \in \mathfrak{g}_\mathbb{C}^-$, for $1 \leq j \leq n$.*

(c) *Every weight $\mu \in \Lambda(V)$ is of the form $\lambda - \nu$, with $\nu \in \mathbb{N}R^+$.*

(d) *The module V has a unique maximal proper submodule W.*

(e) *The module V has a unique non-trivial irreducible quotient.*

Proof: The first assertion of (a) follows from the definition of highest weight vector. We define an increasing sequence of linear subspaces of V inductively by $V_0 = \mathbb{C}v$ and $V_{n+1} = V_n + \pi(\mathfrak{g}_\mathbb{C}^-)V_n$. Let W be the union of the spaces V_n. We claim that W is an invariant subspace of V. To establish the claim, we note that by definition we have $\pi(\mathfrak{g}_\mathbb{C}^-)V_n \subset V_{n+1}$; hence W is $\mathfrak{g}_\mathbb{C}^-$ invariant. The space V_0 is $\mathfrak{t}$- and $\mathfrak{g}_\mathbb{C}^+$-invariant; by induction we will show that the same holds for V_n. Assume that V_n is $\mathfrak{t}$- and $\mathfrak{g}_\mathbb{C}^+$-invariant, and let $v \in V_n,\ Y \in \mathfrak{g}_\mathbb{C}^-$. Then for H in $\mathfrak{t}$ we have $HYv = YHv + [H, Y]v$. Now $v \in V_n$ and by the inductive hypothesis it follows that $Hv \in V_n$. Hence $YHv \in V_{n+1}$. Also $[H, Y] \in \mathfrak{g}_\mathbb{C}^-$ and it follows that $[H, Y]v \in V_{n+1}$. We conclude that $HYv \in V_{n+1}$. It follows from this that

$$\pi(\mathfrak{t})\pi(\mathfrak{g}_\mathbb{C}^-)V_n \subset V_{n+1}.$$

Hence V_{n+1} is $\mathfrak{t}$-invariant.

Let now $v \in V_n$, $Y \in \mathfrak{g}_\mathbb{C}^-$ and $X \in \mathfrak{g}_\mathbb{C}^+$. Then $XYv = YXv + [X, Y]v$. Now $Xv \in V_n$ by the induction hypothesis and we see that $YXv \in V_{n+1}$. Also, $[X, Y] \in \mathfrak{g}_\mathbb{C}$. By the induction hypothesis it follows that $\mathfrak{g}_\mathbb{C}V_n \subset V_{n+1}$. Hence $[X, Y]v \in V_{n+1}$. We conclude that $XYv \in V_{n+1}$. It follows from this that

$$\pi(\mathfrak{g}_\mathbb{C}^+)\pi(\mathfrak{g}_\mathbb{C}^-)V_n \subset V_{n+1}.$$

Hence V_{n+1} is $\mathfrak{g}_\mathbb{C}^+$-invariant. This establishes the claim that W is a $\mathfrak{g}_\mathbb{C}$-invariant subspace of V.

Since W contains the cyclic vector v, it follows that $W = V$. Hence, (b) follows. Let $w = \pi(Y_1)\cdots\pi(Y_n)v$, with $n \in \mathbb{N}$, $Y_j \in \mathfrak{g}_{\mathbb{C}(-\alpha_j)}$, $\alpha_j \in R^+$. Then w belongs to the weight space $V_{\lambda-\nu}$, where $\nu = \alpha_1 + \cdots + \alpha_n \in \mathbb{N}R^+$. Since v and such elements w span $W = V$, we conclude that every weight μ in $\Lambda(V)$ is of the form $\lambda - \nu$ with $\nu \in \mathbb{N}R^+$. This establishes (c). Moreover, it follows from the above description that V equals the vector sum of $\mathbb{C}v$ and V_-, where V_- denotes the sum of the weight spaces V_μ with $\mu \in \Lambda(V) \setminus \{\lambda\}$. This implies that $V_\lambda = \mathbb{C}v$, whence the second assertion of (a).

We now turn to assertion (d). Let U be a submodule of V. In particular, U is a $\mathfrak{t}_{\mathbb{C}}$-invariant subspace. Let $\Lambda(U)$ be the collection of $\mu \in \Lambda(V)$ for which $U_\mu := U \cap V_\mu \neq 0$. In view of Lemma 31.2, U is the direct sum of the spaces U_μ, for $\mu \in \Lambda(U)$. If U is a proper submodule, then $U_\lambda = 0$, hence $\Lambda(U) \subset \Lambda(V) \setminus \{\lambda\}$. It follows that the vector sum W of all proper submodules satisfies $\Lambda(W) \subset \Lambda(V) \setminus \{\lambda\}$ hence is still proper. Therefore, V has W as unique maximal submodule.

The final assertion (e) is equivalent to (d). To see this, let $p : V \to V'$ be a surjective $\mathfrak{g}_{\mathbb{C}}$-module homomorphism onto a non-trivial $\mathfrak{g}_{\mathbb{C}}$-module. Then $U \mapsto p^{-1}(U)$ defines a bijection from the collection of proper submodules of V' onto the collection of proper submodules of V containing $\ker p$. It follows that V' is irreducible if and only if $\ker p$ is a proper maximal submodule of V. The equivalence of (d) and (e) now readily follows. $\square$

Corollary 31.21 *Let V be a finite dimensional irreducible $\mathfrak{g}_{\mathbb{C}}$-module. Then V has a highest weight vector v, which is unique up to a scalar factor. Let λ be the weight of v. Then assertions (a) - (c) of Proposition 31.20 are valid.*

Proof: It follows from Lemma 31.18 that V has a highest weight vector. Let v be any highest weight vector in V and let λ be its weight. By irreducibility of V, the vector v is cyclic. Hence all assertions of Proposition 31.20 are valid.

Let w be a second highest weight vector and let μ be its weight. Then all assertions of Proposition 31.20 are valid. Hence $\mu \in \lambda - \mathbb{N}R^+$ and $\lambda \in \mu - \mathbb{N}R^+$, from which $\mu - \lambda \in \mathbb{N}R^+ \cap (-\mathbb{N}R^+) = \{0\}$. It follows that $\mu = \lambda$; hence $w \in V_\lambda = \mathbb{C}v$. $\square$

Remark 31.22 For obvious reasons the above weight λ is called the *highest weight* of the irreducible $\mathfrak{g}_{\mathbb{C}}$-module V, relative to the choice R^+ of positive roots.

The following theorem is the first step towards the classification of all finite dimensional irreducible representations of $\mathfrak{g}_{\mathbb{C}}$.

Theorem 31.23 *Let V and V' be irreducible $\mathfrak{g}_{\mathbb{C}}$-modules. If V and V' have the same highest weight (relative to R^+), then V and V' are isomorphic (i.e., the associated $\mathfrak{g}_{\mathbb{C}}$-representations are equivalent).*

Proof: We denote the highest weight by λ and fix associated highest weight vectors $v \in V_\lambda \setminus \{0\}$ and $v' \in V'_\lambda \setminus \{0\}$. We consider the direct sum $\mathfrak{g}_{\mathbb{C}}$-module $V \oplus V'$ and denote by W the smallest $\mathfrak{g}_{\mathbb{C}}$-submodule containing the vector $w := (v, v')$. Then w is a cyclic weight vector of W, of weight λ.

Let $p : V \oplus V' \to V$ be the projection onto the first component, and $p' : V \oplus V' \to V'$ the projection onto the second. Then p and p' are $\mathfrak{g}_{\mathbb{C}}$-module homomorphisms. Since $p(w) = v$, it follows that $p|_W$ is surjective onto V. Similarly, $p'|_W$ is surjective onto V'. It follows that V, V' are both irreducible quotients of W, hence isomorphic by Proposition 31.20 (e). $\square$

Remark 31.24 In the above proof it is easy to deduce that in fact W is irreducible, and $p|_W$ and $p'|_W$ are isomorphisms from W onto V and V', respectively.

32 Conjugacy of maximal tori

We retain the notation of the previous section. In this section we shall investigate to what extent the collection $R = R(\mathfrak{g}_{\mathbb{C}})$ depends on the choice of the maximal torus $\mathfrak{t}$. An element $X \in \mathfrak{t}_{\mathbb{C}}$ will be called *regular* if $\alpha(X) \neq 0$ for all $\alpha \in R$. The set of regular elements in $\mathfrak{t}$ and $\mathfrak{t}_{\mathbb{C}}$ will be denoted by $\mathfrak{t}^{\mathrm{reg}}$ and $\mathfrak{t}_{\mathbb{C}}^{\mathrm{reg}}$, respectively. Since R is finite, $\mathfrak{t}^{\mathrm{reg}}$ is an open dense subset of $\mathfrak{t}$; similarly $\mathfrak{t}_{\mathbb{C}}^{\mathrm{reg}}$ is an open dense subset of $\mathfrak{t}_{\mathbb{C}}$.

Lemma 32.1 *Let $\mathfrak{t}$ be a maximal torus in $\mathfrak{g}$, and let $X \in \mathfrak{t}$. Then the following statements are equivalent.*

(a) $X \in \mathfrak{t}^{\mathrm{reg}}$;

(b) $\ker(\mathrm{ad}(X)) = \mathfrak{t}$;

(c) *with respect to any G-invariant inner product on $\mathfrak{g}$ we have* $\mathfrak{t} = \mathrm{im}(\mathrm{ad}(X))^{\perp}$;

(d) *with respect to some G-invariant inner product on $\mathfrak{g}$ we have* $\mathfrak{t} = \mathrm{im}(\mathrm{ad}(X))^{\perp}$;

Proof: Assume (a), and let $Y \in \mathfrak{g}$ commute with X. In the complexification of $\mathfrak{g}$ we may decompose $Y = Y_0 + \sum_{\alpha \in R} Y_\alpha$, with $Y_0 \in \mathfrak{t}_{\mathbb{C}}$ and $Y_\alpha \in \mathfrak{g}_{\mathbb{C}\alpha}$ for $\alpha \in R$. Then

$$0 = [X, Y] = \sum_{\alpha \in R} \alpha(X) Y_\alpha .$$

Since X is regular, $\alpha(X) \neq 0$ for all α, and it follows that $Y_\alpha = 0$ for all $\alpha \in R$. Hence $Y \in \mathfrak{g} \cap \mathfrak{t}_{\mathbb{C}}^{*} = \mathfrak{t}$. This implies $\ker(\mathrm{ad}(X)) \subset \mathfrak{t}$; the converse inclusion is obvious, hence (b) follows.

Next, we assume that (b) holds. Since $\mathrm{ad}(X)$ is anti-symmetric with respect to any invariant inner product, it follows that $\mathrm{im}(\mathrm{ad}(X))^{\perp} = \ker(\mathrm{ad}(X))$. The latter space equals $\mathfrak{t}$ by (b). Hence (c) follows.

That (c) implies (d) is obvious. Now assume that (d) holds. Then it follows that $\mathrm{ad}(X)$ induces a linear automorphism of $\mathfrak{g}/\mathfrak{t}$. All eigenvalues of a linear automorphism must be different from zero, hence $\alpha(X) \neq 0$ for all $\alpha \in R$. $\square$

If $g \in G$, then $\mathrm{Ad}(g)$ is an automorphism of the Lie algebra $\mathfrak{g}$; hence $\mathrm{Ad}(g)\mathfrak{t}$ is a maximal torus in $\mathfrak{g}$. The following result asserts that all maximal tori of $\mathfrak{g}$ arise in this way.

Lemma 32.2 *Let $\mathfrak{t}, \mathfrak{t}'$ be two maximal tori in $\mathfrak{g}$. Then there exists a $g \in G$ such that*

$$\mathfrak{t}' = \mathrm{Ad}(g)\mathfrak{t}.$$

Proof: By the method of averaging over G we see that there exists a G-invariant positive definite inner product on $\mathfrak{g}$; select such an inner product $\langle \cdot \, , \, \cdot \rangle$. Moreover, select regular elements $X \in \mathfrak{t}$ and $Y \in \mathfrak{t}'$. Then by Lemma 32.1 we see that $\mathfrak{t}$ equals the centralizer of X in $\mathfrak{g}$. We consider the smooth function $f : G \to \mathbb{R}$ given by

$$f(x) = \langle \mathrm{Ad}(x)X \, , \, Y \rangle .$$

By compactness of G, the continuous function f attains a minimal value at a point $x_0 \in G$. It follows that for every $Z \in \mathfrak{g}$ the function $t \mapsto f(x_0 \exp tZ)$ has a minimum at $t = 0$, hence

$$0 = \left.\frac{d}{dt}\right|_{t=0} f(x_0 \exp tZ) = \langle \mathrm{Ad}(x_0)[Z, X]\,,\, Y\rangle = -\langle\, \mathrm{ad}(X)(Z)\,,\, \mathrm{Ad}(x_0)^{-1}Y\rangle.$$

By Lemma 32.1 we see that $\mathrm{ad}(X)$ maps $\mathfrak{g}$ onto $\mathfrak{t}^\perp$. Hence $\mathrm{Ad}(x_0)^{-1}Y \in (\mathfrak{t}^\perp)^\perp = \mathfrak{t}$. It follows from this that the maximal torus $\mathfrak{t}'' = \mathrm{Ad}(x_0)\mathfrak{t}$ contains Y; obviously $\mathfrak{t}''$ is contained in the centralizer of Y, which equals $\mathfrak{t}'$, by Lemma 32.1. By maximality of $\mathfrak{t}''$ it follows that $\mathfrak{t}' = \mathfrak{t}'' = \mathrm{Ad}(x_0)\mathfrak{t}$. □

If $g \in G$, then $\mathrm{Ad}(g)$ is an automorphism of the Lie algebra $\mathfrak{g}$. More generally we now consider an automorphism φ of the Lie algebra $\mathfrak{g}$; its complex linear extension, also denoted by φ is an automorphism of the complex Lie algebra $\mathfrak{g}_\mathbb{C}$. If $\mathfrak{t}$ is a maximal torus, then $\mathfrak{t}' = \varphi(\mathfrak{t})$ is a maximal torus as well. The map $\mathfrak{t}^*_\mathbb{C} \to \mathfrak{t}'^*_\mathbb{C}$ given by $\lambda \mapsto \lambda \circ \varphi^{-1}$ is a linear isomorphism, which we again denote by φ. With this notation we have:

Lemma 32.3 *Let φ be an automorphism of the Lie algebra $\mathfrak{g}$. If $\mathfrak{t}$ is a maximal torus in $\mathfrak{g}$, then $\mathfrak{t}' = \varphi(\mathfrak{t})$ is a maximal torus in $\mathfrak{g}$ as well. Moreover, the induced linear isomorphism $\varphi : \mathfrak{t}^*_\mathbb{C} \to \mathfrak{t}'^*_\mathbb{C}$ maps $R = R(\mathfrak{g}_\mathbb{C}, \mathfrak{t})$ bijectively onto $R' = R(\mathfrak{g}_\mathbb{C}, \mathfrak{t}')$. Finally, if $\alpha \in R$, then*

$$\varphi(\mathfrak{g}_{\mathbb{C}\alpha}) = \mathfrak{g}_{\mathbb{C}\varphi(\alpha)}.$$

Proof: Let $\alpha \in R$ and let $X \in \mathfrak{g}_{\mathbb{C}\alpha}$. Then, for every $H' \in \mathfrak{t}'$,

$$[H', \varphi(X)] = \varphi([\varphi^{-1}(H'), X]) = \varphi(\alpha(\varphi^{-1}(H')X) = [\varphi(\alpha)](H')\varphi(X).$$

From this we see that $\varphi(\alpha) \in R'$ and $\varphi(\mathfrak{g}_{\mathbb{C}\alpha}) \subset \mathfrak{g}_{\mathbb{C}\varphi(\alpha)}$. The proof is completed by applying the same reasoning to the inverse of φ. □

Corollary 32.4 *Let R, R' be the collections of roots associated with two maximal tori $\mathfrak{t}, \mathfrak{t}'$ of $\mathfrak{g}$. Then there exists a bijective linear map from $i\mathfrak{t}^*$ onto $i\mathfrak{t}'^*$ which maps R onto R'.*

Proof: By Lemma 32.2 there exists a $g \in G$ such that $\mathrm{Ad}(g)\mathfrak{t} = \mathfrak{t}'$. The map $\varphi = \mathrm{Ad}(g)$ is an automorphism of $\mathfrak{g}$. By Lemma 32.3 the induced isomorphism from $\mathfrak{t}^*_\mathbb{C}$ onto $\mathfrak{t}'^*_\mathbb{C}$ satisfies all requirements. □

33 Automorphisms of a Lie algebra

In this section we assume that $\mathfrak{g}$ is a finite dimensional real or complex Lie algebra. We denote by $\mathrm{Aut}(\mathfrak{g})$ the *group of automorphisms* of the Lie algebra $\mathfrak{g}$. This is clearly a subgroup of $\mathrm{GL}(\mathfrak{g})$. In fact, $\mathrm{Aut}(\mathfrak{g})$ is the intersection, for $X, Y \in \mathfrak{g}$, of the subsets $A_{X,Y}$ consisting of $\varphi \in \mathrm{GL}(\mathfrak{g})$ with $\varphi([X, Y]) - [\varphi(X), \varphi(Y)] = 0$. All of these subsets are closed, hence $\mathrm{Aut}(\mathfrak{g})$ is a closed subgroup of $\mathrm{GL}(\mathfrak{g})$. Its Lie algebra is a sub Lie algebra of $\mathrm{End}(\mathfrak{g})$, equipped with the commutator bracket.

A *derivation* of $\mathfrak{g}$ is by definition a linear map $D \in \mathrm{End}(\mathfrak{g})$ such that

$$D([X, Y]) = [D(X), Y] + [X, D(Y)] \qquad (X, Y \in \mathfrak{g}).$$

One readily sees that the space $\mathrm{Der}(\mathfrak{g})$ of all derivations of $\mathfrak{g}$ is a Lie subalgebra of $\mathrm{End}(\mathfrak{g})$.

Proposition 33.1 $\mathrm{Der}(\mathfrak{g})$ *is the Lie algebra of* $\mathrm{Aut}(\mathfrak{g})$.

Proof: Let D be an element in the Lie algebra of $\mathrm{Aut}(\mathfrak{g})$. Then $\exp(tD) \in \mathrm{Aut}(\mathfrak{g})$ for all $t \in \mathbb{R}$. Let $X, Y \in \mathfrak{g}$, then it follows that

$$e^{tD}[X, Y] = [e^{tD}X, e^{tD}Y].$$

Differentiating this expression with respect to t at $t = 0$ we obtain that $D[X, Y] = [DX, Y] + [X, DY]$; hence D is a derivation. It follows that the Lie algebra of $\mathrm{Aut}(\mathfrak{g})$ is contained in $\mathrm{Der}(\mathfrak{g})$.

To prove the converse inclusion, let $D \in \mathrm{Der}(\mathfrak{g})$, and let $X, Y \in \mathfrak{g}$. Consider the function $c : \mathbb{R} \to \mathfrak{g}$ defined by

$$c(t) = e^{-tD}[e^{tD}X, e^{tD}Y].$$

Then c is differentiable, and its derivative is given by

$$\begin{aligned} c'(t) &= (\partial_t\, e^{-tD})[e^{tD}X, e^{tD}Y] + e^{-tD}[(\partial_t\, e^{tD}X), e^{tD}Y] + e^{-tD}[e^{tD}X, (\partial_t\, e^{tD}Y)] \\ &= -e^{-tD}D([e^{tD}X, e^{tD}Y]) + e^{-tD}[De^{tD}X, e^{tD}Y] + e^{-tD}[e^{tD}X, De^{tD}Y] \\ &= 0. \end{aligned}$$

Hence c is constantly equal to $c(0) = [X, Y]$. It follows from this that $e^{tD} \in \mathrm{Aut}(\mathfrak{g})$ for all $t \in \mathbb{R}$, hence D belongs to the Lie algebra of $\mathrm{Aut}(\mathfrak{g})$. $\square$

Corollary 33.2 *The homomorphism* ad *maps* $\mathfrak{g}$ *into* $\mathrm{Der}(\mathfrak{g})$. *If* $X \in \mathfrak{g}$, *then* $e^{\,\mathrm{ad}X}$ *is an automorphism of the Lie algebra* $\mathfrak{g}$.

Proof: The first assertion follows from the Jacobi identity. The last statement is now a consequence of the above lemma; indeed $e^{D} \in \mathrm{Aut}(\mathfrak{g})$, for $D \in \mathrm{Der}(\mathfrak{g})$. $\square$

The subgroup of $\mathrm{Aut}(\mathfrak{g})$ generated by $e^{\,\mathrm{ad}(X)}$, $X \in \mathfrak{g}$ is called the group of *interior automorphisms* of $\mathfrak{g}$, notation: $\mathrm{Int}(\mathfrak{g})$. Its Lie algebra equals $\mathrm{ad}(\mathfrak{g})$, see Section 7.

34 The Killing form

Let $\mathfrak{g}$ be a finite dimensional Lie algebra over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Its *Killing form* is by definition the bilinear form $\mathfrak{g} \times \mathfrak{g} \to \mathbb{K}$ defined by

$$B(X, Y) = \mathrm{tr}(\,\mathrm{ad}(X) \circ \mathrm{ad}(Y)), \qquad (X, Y \in \mathfrak{g}).$$

Lemma 34.1 *The Killing form is symmetric. Morever, if* $\varphi \in \mathrm{Aut}(\mathfrak{g})$, *then*

$$B(\varphi(X), \varphi(Y)) = B(X, Y) \qquad (X, Y \in \mathfrak{g}). \tag{48}$$

Finally,

$$B([Z, X], Y) = -B(X, [Z, Y]) \qquad (X, Y, Z \in \mathfrak{g}). \tag{49}$$

Proof: If A, B are endomorphisms of a linear space, it is well known that $AB - BA$ has trace 0. Hence $\mathrm{tr}(\mathrm{ad}(X) \circ \mathrm{ad}(Y)) = \mathrm{tr}(\mathrm{ad}(Y) \circ \mathrm{ad}(X))$, for $X, Y \in \mathfrak{g}$, and the symmetry of B follows.

If φ is a Lie algebra automorphism of $\mathfrak{g}$, then it follows that $\varphi \circ \mathrm{ad}X = \mathrm{ad}(\varphi(X)) \circ \varphi$. Hence $\mathrm{ad}(\varphi(X)) = \varphi \circ \mathrm{ad}(X) \circ \varphi^{-1}$. Using this and conjugation invariance of the trace (48) follows.

Let $t \in \mathbb{R}$, then $e^{t\,\mathrm{ad}Z} \in \mathrm{Aut}(\mathfrak{g})$; thus (48) holds with $e^{t\,\mathrm{ad}Z}$ inserted for φ. Differentiation of the resulting identity with respect to t at $t = 0$ yields (49).

The latter identity can also be derived algebraically, as follows. We note that $\mathrm{ad}([Z, X]) = \mathrm{ad}(Z)\,\mathrm{ad}(X) - \mathrm{ad}(X)\,\mathrm{ad}(Z)$, hence

$$\begin{aligned} B([Z, X], Y) &= \mathrm{tr}(\mathrm{ad}(Z)\,\mathrm{ad}(X)\,\mathrm{ad}(Y)) - \mathrm{tr}(\mathrm{ad}(X)\,\mathrm{ad}(Z)\,\mathrm{ad}(Y)) \\ &= \mathrm{tr}(\mathrm{ad}(X)\,\mathrm{ad}(Y)\,\mathrm{ad}(Z)) - \mathrm{tr}(\mathrm{ad}(Y)\,\mathrm{ad}(Z)\,\mathrm{ad}(X)) \\ &= \mathrm{tr}\,(\mathrm{ad}(X)\,\mathrm{ad}([Y, Z])) = -B(X, [Z, Y]). \end{aligned}$$

$\square$

The identity (49) is known as *invariance of the Killing form*. If $\mathfrak{v}$ is a linear subspace of $\mathfrak{g}$, then by $\mathfrak{v}^\perp$ we shall denote its orthocomplement with respect to B, i.e., the collection of $Y \in \mathfrak{g}$ such that $B(X, Y) = 0$ for all $X \in \mathfrak{v}$. Note that from the invariance of B the following lemma is an immediate consequence.

Lemma 34.2 *Let* $\mathfrak{v} \subset \mathfrak{g}$. *If* $\mathfrak{v}$ *is an ideal, then so is* $\mathfrak{v}^\perp$.

35 Compact and reductive Lie algebras

Throughout this section $\mathfrak{g}$ will be a real finite dimensional Lie algebra.

The algebra $\mathfrak{g}$ is called *compact* if it is isomorphic to the Lie algebra of a compact Lie group. The purpose of this section is to derive useful criteria for a Lie algebra to be compact. Our starting point is the following result.

Let B denote the Killing form of $\mathfrak{g}$. We recall from Lemma 34.1 that B is a symmetric bilinear form. Since ad is a Lie algebra homomorphism, its kernel $\mathfrak{z} := \ker \mathrm{ad}$ is an ideal in $\mathfrak{g}$. This ideal is called the *center* of $\mathfrak{g}$.

Lemma 35.1 *Let* $\mathfrak{g}$ *be compact. Then the Killing form* B *is negative semi-definite. Moreover,* $\mathfrak{g}^\perp = \mathfrak{z}$.

Proof: We may assume that $\mathfrak{g}$ is the Lie algebra of the compact group G. The representation Ad of G in $\mathfrak{g}_{\mathbb{C}}$ is unitarizable, hence there exists a positive definite inner product on $\mathfrak{g}_{\mathbb{C}}$ that is $\mathrm{Ad}(G)$-invariant. With respect to this inner product we have $\mathrm{Ad}(G) \subset \mathrm{U}(\mathfrak{g}_{\mathbb{C}})$.

Since ad is the infinitesimal representation of $\mathfrak{g}$ in $\mathfrak{g}_{\mathbb{C}}$ associated with Ad, it follows that $\mathrm{ad}(\mathfrak{g}) \subset \mathfrak{u}(\mathfrak{g}_{\mathbb{C}})$. Hence, $\mathrm{ad}(X)$ is an anti-symmetric Hermitian endomorphism of $\mathfrak{g}_{\mathbb{C}}$, for $X \in \mathfrak{g}$. This implies that $\mathrm{ad}X$ has a an orthonormal basis of eigenvectors and imaginary eigenvalues. Hence $\mathrm{ad}(X)^2$ has the same orthonormal basis of eigenvectors with eigenvalues ≤ 0. It follows that $B(X, X) = \mathrm{tr}\, \mathrm{ad}(X)^2 \leq 0$. Hence, B is negative semi-definite. Moreover, if $B(X, X) = 0$ then $\mathrm{tr}\, \mathrm{ad}(X)^2 = 0$ and it follows that all eigenvalues of $\mathrm{ad}(X)^2$, hence of $\mathrm{ad}(X)$ are zero. Hence, $\mathrm{ad}(X) = 0$. This shows that $\mathfrak{g}^\perp \subset \mathfrak{z}$. The converse inclusion is obvious. $\square$

If $\mathfrak{v}, \mathfrak{w}$ are subspaces of $\mathfrak{g}$, then by $[\mathfrak{v}, \mathfrak{w}]$ we denote the subspace of $\mathfrak{g}$ spanned by the elements $[X, Y]$, where $X \in \mathfrak{v}, Y \in \mathfrak{w}$. If $\mathfrak{v}, \mathfrak{w}$ are ideals, then $[\mathfrak{v}, \mathfrak{w}]$ is an ideal of $\mathfrak{g}$. Indeed, this follows by a straightforward application of the Jacobi identity. In particular

$$\mathcal{D}\mathfrak{g} := [\mathfrak{g}, \mathfrak{g}]$$

is an ideal of $\mathfrak{g}$, called the *commutator ideal.*

If $\mathfrak{a}, \mathfrak{b}$ are ideals of $\mathfrak{g}$, then $\mathfrak{a} + \mathfrak{b}$ is an ideal of $\mathfrak{g}$ as well. Two ideals $\mathfrak{a}$ and $\mathfrak{b}$ of $\mathfrak{g}$ are said to be complementary if

$$\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b} \tag{50}$$

as linear spaces.

Lemma 35.2 *If $\mathfrak{a}, \mathfrak{b}$ are ideals of $\mathfrak{g}$, then $[\mathfrak{a}, \mathfrak{b}] \subset \mathfrak{a} \cap \mathfrak{b}$. If $\mathfrak{a}$ and $\mathfrak{b}$ are complementary, then $[\mathfrak{a}, \mathfrak{b}] = 0$. In that case, (50) is a direct sum of Lie algebras.*

Proof: Since $\mathfrak{a}$ is an ideal, $[\mathfrak{a}, \mathfrak{b}] = [\mathfrak{b}, \mathfrak{a}] \subset \mathfrak{a}$. Similarly, $[\mathfrak{a}, \mathfrak{b}] \subset \mathfrak{b}$. Hence $[\mathfrak{a}, \mathfrak{b}] \subset \mathfrak{a} \cap \mathfrak{b}$. The last two assertions now readily follow. $\square$

Lemma 35.3 *Let $\mathfrak{g}$ be a compact Lie algebra. Then every ideal of $\mathfrak{g}$ has a complementary ideal.*

Proof: As in the proof of Lemma 35.1 there exists a positive definite inner product $\langle \cdot\, , \,\cdot \rangle$ on $\mathfrak{g}$ with respect to which $\mathrm{Ad}(G) \subset \mathrm{O}(\mathfrak{g})$. It follows that $\mathrm{ad}(\mathfrak{g}) \subset \mathfrak{o}(\mathfrak{g})$, or, equivalently, that $\langle [X, U]\, , \, V \rangle = -\langle U\, , \, [X, V] \rangle$, for all $X, U, V \in \mathfrak{g}$. By this property, if $\mathfrak{a} \subset \mathfrak{g}$ is an ideal, then $\mathfrak{a}^\perp$ is an ideal; moreover, $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$. $\square$

Lemma 35.4 *Let $\mathfrak{g}$ have the property that every ideal has a complementary ideal. Then*

$$\mathfrak{g} = \mathfrak{z} \oplus \mathcal{D}\mathfrak{g}.$$

Proof: The ideal $\mathcal{D}\mathfrak{g}$ has a complementary ideal, say $\mathfrak{a}$. Since obviously $[\mathfrak{g}, \mathfrak{a}] \subset \mathcal{D}\mathfrak{g}$, we have $[\mathfrak{g}, \mathfrak{a}] \subset \mathfrak{a} \cap \mathcal{D}\mathfrak{g} = 0$, from which we conclude that $\mathfrak{a} \subset \mathfrak{z}$. It follows that $\mathfrak{g} = \mathfrak{z} + \mathcal{D}\mathfrak{g}$.

The ideal $\mathfrak{z}$ has a complementary ideal, say $\mathfrak{b}$. Thus, $\mathfrak{g} = \mathfrak{z} \oplus \mathfrak{b}$. Now $\mathcal{D}\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}] \subset [\mathfrak{g}, \mathfrak{z}] + [\mathfrak{g}, \mathfrak{b}] \subset \mathfrak{b}$, from which we conclude that $\mathfrak{z} \cap \mathcal{D}\mathfrak{g} = 0$. $\square$

Theorem 35.5 *The following assertions are equivalent.*

(a) $\mathfrak{g}$ *is compact*

(b) $\mathfrak{g} = \mathfrak{z} \oplus \mathcal{D}\mathfrak{g}$ *and* B *is negative definite on* $\mathcal{D}\mathfrak{g}$.

(c) *There exists a subalgebra* $\mathfrak{g}' \subset \mathfrak{g}$ *such that* $\mathfrak{g} = \mathfrak{z} + \mathfrak{g}'$ *and such that* B *is negative definite on* $\mathfrak{g}'$.

Finally, if $\mathfrak{g}'$ *is as in (c) then* $\mathfrak{g}' = \mathcal{D}\mathfrak{g}$.

Proof: First, assume that (a) is valid. Then $\mathfrak{g} = \mathfrak{z} \oplus \mathcal{D}\mathfrak{g}$ by Lemma 35.4. By Lemma 35.1 it follows that $B < 0$ on $\mathcal{D}\mathfrak{g}$. Hence (b). The implication (b) $\Rightarrow$ (c) is obvious. The implication (c) $\Rightarrow$ (a) and the final assertion will be established in the following lemma. □

Lemma 35.6 *Let the Killing form of* $\mathfrak{g}$ *be negative definite. Then* $\mathrm{Aut}(\mathfrak{g})$ *is compact. Moreover,* ad *is a Lie algebra isomorphism from* $\mathfrak{g}$ *onto* $\mathrm{Der}(\mathfrak{g})$. *In particular it follows that* $\mathfrak{g}$ *is compact and has trivial center.*

Proof: Let $\mathrm{O}(\mathfrak{g})$ denote the group of invertible transformations of $\mathfrak{g}$ that are orthogonal relative to the positive definite inner product $-B$. Then $\mathrm{O}(\mathfrak{g})$ is compact. From (48) it follows that the closed subgroup $\mathrm{Aut}(\mathfrak{g})$ of $\mathrm{GL}(\mathfrak{g})$ is contained in the compact group $\mathrm{O}(\mathfrak{g})$, hence is compact.

By definition of the Killing form, $\ker \mathrm{ad} \subset \mathfrak{g}^{\perp}$; since B is non-degenerate, it follows that ad is an injective Lie algebra homomorphism. It follows from the Jacobi identity that ad maps $\mathfrak{g}$ into $\mathrm{Der}(\mathfrak{g})$.

If $D \in \mathrm{Der}(\mathfrak{g})$, then for $X \in \mathfrak{g}$ we have that $[D \circ \mathrm{ad}(X)](Y) = \mathrm{ad}(DX)Y + [\mathrm{ad}(X) \circ D](Y)$ for $Y \in \mathfrak{g}$. Hence

$$[D, \mathrm{ad}(X)] = \mathrm{ad}(DX). \tag{51}$$

It follows that $\mathrm{ad}(\mathfrak{g})$ is an ideal in $\mathrm{Der}(\mathfrak{g})$.

Now $\mathrm{Der}(\mathfrak{g})$ is the Lie algebra of the compact group $\mathrm{Aut}(\mathfrak{g})$, see Proposition 33.1. It follows that $\mathrm{Der}(\mathfrak{g})$ is compact. By application of Lemma 35.3 it follows that $\mathrm{ad}(\mathfrak{g})$ has a complementary ideal $\mathfrak{b}$ in $\mathrm{Der}(\mathfrak{g})$. Let $D \in \mathfrak{b}$. Then D commutes with $\mathrm{ad}(\mathfrak{g})$, hence from (51) we see that $\mathrm{ad}(DX) = 0$, whence $DX = 0$ for all $X \in \mathfrak{g}$. Hence $D = 0$. We conclude that $\mathfrak{b} = 0$, hence $\mathrm{ad}(\mathfrak{g}) = \mathrm{Der}(\mathfrak{g})$.

It follows that ad is an isomorphism from $\mathfrak{g}$ onto $\mathrm{Der}(\mathfrak{g})$; the latter is the Lie algebra of the compact group $\mathrm{Aut}(\mathfrak{g})$. Hence $\mathfrak{g}$ is compact. □

Completion of the proof of Theorem 35.5: Assume that (c) holds. Then $\mathfrak{g}'$ has negative definite Killing form, hence is compact. Since $\mathfrak{z} = \ker \mathrm{ad} \subset \mathfrak{g}^{\perp}$ it follows that $\mathfrak{z} \cap \mathfrak{g}' = 0$. Hence, $\mathfrak{g} = \mathfrak{z} \oplus \mathfrak{g}'$ as linear spaces. Since obviously $[\mathfrak{z}, \mathfrak{g}'] = 0$, the mentioned direct sum is a direct sum of Lie algebras.

Let G' be a compact Lie group with algebra isomorphic to $\mathfrak{g}'$. Let $n = \dim \mathfrak{z}$. Then $\mathfrak{z} \simeq \mathbb{R}^n$ as abelian Lie algebras. Hence, the compact torus $T := \mathbb{R}^n/\mathbb{Z}^n$ has Lie algebra isomorphic $\mathfrak{z}$. It follows that the compact group $G := T \times G'$ has Lie algebra isomorphic to $\mathfrak{z} \oplus \mathfrak{g}' = \mathfrak{g}$. (a) follows.

Finally, let $\mathfrak{g}'$ be as in (c). Then by the above reasoning, $\mathfrak{g} = \mathfrak{z} \oplus \mathfrak{g}'$ as Lie algebras. It follows that $\mathcal{D}\mathfrak{g} \subset [\mathfrak{g}', \mathfrak{g}'] \subset \mathfrak{g}'$. Now apply (b) to conclude that $\mathcal{D}\mathfrak{g} = \mathfrak{g}'$. □

The Lie algebra $\mathfrak{g}$ is called *simple* if it is not abelian and has no ideals besides 0 and $\mathfrak{g}$. It is called *semisimple* if it is a direct sum of simple ideals.

Lemma 35.7 *Let $\mathfrak{g}$ be semisimple, then $\mathfrak{z} = 0$ and $\mathcal{D}\mathfrak{g} = \mathfrak{g}$.*

Proof: Let $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n$ be a decomposition into simple ideals. We observe that each ideal $\mathfrak{g}_j$ is non-abelian, hence $[\mathfrak{g}_j, \mathfrak{g}_j]$ is a non-trivial ideal in $\mathfrak{g}_j$. Since the latter is simple, we conclude that $\mathcal{D}\mathfrak{g}_j = \mathfrak{g}_j$. Since the $\mathfrak{g}_j$ mutually commute, it follows that $\mathcal{D}\mathfrak{g} = \mathcal{D}\mathfrak{g}_1 + \cdots + \mathcal{D}\mathfrak{g}_n = \mathfrak{g}$. If X belongs to the center of $\mathfrak{g}$, write $X = X_1 + \cdots + X_n$, according to the decomposition (52). Then X commutes with $\mathfrak{g}_i$ and each X_j, for $j \neq i$, commutes with $\mathfrak{g}_i$. Hence, X_i commutes with $\mathfrak{g}_i$ as well. Hence X belongs to the center $\mathfrak{z}_i$ of $\mathfrak{g}_i$. This center is an ideal different from $\mathfrak{g}_i$, since $\mathfrak{g}_i$ is not abelian. Since $\mathfrak{g}_i$ is simple, $\mathfrak{z}_i = 0$. We conclude that $X = 0$. Hence, $\mathfrak{z} = 0$. ☐

Proposition 35.8 *The following assertions are equivalent.*

(a) *The algebra $\mathfrak{g}$ is compact and has trivial center;*

(b) *The Killing form of $\mathfrak{g}$ is negative definite;*

(c) *The algebra $\mathfrak{g}$ is compact and semisimple.*

Proof: Assume (a). Then by Theorem 35.5, $\mathfrak{g} = \mathcal{D}\mathfrak{g}$ and (b) follows. Since the implication (c) $\Rightarrow$ (a) follows from Lemma 35.7, it remains to establish the implication (b) $\Rightarrow$ (c). Assume (b). If $\mathfrak{a}, \mathfrak{b}$ are $\mathrm{ad}(\mathfrak{g})$-invariant subspaces of $\mathfrak{g}$ with $\mathfrak{a} \subset \mathfrak{b}$, then $\mathfrak{a}_1 := \mathfrak{a}^\perp \cap \mathfrak{b}$ is an $\mathrm{ad}(\mathfrak{g})$-invariant subspace of $\mathfrak{g}$ and $\mathfrak{b} = \mathfrak{a} \oplus \mathfrak{a}_1$ is a direct sum decomposition of $\mathfrak{b}$. Applying this observation repeatedly, we obtain a direct sum decomposition

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n \tag{52}$$

of non-trivial $\mathrm{ad}(\mathfrak{g})$-invariant subspaces, such that $\mathfrak{g}_j$ has no $\mathrm{ad}(\mathfrak{g})$-invariant subspaces besides 0 and $\mathfrak{g}_j$, for each $j \in \mathfrak{g}$. The assertion that $\mathfrak{g}_j$ is $\mathrm{ad}(\mathfrak{g})$-invariant is equivalent to the assertion that $\mathfrak{g}_j$ is an ideal in $\mathfrak{g}$. Hence, (52) is a direct sum of ideals. It follows that $[\mathfrak{g}_i, \mathfrak{g}_j] = 0$ for $i \neq j$. Hence every ideal $\mathfrak{a}$ of $\mathfrak{g}_i$ is also an ideal of $\mathfrak{g}$. We see that each algebra $\mathfrak{g}_i$ has no ideals besides 0 and itself. If $\mathfrak{g}_i$ were abelian, it would centralize $\mathfrak{g}_i$ and $\mathfrak{g}_j$ for all $j \neq i$, hence $\mathfrak{g}$. This would imply that $\mathfrak{g}_i \subset \mathfrak{g}^\perp = 0$, contradiction. Thus, each ideal $\mathfrak{g}_i$ is simple, and it follows that $\mathfrak{g}$ is semisimple. ☐

Lemma 35.9 *Let $\mathfrak{g}$ be a finite dimensional Lie algebra, and let $\mathfrak{a}$ be a simple ideal of $\mathfrak{g}$. If $\mathfrak{g} = \mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_n$ is a direct sum of ideals, then there exists a unique $1 \leq j \leq n$ such that $\mathfrak{g}_j \supset \mathfrak{a}$.*

Proof: We note that $[\mathfrak{g}, \mathfrak{a}] \subset \mathfrak{a}$ since $\mathfrak{a}$ is an ideal, and $[\mathfrak{g}, \mathfrak{a}] \supset [\mathfrak{a}, \mathfrak{a}] = \mathfrak{a}$ since $\mathfrak{a}$ is simple. Hence $[\mathfrak{g}, \mathfrak{a}] = \mathfrak{a}$. From the direct sum decomposition it now follows that

$$\mathfrak{a} = [\mathfrak{g}_1, \mathfrak{a}] + \cdots + [\mathfrak{g}_n, \mathfrak{a}].$$

Hence there exists a j such that $[\mathfrak{g}_j, \mathfrak{a}] \neq 0$. Since $\mathfrak{a}$ is simple and $[\mathfrak{g}_j, \mathfrak{a}]$ is an ideal in $\mathfrak{a}$ we must have $[\mathfrak{g}_j, \mathfrak{a}] = \mathfrak{a}$. This implies that $\mathfrak{a} = [\mathfrak{a}, \mathfrak{g}_j] \subset \mathfrak{g}_j$. Of course j is uniquely determined by the latter property. ☐

Lemma 35.10 *Let $\mathfrak{g}$ be semisimple, and let S be the collection of simple ideals in $\mathfrak{g}$. Every ideal $\mathfrak{a} \subset \mathfrak{g}$ is the direct sum of the ideals from S that are contained in $\mathfrak{a}$. In particular, $\mathfrak{g}$ is the direct sum of the ideals from S.*

Proof: We may express $\mathfrak{g}$ as a direct sum of simple ideals of the form (52). If $\mathfrak{a} \in S$ then, by the previous lemma, $\mathfrak{a} \subset \mathfrak{g}_j$ for some j. Since $\mathfrak{g}_j$ is simple, it follows that $\mathfrak{a} = \mathfrak{g}_j$. We conclude that $S = \{\mathfrak{g}_1, \ldots, \mathfrak{g}_n\}$.

Let now $\mathfrak{b} \subset \mathfrak{g}$ be any ideal. We will show that $\mathfrak{b}$ is the direct sum of the simple ideals from $S(\mathfrak{b}) := \{\mathfrak{a} \in S \mid \mathfrak{a} \subset \mathfrak{b}\}$ by induction on $\#S$. First, assume that $\#S = 1$. Then $\mathfrak{g}$ is simple, hence $\mathfrak{b} = 0$ or $\mathfrak{b} = \mathfrak{g}$ and the result follows. Now assume that $\#S > 1$ and that the result has been established for $\mathfrak{g}$ with S of strictly smaller cardinality. If $\mathfrak{b} = 0$ there is nothing to prove. If $\mathfrak{b} \neq 0$, then $[\mathfrak{g}, \mathfrak{b}] \neq 0$ since $\mathfrak{z} = 0$. It follows that $[\mathfrak{g}_j, \mathfrak{b}] \neq 0$ for some j. But $[\mathfrak{g}_j, \mathfrak{b}]$ is an ideal in the simple algebra $\mathfrak{g}_j$, hence $\mathfrak{g}_j = [\mathfrak{g}_j, \mathfrak{b}] \subset \mathfrak{b}$. Let $\mathfrak{g}'$ be the direct sum of the ideals from the non-empty set $S(\mathfrak{b})$, and let $\mathfrak{g}''$ be the direct sum of the ideals from $S \setminus S(\mathfrak{b})$. Then $\mathfrak{g}' \subset \mathfrak{b}$ and $\mathfrak{g} = \mathfrak{g}' \oplus \mathfrak{g}''$, hence $\mathfrak{b} = \mathfrak{g}' \oplus (\mathfrak{b} \cap \mathfrak{g}'')$. Now $\mathfrak{b} \cap \mathfrak{g}''$ is an ideal in the semisimple algebra $\mathfrak{g}''$. By the induction hypothesis, $\mathfrak{b} \cap \mathfrak{g}''$ is the direct sum of the ideals from S contained in both $\mathfrak{b}$ and $\mathfrak{g}''$. This set is empty, hence $\mathfrak{b} \cap \mathfrak{g}'' = 0$ and the result follows. $\square$

36 Root systems for compact algebras

In this section we assume that $\mathfrak{g}$ is a compact Lie algebra with trivial center. Then B, the Killing form of $\mathfrak{g}$, is symmetric and negative definite, see Lemma 34.1 and Proposition 35.8. The extension of B to a complex bilinear form on $\mathfrak{g}_{\mathbb{C}}$ equals the Killing form of $\mathfrak{g}_{\mathbb{C}}$ and is also denoted by B. We fix a maximal torus $\mathfrak{t}$ in $\mathfrak{g}$. Let R be the associated root system and let R^+ be a choice of positive roots.

If $\alpha \in R$, then $\alpha \in i\mathfrak{t}^*$. Therefore, $\ker\alpha$ is a hyperplane in $\mathfrak{t}_{\mathbb{C}}$, which is the complexification of the hyperplane $\ker\mathfrak{a} \cap i\mathfrak{t}$ in $i\mathfrak{t}$. The Killing form B is negative definite on $\mathfrak{t}$, hence positive definite on $i\mathfrak{t}$. It follows that there exists a unique element $H_\alpha \in i\mathfrak{t}$ with the properties

$$H_\alpha \perp \ker\alpha \qquad \text{and} \qquad \alpha(H_\alpha) = 2. \tag{53}$$

Lemma 36.1 *Let $\lambda, \mu \in \mathfrak{t}_{\mathbb{C}}^*$.*

(a) *If $\lambda + \mu \neq 0$, then $B = 0$ on $\mathfrak{g}_{\mathbb{C}\lambda} \times \mathfrak{g}_{\mathbb{C}\mu}$.*

(b) *If $\alpha \in R$ and $X \in \mathfrak{g}_{\mathbb{C}\alpha}$, $Y \in \mathfrak{g}_{\mathbb{C}-\alpha}$, then $[X, Y] \in \mathfrak{t}_{\mathbb{C}}$ and*

$$B([X, Y], H) = B(X, Y)\,\alpha(H) \qquad (H \in \mathfrak{t}_{\mathbb{C}}). \tag{54}$$

(c) $[\mathfrak{g}_{\mathbb{C}\alpha}, \mathfrak{g}_{\mathbb{C}-\alpha}] \subset \mathbb{C}H_\alpha$.

Proof: Let $X \in \mathfrak{g}_{\mathbb{C}\lambda}$ and $Y \in \mathfrak{g}_{\mathbb{C}\mu}$. Then by invariance of the Killing form we have, for all $H \in \mathfrak{t}_{\mathbb{C}}$,

$$[\lambda(H) + \mu(H)]\, B(X, Y) = B([H, X], Y) + B(X, [H, Y]) = 0.$$

From this, (a) follows.

Let now α, X, Y be as in (b). Then $[X, Y] \in \mathfrak{g}_{\mathbb{C}0} = \mathfrak{t}_{\mathbb{C}}$, by Corollary 31.11 and Lemma 31.4. Moreover, for all $H \in \mathfrak{t}_{\mathbb{C}}$ we have

$$B([X, Y], H) = -B(Y, [X, H]) = B(Y, [H, X]) = B(Y, \alpha(H)X) = B(X, Y)\alpha(H).$$

Hence, (b).

Finally, for (c) we note that (b) implies that $[X, Y] \perp \ker \alpha$, relative to B, for $X \in \mathfrak{g}_{\mathbb{C}\alpha}$ and $Y \in \mathfrak{g}_{\mathbb{C}-\alpha}$. In view of (53) this implies that $[\mathfrak{g}_{\mathbb{C}\alpha}, \mathfrak{g}_{\mathbb{C}-\alpha}] \subset \mathbb{C}H_\alpha$.

$\square$

Let $\tau : \mathfrak{g}_{\mathbb{C}} \to \mathfrak{g}_{\mathbb{C}}$ be the conjugation with respect to the real form $\mathfrak{g}$ of $\mathfrak{g}_{\mathbb{C}}$. Thus, $\tau = I$ on $\mathfrak{g}$ and $\tau = -I$ on $i\mathfrak{g}$. We recall that $\tau(\mathfrak{g}_{\mathbb{C}\alpha}) = \mathfrak{g}_{\mathbb{C}-\alpha}$, for all $\alpha \in R$, see proof of Lemma 31.14. We denote the positive definite inner product $-B(\,\cdot\,,\,\cdot\,)|_{\mathfrak{g}}$ by $\langle\,\cdot\,,\,\cdot\,\rangle$ and extend it to a Hermitian positive definite inner product on $\mathfrak{g}_{\mathbb{C}}$. Then

$$\langle X\,,\,Y\rangle = -B(X, \tau Y), \qquad (X, Y \in \mathfrak{g}_{\mathbb{C}}).$$

The following result is now immediate.

Corollary 36.2 *The root space decomposition*

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{t}_{\mathbb{C}} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_{\mathbb{C}\alpha}$$

is orthogonal with respect to the inner product $\langle\,\cdot\,,\,\cdot\,\rangle$.

Lemma 36.3 *Let $\alpha \in R$. Then there exists a $X_\alpha \in \mathfrak{g}_{\mathbb{C}\alpha}$ such that H_α, X_α and $Y_\alpha := -\tau X_\alpha$ form a standard $\mathfrak{sl}(2, \mathbb{C})$-triple.*

Proof: We observe that $\langle\,\cdot\,,\,\cdot\,\rangle$ is positive definite on $\mathfrak{g}_{\mathbb{C}\alpha} \oplus \mathfrak{g}_{\mathbb{C}-\alpha}$. Let X be a non-zero element in $\mathfrak{g}_{\mathbb{C}\alpha}$. Then $X + \tau X \neq 0$, hence $0 < \langle X + \tau X\,,\,X + \tau X\rangle = -B(X + \tau X, X + \tau X) = -2B(X, \tau X)$, since τX belongs to the space $\mathfrak{g}_{\mathbb{C}-\alpha}$ which is perpendicular to $\mathfrak{g}_{\mathbb{C}-\alpha}$ with respect to the Killing form. Put $Y = -\tau X$, then

$$B(X, Y) > 0.$$

Moreover, $\tau[X, \tau X] = [\tau X, X] = -[X, \tau X]$, hence $[X, Y] = -[X, \tau X] \in i\mathfrak{g} \cap \mathfrak{t}_{\mathbb{C}} = i\mathfrak{t}$. It now follows from Lemma 36.1 (c) that

$$[X, Y] = cH_\alpha,$$

for some $c \in \mathbb{R}$. Substituting this in (54) with $H = H_\alpha$, we obtain $B(cH_\alpha, H_\alpha) = 2B(X, Y)$. Since B is positive definite on $i\mathfrak{t}$, we have $B(H_\alpha, H_\alpha) > 0$. Hence $c > 0$. Take $X_\alpha = \sqrt{c}^{-1} X$. Then $[X_\alpha, -\tau X_\alpha] = c^{-1}\,[X, -\tau X] = H_\alpha$. Hence, X_α satisfies the requirements. $\square$

Example 36.4 Let $\mathfrak{g} = \mathfrak{su}(2)$. Then $\mathfrak{g}_{\mathbb{C}} = \mathfrak{sl}(2, \mathbb{C})$ and the conjugation τ is given by $\tau A = -A^*$, where the star indicates that the Hermitian adjoint is taken. Let H, X, Y be the usual standard triple in $\mathfrak{sl}(2, \mathbb{C})$. Thus, H is diagonal with entries $+1, -1$ and X is upper triangular with 1 in the upper right corner, Y is lower triangular with 1 in the lower left corner. Then $\mathfrak{t} = i\mathbb{R}H$ is a maximal torus in $\mathfrak{g}$. Moreover, $R = \{\alpha, -\alpha\}$, where $\alpha \in i\mathfrak{t}^*$ is determined by $\alpha(H) = 2$. Finally, $\tau X = -Y$, and we see that the above lemma with $X_\alpha = X$ gives us the usual standard triple.

If $\mathfrak{l}$ is a Lie subalgebra of $\mathfrak{g}$, then via the adjoint representation, $\mathfrak{g}_{\mathbb{C}}$ may be viewed as a $\mathfrak{l}$-module.

Lemma 36.5 *Let $\mathfrak{l}$ be a Lie subalgebra of $\mathfrak{g}$ and let $V \subset \mathfrak{g}_{\mathbb{C}}$ be a $\mathrm{ad}(\mathfrak{l})$-invariant subspace. Then V decomposes as a direct sum of irreducible $\mathfrak{l}$-modules.*

Proof: We observe that, for every $X \in \mathfrak{l}$, the endomorphism $\mathrm{ad}(X)$ of $\mathfrak{g}_{\mathbb{C}}$ is anti-symmetric with respect to $\langle \cdot\,,\,\cdot \rangle$. Indeed, this follows from invariance of the Killing form. Thus, if $W \subset V$ is a $\mathrm{ad}(\mathfrak{l})$-invariant subspace, then so is $W^\perp \cap V$. The lemma follows by repeated application of this observation. $\square$

Proposition 36.6 *Let $\alpha \in R$. Then $\dim_{\mathbb{C}} \mathfrak{g}_{\mathbb{C}\alpha} = 1$. Moreover, $R \cap \mathbb{R}\alpha = \{-\alpha, \alpha\}$. The algebra*

$$\mathfrak{s}_\alpha := \mathfrak{g}_{\mathbb{C}-\alpha} \oplus \mathbb{C}H_\alpha \oplus \mathfrak{g}_{\mathbb{C}\alpha}$$

is isomorphic with $\mathfrak{sl}(2, \mathbb{C})$. Its intersection with $\mathfrak{g}$ is isomorphic with $\mathfrak{su}(2)$. Finally, $\mathfrak{t}_\alpha := i\mathbb{R}H_\alpha$ is a maximal torus in $\mathfrak{s}_\alpha \cap \mathfrak{g}$ and the associated root system is equal to $\{\alpha|_{\mathfrak{t}_\alpha}, -\alpha|_{\mathfrak{t}_\alpha}\}$.

Proof: We fix $X_\alpha \in \mathfrak{g}_{\mathbb{C}\alpha}$ as in Lemma 36.3 and put $\mathfrak{s} = \mathbb{C}H_\alpha \oplus \mathbb{C}X_\alpha \oplus \mathbb{C}Y_\alpha$. Then $\mathfrak{s}$ is isomorphic with $\mathfrak{sl}(2, \mathbb{C})$. Moreover, $\mathfrak{s}$ is invariant under τ and $\mathfrak{g} \cap \mathfrak{s} = \ker(\tau - I) \cap \mathfrak{s} = i\mathbb{R}H_\alpha + \mathbb{R}(X_\alpha - Y_\alpha) + i\mathbb{R}(X_\alpha + Y_\alpha) \simeq \mathfrak{su}(2)$. The two last assertions of the proposition hold with $\mathfrak{s}$ in place of $\mathfrak{s}_\alpha$.

We consider the subspace

$$V = V(\alpha) := \sum_{\beta \in R \cap \mathbb{R}\alpha} \mathfrak{g}_{\mathbb{C}\beta} \oplus \mathbb{C}H_\alpha.$$

and leave it to the reader to verify that V is invariant under the adjoint action of $\mathfrak{s}$. It follows that V splits as a direct sum of irreducible $\mathfrak{s}$-modules. The decomposition of each irreducible $\mathfrak{s}$-submodule in $\mathbb{C}H_\alpha$-weight spaces is compatible with the given weight space decomposition of V. All weights of the irreducible representations of $\mathfrak{s}$ belong to $\frac{1}{2}\mathbb{Z}\alpha_0$, with $\alpha_0 = \alpha|_{\mathbb{C}H_\alpha}$. Thus, if $\beta \in R \cap \mathbb{R}\alpha$, then $\beta|_{\mathbb{C}H_\alpha} \in \frac{1}{2}\alpha_0$, from which we conclude that $\beta \in R \cap \frac{1}{2}\mathbb{Z}\alpha$. It follows that $R \cap \mathbb{R}\alpha = R \cap \frac{1}{2}\mathbb{Z}\alpha$. Put

$$V_{ev} = \sum_{\beta \in R \cap \mathbb{Z}\alpha} \mathfrak{g}_{\mathbb{C}\beta} \oplus \mathbb{C}H_\alpha.$$

and let V_{odd} be the sum of the remaining root spaces in V. Then both V_{ev} and V_{odd} are $\mathfrak{s}_\alpha$-invariant. The first of these spaces splits as a direct sum of irreducible $\mathfrak{s}$-modules all of whose weights belong to $\mathbb{Z}\alpha_0$. By the classification of irreducible $\mathfrak{sl}(2,\mathbb{C})$-modules it follows that each of the irreducible summands has a zero weight space, which must be contained in $\mathbb{C}H_\alpha$. It follows that V_{ev} has only one irreducible summand, hence is irreducible. Since $\mathfrak{s} \subset V_{ev}$ is an invariant subspace, it follows that $\mathfrak{s} = V_{ev}$. This implies that $R \cap \mathbb{Z}\alpha = \{\alpha, -\alpha\}$ and $\mathfrak{s} = \mathfrak{s}_\alpha$.

It remains to be shown that V_{odd} is the zero space. Assume not. Then V has a $\mathbb{C}H_\alpha$-weight of the form $(2n+1)/2\alpha_0$, with $n \in \mathbb{Z}$. This weight occurs in an irreducible summand of the $\mathfrak{s}_\alpha$-module V_{odd}. From the classification of the irreducible $\mathfrak{sl}(2,\mathbb{C})$-modules, we see that $\frac{1}{2}\alpha_0$ then also occurs as a weight in the irreducible summand, hence in V. Put $\alpha' = \frac{1}{2}\alpha$. Then it follows that $\mathfrak{g}_{\mathbb{C}\alpha'} \neq 0$, hence $\alpha' \in R$. Define $\mathfrak{s}_{\alpha'}$ as above, with α' in place of α. Then $V(\alpha) = V(\alpha') = V(\alpha')_{even}$. By the first part of the proof, applied with α' in place of α, it follows that $V(\alpha) = V(\alpha')_{ev} = \mathfrak{s}_{\alpha'}$. This contradicts the fact that $\mathfrak{g}_{\mathbb{C}\alpha} \subset V(\alpha)$.

We conclude that $V_{odd} = 0$. Hence, $V = V_{ev} = \mathfrak{s}_\alpha = \mathfrak{s}$ and all assertions follow. □

Let $\alpha \in R$. Then by s_α we denote the B-orthogonal reflection in the hyperplane $\ker\alpha$ in $i\mathfrak{t}$. Thus, $s_\alpha(H_\alpha) = -H_\alpha$ and $s_\alpha = I$ on $\ker\alpha$, from which one readily deduces that

$$s_\alpha(H) = H - \alpha(H)H_\alpha \qquad (H \in i\mathfrak{t}).$$

The complex linear extension of s_α to $\mathfrak{t}_\mathbb{C}$, also denoted s_α, is given by the same formula, for $H \in \mathfrak{t}_\mathbb{C}$.

If V is a finite dimensional real linear space, equipped with a positive definite inner product $\langle \cdot\,,\,\cdot \rangle$, then the map $v \mapsto \langle v\,,\,\cdot\rangle$ defines a linear isomorphism $j : V \to V^*$. We equip V^* with the so called dual inner product. This is defined to be the unique inner product that makes j orthogonal. Thus, if $\lambda, \mu \in V^*$ then

$$\langle \lambda\,,\,\mu\rangle = \langle j^{-1}\lambda\,,\,j^{-1}\mu\rangle = \lambda(j^{-1}\mu) = \mu(j^{-1}\lambda).$$

If $A : V \to V$ is orthogonal, then so is $j \circ A \circ j^{-1} : V^* \to V^*$. Using the definitions one readily verifies that $j \circ A \circ j^{-1} = A^{-1*}$. In this case we agree to write A for the orthogonal map $A^{-1*} : V^* \to V^*$. Thus, for $\eta \in V^*$ we write $A\eta = \eta \circ A^{-1}$.

Following the above convention for $V = i\mathfrak{t}$ equipped with the positive definite inner product B, we obtain an orthogonal map $s_\alpha : i\mathfrak{t}^* \to i\mathfrak{t}^*$ defined by $s_\alpha\lambda = \lambda \circ s_\alpha^{-1} = \lambda \circ s_\alpha$, for $\lambda \in i\mathfrak{t}^*$. Let $H \in i\mathfrak{t}$; then it follows by application of the above formula for the reflection s_α that $s_\alpha(\lambda)(H) = \lambda(H - \alpha(H)H_\alpha) = \lambda(H) - \lambda(H_\alpha)\alpha(H)$. From this we see that

$$s_\alpha\lambda = \lambda - \lambda(H_\alpha)\alpha, \qquad (\lambda \in i\mathfrak{t}^*).$$

Thus, s_α maps α to $-\alpha$ and is the identity on the hyperplane $H_\alpha^0 := \{\lambda \mid \lambda(H_\alpha) = 0\}$. Since s_α is orthogonal it follows that $H_\alpha^0 = \alpha^\perp$ and that s_α is the orthogonal reflection in the hyperplane $\alpha^\perp$. The reflection $s_\alpha \in \mathrm{End}(i\mathfrak{t}^*)$ is therefore also given by the formula

$$s_\alpha\lambda = \lambda - 2\frac{\langle\lambda\,,\,\alpha\rangle}{\langle\alpha\,,\,\alpha\rangle}\alpha, \qquad (\lambda \in i\mathfrak{t}^*).$$

Comparing this formula with the previous one we see that $j(H_\alpha) = 2\alpha/\langle\alpha\,,\,\alpha\rangle$.

Lemma 36.7 *Let $\alpha \in R$. There exists an automorphism φ of $\mathfrak{g}_{\mathbb{C}}$, which leaves $\mathfrak{t}_{\mathbb{C}}$ invariant and has restriction s_α to this space. The induced endomorphism s_α of $\mathfrak{t}_{\mathbb{C}}^*$ leaves R invariant. Moreover, if $\alpha \in R$ then $\beta - s_\alpha(\beta) \in \mathbb{Z}\alpha$.*

Proof: We fix $X_\alpha, Y_\alpha \in \mathfrak{s}_\alpha$ such that $H_\alpha, X_\alpha, Y_\alpha$ is a standard triple. Moreover, we put

$$U_\alpha = \frac{\pi}{2}(X_\alpha - Y_\alpha)$$

and $\varphi := e^{\,\mathrm{ad} U_\alpha}$. Then φ is an automorphism of $\mathfrak{g}_{\mathbb{C}}$. Since $\mathrm{ad} U_\alpha$ annihilates every element of the subset $\ker\alpha$ it follows that $\varphi = I$ on $\ker\alpha$. On the other hand, we claim that $\varphi(H_\alpha) = -H_\alpha$.

To establish the claim, we observe that the identity is entirely formulated in terms of the structure of the Lie algebra $\mathfrak{s}_\alpha$. By isomorphism it suffices to show the similar identity in case $H_\alpha, X_\alpha, Y_\alpha$ is the usual standard triple in $\mathfrak{sl}(2, \mathbb{C})$. The advantage in that situation is that we can use computations in the group $\mathrm{SL}(2, \mathbb{C})$. In fact, we have

$$U_\alpha = \frac{\pi}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

hence

$$\exp U_\alpha = \begin{pmatrix} \cos\frac{\pi}{2} & \sin\frac{\pi}{2} \\ -\sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

from which we see that $e^{\,\mathrm{ad} U_\alpha} H_\alpha = \mathrm{Ad}(\exp(U_\alpha)) H_\alpha = \exp(U_\alpha) H_\alpha \exp(U_\alpha)^{-1} = -H_\alpha$. This establishes the claim.

We conclude that $\varphi \in \mathrm{Aut}(\mathfrak{g}_{\mathbb{C}})$, $\varphi(\mathfrak{t}_{\mathbb{C}}) = \mathfrak{t}_{\mathbb{C}}$ and $\varphi|_{\mathfrak{t}_{\mathbb{C}}} = s_\alpha$. It follows from Lemma 32.3 that the induced map $s_\alpha \in \mathrm{GL}(\mathfrak{t}_{\mathbb{C}}^*)$ maps R to itself.

Finally, let $\beta \in R$. Then

$$V = \sum_{k \in \mathbb{Z}} \mathfrak{g}_{\mathbb{C}(\beta + k\alpha)}$$

is readily seen to be an $\mathrm{ad}(\mathfrak{s}_\alpha)$-invariant subspace. Since $U_\alpha \in \mathfrak{s}_\alpha$ it follows that V is $\mathrm{ad}(U_\alpha)$-invariant, and since V is closed, it follows that φ leaves V invariant. Therefore,

$$\mathfrak{g}_{\mathbb{C}s_\alpha(\beta)} = \mathfrak{g}_{\mathbb{C}\varphi(\beta)} = \varphi(\mathfrak{g}_{\mathbb{C}\beta}) \subset V$$

and we conclude that $s_\alpha(\beta) \in \beta + \mathbb{Z}\alpha$. $\square$

Definition 36.8 The subgroup $W = W(\mathfrak{g}, \mathfrak{t})$ of $\mathrm{GL}(i\,\mathfrak{t}^*)$ generated by the reflections s_α, for $\alpha \in R$, is called the *Weyl group* of $(\mathfrak{g}, \mathfrak{t})$.

Lemma 36.9 *The Weyl group W is finite.*

Proof: By Lemma 36.7, each reflection s_α leaves R invariant. Hence $w(R) \subset R$ for each $w \in W$. Since w is injective and R finite, it follows that $w|_R$ belongs to the group $\mathrm{Sym}(R)$ of bijections from R onto itself. Clearly the restriction map $r : w \mapsto w|_R$, $W \to \mathrm{Sym}(R)$, is a group homomorphism. Moreover, since R spans $\mathfrak{t}^*$, by Lemma 31.9, it follows that $\ker r$ is trivial. Hence $\#W \leq \#\mathrm{Sym}(R) < \infty$. $\square$

Let E be a finite dimensional linear space. If $\alpha \in E \setminus \{0\}$ then by a *reflection* in α we mean a linear map $s : E \to E$ with $s(\alpha) = -\alpha$ and

$$E = \mathbb{R}\alpha \oplus \ker(s - I).$$

Note that any reflection satisfies $s^2 = I$. Hence $s \in \mathrm{GL}(E)$ and $s^{-1} = s$.

Lemma 36.10 *Let E be a finite dimensional real linear space, and $R \subset E$ a finite subset that spans E. Then for every $\alpha \in R$ there is at most one reflection s in α such that $s(R) = R$.*

Proof: Let K be the group of $A \in \mathrm{GL}(E)$ with $A(R) = R$. The restriction map $r : A \mapsto A|_R$ is a group homomorphism from K to the group of bijections of R. Moreover, r has trivial kernel, since R spans E. It follows that K is a finite group. Hence, there exists an inner product on E for which K acts by orthogonal transformations (use averaging). If s is any reflection in a non-zero element α of E which preserves R, then it must be an orthogonal transformation, hence the orthogonal reflection in the hyperplane $\alpha^\perp$. In particular, there exists at most one such reflection. $\square$

Definition 36.11 A (general) *root system* is a pair (E, R) consisting of a finite dimensional real linear space E and a finite subset $R \subset E \setminus \{0\}$ such that the following conditions are fulfilled.

(a) R spans E.

(b) If $\alpha \in R$, then $R \cap \mathbb{R}\alpha = \{-\alpha, \alpha\}$.

(c) If $\alpha \in R$ then there exists a (necessarily unique) reflection s_α in α that maps R to itself.

(d) If $\alpha, \beta \in R$ then $s_\alpha(\beta) \in \beta + \mathbb{Z}\alpha$.

According to the results of this section, the pair consisting of $E = i\mathfrak{t}^*$ and $R = R(\mathfrak{g}, \mathfrak{t})$ is a root system in the sense of the above definition.

For a general root system, the subgroup W of $\mathrm{GL}(E)$, generated by the reflections s_α, for $\alpha \in R$, is called the *Weyl group* of the root system (E, R). By the same proof as that of Lemma 36.9, it follows that W is finite. By averaging we see that E may be equipped with a positive definite inner product $\langle \cdot\, , \, \cdot \rangle$ that is W-invariant. It follows that each reflection s_α, for $\alpha \in R$ is orthogonal relative to $\langle \cdot\, , \, \cdot \rangle$. Hence, it is given by the formula

$$s_\alpha(\lambda) = \lambda - 2\frac{\langle \lambda\, , \, \alpha \rangle}{\langle \alpha\, , \, \alpha \rangle}\alpha, \qquad (\lambda \in E). \tag{55}$$

In terms of the inner product the condition (d) in the definition of root system may therefore be rephrased as

$$2\frac{\langle \beta\, , \, \alpha \rangle}{\langle \alpha\, , \, \alpha \rangle} \in \mathbb{Z} \qquad (\alpha, \beta \in R).$$

Two root systems (E, R) and (E', R') are called *isomorphic* if there exists a linear isomorphism $T : E \to E'$ with $T(R) = R'$. If $\mathfrak{g}$ is a compact semisimple Lie algebra, then it follows from Lemmas 32.2 and 32.3 that the isomorphism class of the root system $R(\mathfrak{g}, \mathfrak{t})$ is independent of the choice of the maximal torus $\mathfrak{t}$.

We now have the following result, which we state without proof. It reduces the classification of all compact semisimple Lie algebras to the classification of all root systems.

Theorem 36.12 *The map $\mathfrak{g} \to R(\mathfrak{g}_{\mathbb{C}}, \mathfrak{t}_{\mathbb{C}})$ induces a map from (a) the isomorphism classes of real Lie algebras with negative definite Killing form to (b) the isomorphism classes of root systems. This map is a bijection.*

37 Weyl's formulas

We retain the notation of the previous section. In this section we will describe the classification of of all irreducible representations of the compact semisimple Lie algebra $\mathfrak{g}$. Moreover, in terms of this classification we will state the beautiful character and dimension formulas due to Hermann Weyl.

The *weight lattice* $\Lambda = \Lambda(\mathfrak{g}, \mathfrak{t})$ of the pair $(\mathfrak{g}, \mathfrak{t})$ is defined as the set

$$\Lambda = \{\lambda \in i\mathfrak{t}^* \mid \forall \alpha \in R : s_\alpha \lambda \in \lambda + \mathbb{Z}\alpha\}.$$

Equip $i\mathfrak{t}^*$ with any W-invariant positive definite inner product $\langle \cdot\,,\,\cdot \rangle$. Then from (55) we see that, alternatively, $\Lambda(\mathfrak{g}, \mathfrak{t})$ may be defined as the set of elements $\lambda \in i\mathfrak{t}^*$ such that

$$2\frac{\langle \lambda\,,\,\alpha \rangle}{\langle \alpha\,,\,\alpha \rangle} \in \mathbb{Z} \qquad \text{for all} \qquad \alpha \in R.$$

It follows from the definition of root system that the $\mathbb{Z}$-lattice spanned by R is contained in Λ.

The collection of *dominant weights* (relative to R^+) is defined by

$$\Lambda^+ = \{\lambda \in i\mathfrak{t}^* \mid \forall \alpha \in R^+ :\ s_\alpha \lambda \in \lambda + \mathbb{N}(-\alpha)\}.$$

Thus, Λ^+ consists of the collection of weights in Λ that are contained in the convex cone

$$C^+ = \{\lambda \in i\mathfrak{t}^* \mid \langle \lambda\,,\,\alpha \rangle \geq 0 \quad \text{for all} \quad \alpha \in R^+\}.$$

The following results amount to the classification of all irreducible representations of $\mathfrak{g}$.

Theorem 37.1 *For every $\lambda \in \Lambda^+$ there exists a unique (up to equivalence) irreducible representation π_λ of $\mathfrak{g}$ with highest weight λ.*

From this result combined with Theorem 31.23 we obtain the following.

Corollary 37.2 *The map $\lambda \mapsto \pi_\lambda$ induces a bijection from Λ^+ onto the collection of equivalence classes of irreducible representations of $\mathfrak{g}$.*

Let now G be a compact connected Lie group with algebra $\mathfrak{g}$. Let π be a finite dimensional irreducible representation of G. Then the associated representation of the Lie algebra is equivalent to π_λ for a unique $\lambda \in \Lambda^+$. It turns out that in terms of this parametrization, there exist beautiful formulas for the character and dimension of π. The character χ_π of π is conjugation invariant. In view of the following result, which we state without proof, it is completely determined by its restriction to $T := \exp(\mathfrak{t})$.

Proposition 37.3 *The group $T = \exp(\mathfrak{t})$ is a compact torus in G. Moreover, each element of G is conjugate to an element of T.*

If $w \in W$ we write $\epsilon(w) = \det(w)$ for the determinant of w; since w is orthogonal with respect to a suitable inner product, we have $\epsilon(w) = \pm 1$. We define the element $\delta \in i\mathfrak{t}^*$ by

$$\delta = \frac{1}{2} \sum_{\alpha \in R^+} \alpha.$$

Theorem 37.4 *(Weyl's formulas). Let π be an irreducible representation of G of highest weight λ. Then the character χ_π is given by*

$$\chi_\pi(\exp X) = \frac{\sum_{w \in W} \epsilon(w) e^{w(\lambda+\delta)(X)}}{\sum_{w \in W} \epsilon(w) e^{w\delta(X)}},$$

for all $X \in \mathfrak{t}$ for which the denominator is non-zero; these X form an open dense subset (Weyl's character formula). Moreover, the dimension of π is given by

$$\dim \pi = \prod_{\alpha \in R^+} \frac{\langle \lambda + \delta \,,\, \alpha \rangle}{\langle \delta \,,\, \alpha \rangle}$$

(Weyl's dimension formula).

Example 37.5 We consider the example of $\mathfrak{g} = \mathfrak{su}(2)$, with the associated standard triple H, X, Y in $\mathfrak{g}_{\mathbb{C}} = \mathfrak{sl}(2, \mathbb{C})$. We recall that $\mathfrak{t} = i\mathbb{R}H$ is a maximal torus in $\mathfrak{g}$. Moreover, $R = \{-\alpha, \alpha\}$, where $\alpha \in i\mathfrak{t}^*$ is determined by $\alpha(H) = 2$. Also, $R^+ = \{\alpha\}$ is a choice of positive roots. The associated Weyl group consists of two elements, 1 and s_α. Moreover, $\delta = \frac{1}{2}\alpha$. Hence,

$$\Lambda^+ = \{n\delta \mid n \in \mathbb{N}\}.$$

The representation with highest weight $n\delta$ was earlier denoted by π_n. We note that $(n\delta + \delta)(H) = n + 1$ and $[s_\alpha(n\delta + \delta)](H) = -(n + 1)$. According to the above formula the character of π_n is therefore given by the formula

$$\pi_n(\exp itH) = \frac{e^{i(n+1)t} - e^{-i(n+1)t}}{e^{it} - e^{-it}}$$

which is consistent with what we computed earlier. The dimension of π_n is given by

$$\dim(\pi_n) = \frac{\langle n\delta + \delta \,,\, \alpha \rangle}{\langle \delta \,,\, \alpha \rangle} = n + 1,$$

consistent with what we discussed before.

38 The classification of root systems

38.1 Cartan integers

In this section we shall study some aspects of the theory of root systems. In particular we shall describe the first step towards their classification. The starting point of the theory is the definition of a root system as given in Definition 36.11. In the rest of this section we assume that (E, R) is such a root system. The dimension of E is called the *rank* of the root system.

By the process of averaging over the Weyl group W of the given root system, we select a W-invariant positive definite inner product $\langle \cdot\, ,\ \cdot \rangle$ om E. Then, for every $\alpha \in R$ the reflection s_α is given by the following formula, for $\lambda \in E$,

$$s_\alpha(\lambda) = \lambda - 2\frac{\langle \lambda\, ,\, \alpha \rangle}{\langle \alpha\, ,\, \alpha \rangle}\alpha.$$

For two roots $\alpha, \beta \in R$ we define $n_{\alpha\beta}$ to be the integer determined by

$$s_\alpha(\beta) = \beta - n_{\alpha\beta}\alpha, \tag{56}$$

see Definition 36.11 (d). These integers are called the *Cartan integers* for the root system. It follows from the above representation of the reflection in terms of the inner product that the Cartan integers are alternatively given by

$$n_{\alpha\beta} = 2\frac{\langle \beta\, ,\, \alpha \rangle}{\langle \alpha\, ,\, \alpha \rangle}. \tag{57}$$

Lemma 38.1 *Let φ be an isomorphism from (E, R) onto a second root system (E', R'). Then, for all $\alpha, \beta \in R$,*

(a) $\varphi \circ s_\alpha = s_{\varphi(\alpha)} \circ \varphi$;

(b) $n_{\varphi(\alpha)\,\varphi(\beta)} = n_{\alpha\beta}$.

Proof: It is readily seen that $s := \varphi \circ s_\alpha \circ \varphi^{-1} : E' \to E$ is a reflection in $\varphi(\alpha)$. Since $s(R') = \varphi s_\alpha \varphi^{-1}(\varphi R) = R'$, (a) follows. Assertion (b) follows by application of (56). $\square$

We shall now discuss the possible values of the Cartan integers. If $\alpha, \beta \in E \setminus \{0\}$, then by the Cauchy-Schwarz inequality there is a unique $\varphi_{\alpha\beta} \in [0, \pi]$ such that

$$\langle \alpha\, ,\, \beta \rangle = |\alpha||\beta| \cos \varphi_{\alpha\beta}.$$

The number $\varphi_{\alpha\beta}$ is called the *angle* between α and β (with respect to the given inner product).

Assume that $\alpha, \beta \in R$ and $\alpha \neq \pm\beta$. Then

$$2\frac{|\beta|}{|\alpha|} \cos \varphi_{\alpha\beta} = 2\frac{\langle \beta\, ,\, \alpha \rangle}{\langle \alpha\, ,\, \alpha \rangle} = n_{\alpha\beta} \in \mathbb{Z}.$$

It follows that

$$n_{\alpha\beta} n_{\beta\alpha} = 4\cos^2 \varphi_{\alpha\beta} \in \mathbb{Z}.$$

From this formula we see that the value of $\varphi_{\alpha\beta}$ is independent of the particular choice of W-invariant inner product on E. By Definition 36.11(b) the roots α, β are not proportional, hence $|\cos \varphi_{\alpha\beta}| < 1$. It follows that

$$n_{\alpha\beta} n_{\beta\alpha} \in \{0, 1, 2, 3\}.$$

After renaming we may assume that $|\alpha| \leq |\beta|$. It then follows from (57) that $|n_{\alpha\beta}| \geq |n_{\beta\alpha}|$. By integrality of the Cartan integers we find that either $\alpha \perp \beta$ or $n_{\beta\alpha} = \pm 1$. This leads to the following table of possibilities for $n_{\alpha\beta}$ and $\varphi_{\alpha\beta}$.

Lemma 38.2 *Let $\alpha, \beta \in R$ be non-proportional roots with $|\alpha| \leq |\beta|$. Then the following table contains all possible combinations of values of $n_{\alpha\beta}$, $n_{\beta\alpha}$ and $\varphi_{\alpha\beta}$. The question mark indicates that the value involved is undetermined.*

$n_{\alpha\beta} n_{\beta\alpha}$	$n_{\alpha\beta}$	$n_{\beta\alpha}$	$\cos \varphi_{\alpha\beta}$	$\varphi_{\alpha\beta}$	$\frac{\lvert\beta\rvert^2}{\lvert\alpha\rvert^2} = \frac{n_{\alpha\beta}}{n_{\beta\alpha}}$
0	0	0	0	$\frac{\pi}{2}$	?
1	1	1	$\frac{1}{2}$	$\frac{\pi}{3}$	1
1	-1	-1	$-\frac{1}{2}$	$\frac{2\pi}{3}$	1
2	2	1	$\frac{1}{2}\sqrt{2}$	$\frac{\pi}{4}$	2
2	-2	-1	$-\frac{1}{2}\sqrt{2}$	$\frac{3\pi}{4}$	2
3	3	1	$\frac{1}{2}\sqrt{3}$	$\frac{\pi}{6}$	3
3	-3	-1	$-\frac{1}{2}\sqrt{3}$	$\frac{5\pi}{6}$	3

Example 38.3 Let $E = \mathbb{R}^2$, equipped with the standard inner product. Let α be the first standard basis vector $(1, 0)$, and $\beta = (-\frac{1}{2}, \frac{1}{2}\sqrt{3})$. Then $|\beta| = |\alpha| = 1$ and $\varphi_{\alpha\beta} = 2\pi/3$. Moreover, $\alpha + \beta = (\frac{1}{2}, \frac{1}{2}\sqrt{3})$ has angle $\pi/3$ with both α and β. It is easily verified that $R = \{\pm\alpha, \pm\beta, \pm(\alpha + \beta)\}$ is a root system. Note that $n_{\alpha\beta} = n_{\beta\alpha} = -1$. This root system, called A_2, is depicted in the illustration following Lemma 38.16. Let $r = s_{\alpha+\beta} s_\alpha$. Then r is the rotation over angle $2\pi/3$. The reflection $s = s_\alpha$ is the reflection in the line $\alpha^\perp = \mathbb{R}(0, 1)$. The Weyl group W equals $\{1, r, r^{-1}, s, sr, sr^{-1}\}$.

The following lemma will be extremely useful in the further development of the theory.

Lemma 38.4 *Let $\alpha, \beta \in R$ be non-proportional roots.*

(a) *If $\langle \alpha \,,\, \beta \rangle > 0$ then $\alpha - \beta \in R$.*

(b) *If* $\langle \alpha , \beta \rangle < 0$ *then* $\alpha + \beta \in R$.

Proof: It suffices to establish (a). Then (b) follows by replacing β with $-\beta$. Since $\alpha - \beta \in R$ is equivalent to $\beta - \alpha \in R$ we may as well assume that $|\alpha| \le |\beta|$. Then it follows that $0 < n_{\beta\alpha} \le n_{\alpha\beta}$, hence $n_{\beta\alpha} = 1$. In view of (56) this implies that $s_\beta(\alpha) = \alpha - \beta$. Now use Definition 36.11 to conclude that $\alpha - \beta \in R$. $\square$

Given non-proportional roots $\alpha, \beta \in R$ we define the α-string through β to be the set

$$L_\alpha(\beta) := (\beta + \mathbb{Z}\alpha) \cap R.$$

The following lemma expresses that root strings have no interruptions and have at most 4 elements.

Lemma 38.5 *Let* $\alpha, \beta \in R$ *be non-proportional.*

(a) *There exist unique* $p, q \in \mathbb{Z}$ *with* $p \le q$ *such that* $L_\alpha(\beta) = \{\beta + k\alpha \mid p \le k \le q\}$.

(b) $p \le 0 \le q$ *and* $p + q = -n_{\alpha\beta}$.

(c) $\# L_\alpha(\beta) \le 4$.

Proof: We first establish (a). Write $\lambda_j := \beta - j\alpha$, for $j \in \mathbb{Z}$. Assume (a) does not hold. Then there exist integers $k < l$ such that $\lambda_k, \lambda_l \in R$ but $\lambda_{k+1}, \lambda_{l-1} \notin R$. It follows by application of Lemma 38.4 that

$$\langle \lambda_k , \alpha \rangle \ge 0 \quad \text{and} \quad \langle \lambda_l , \alpha \rangle \le 0.$$

On the other hand,

$$\langle \lambda_k , \alpha \rangle = \langle \beta , \alpha \rangle + k|\alpha|^2 < \langle \beta , \alpha \rangle + l|\alpha|^2 = \langle \lambda_l , \alpha \rangle,$$

contradiction. We conclude that (a) holds. Since $\beta \in L_\alpha(\beta)$, the first assertion of (b) follows. For the other assertion, we note that s_α maps $L_\alpha(\beta)$ bijectively onto itself. Hence $s_\alpha(\beta + p\alpha) = \beta + q\alpha$, from which it follows that $-n_{\alpha\beta}\alpha - p\alpha = q\alpha$. This establishes (b).

For (c) we note that $\gamma = \beta + p\alpha$ is a root. Clearly, $L_\alpha(\beta) = L_\alpha(\gamma) = \{\gamma + j\alpha \mid 0 \le j \le q - p\}$, so that $\# L_\alpha(\gamma) = q - p + 1$. It now follows from (b) applied to the pair α, γ that $q - p = -n_{\alpha\gamma}$, hence $n_{\alpha\gamma} \le 0$ and $q - p \in \{0, 1, 2, 3\}$. $\square$

38.2 Fundamental and positive systems

If F is a finite subset of E we write

$$\mathbb{N}F = \{\sum_{f \in F} n_f f \mid n_f \in \mathbb{N}\}.$$

Here $\mathbb{N} = \{0, 1, 2, \ldots\}$.

Definition 38.6 A *fundamental system* or *basis* for (E, R) is a subset $S \subset R$ such that

(a) S is a basis for E;

(b) $R \subset \mathbb{N}S \cup \mathbb{N}(-S)$.

Conditions (a) and (b) of the above definition may be restated as follows. Every root $\beta \in R$ admits a unique expression of the form

$$\beta = \sum_{\alpha \in S} k_\alpha \alpha,$$

with $k_\alpha \in \mathbb{R}$. Moreover, either $k_\alpha \in \mathbb{N}$ for all $\alpha \in S$ or $k_\alpha \in -\mathbb{N}$ for all $\alpha \in S$. The *height* of β relative to S is defined by

$$\mathrm{ht}(\beta) = \sum_{\alpha \in S} k_\alpha$$

Lemma 38.7 *Let S be a fundamental system for the root system R. Then for all roots $\alpha, \beta \in S$ with $\alpha \neq \beta$ one has $\langle \alpha \, , \, \beta \rangle \leq 0$ (or, equivalently, $\varphi_{\alpha\beta} \geq \pi/2$).*

Proof: Since $\alpha - \beta$ is a linear combination of the elements of S with both plus and minus signs, it cannot be a root. It follows from Lemma 38.4 that $\langle \alpha \, , \, \beta \rangle \leq 0$. $\square$

To ensure the existence of fundamental systems, we introduce the notion of a positive system. By an *open half space* of E we mean a set of the form $E^+(\xi) = \{x \in E \mid \xi(x) > 0\}$, where ξ is a non-zero element of the dual space $E^* := \mathrm{Hom}_{\mathbb{R}}(E, \mathbb{R})$. Via the given inner product, we sometimes identify E with E^*. Accordingly, if $\gamma \in E \setminus \{0\}$, we write

$$E^+(\gamma) = \{x \in E \mid \langle x \, , \, \gamma \rangle > 0\}.$$

Definition 38.8 A *positive system* or *choice of positive roots* for R is a subset $P \subset R$ with the following properties.

(a) There exists an open half space containing P.

(b) $R \subset P \cup (-P)$.

Let P be a positive system for R. An *indecomposable* or *simple root* in P is defined to be a root that cannot be written as the sum of two roots from P. The set of these simple roots is denoted by $S(P)$.

Lemma 38.9 *Let P be a positive system for R. Then $S(P)$ is a fundamental system for R and $P = \mathbb{N}S(P) \cap R$. The map $P \mapsto S(P)$ is a bijective map from the collection $\mathcal{P}$ of positive systems for R onto the collection $\mathcal{S}$ of fundamental systems for R.*

Proof: Put $S = S(P)$. Let $\alpha \in P$. Then either $\alpha \in S$, or α can be written as a sum $\beta + \gamma$ with $\beta, \gamma \in P$. Proceeding in this way we see that $P \subset \mathbb{N}S$, hence condition (b) of Definition 38.6 holds. In particular, it follows that S spans E. It remains to be shown that the elements of S are linearly independent. Let α, β be distinct roots in S. By definition of S, neither $\alpha - \beta$ nor $\beta - \alpha$ does belong to P. Hence, $\alpha - \beta \notin R$. It follows by application of Lemma 38.4 that $\langle \alpha\, , \, \beta \rangle \leq 0$. From the lemma below it now follows that the elements of S are linearly independent. Hence S is a fundamental system. In the above we established $P \subset \mathbb{N}S$, whence $-P \subset \mathbb{N}(-S)$ and since $R = P \cup (-P)$ it follows that $P = \mathbb{N}S \cap R$.

We have shown that the map $P \mapsto S(P)$ is injective $\mathcal{P} \to \mathcal{S}$ and will finish the proof by establishing its surjectivity. For S a fundamental system of R we define $R^+ = R^+(S) = \mathbb{N}S \cap R$. Since S is a basis for E, the linear functionals $\langle \alpha\, , \, \cdot \rangle$, for $\alpha \in S$, form a basis for the dual space E^*. It follows that there exists a $\gamma \in E$ such that $\langle \alpha\, , \, \gamma \rangle > 0$ for all $\alpha \in S$. We conclude that S, hence R^+, is contained in a half space. From $R \subset \mathbb{N}S \cup (-\mathbb{N}S)$ it follows that $R \subset R^+ \cup (-R^+)$. Hence $R^+(S)$ is a positive system. From $S \subset R^+(S) \subset \mathbb{N}S$ it follows that $S(R^+(S)) \subset S$. Both sets of of this inclusion are bases for E; hence, they must be equal. We conclude that the map $R \mapsto S(P)$, $\mathcal{P} \to \mathcal{S}$ is bijective with inverse $S \mapsto R^+(S)$. $\square$

Lemma 38.10 *Let E be a finite dimensional real linear space, equipped with a positive definite inner product. Let $S \subset E$ be a finite subset contained in a fixed open half space and such that $\langle \alpha\, , \, \beta \rangle \leq 0$ for all distinct $\alpha, \beta \in S$. Then the elements of S are linearly independent.*

Proof: There exists a $\xi \in E$ such that $\langle \xi\, , \, \alpha \rangle > 0$ for all $\alpha \in S$. Let $\lambda_\alpha \in \mathbb{R}$, for $\alpha \in S$, be such that $\sum_{\alpha \in S} \lambda_\alpha \alpha = 0$. We define $S_\pm := \{\alpha \in S \mid \pm\lambda_\alpha > 0\}$. Then S_+ and S_- are disjoint. We define $\epsilon_\pm := \sum_{\alpha \in S_\pm} |\lambda_\alpha|\alpha$. If one of the sets of summation is empty, the sum is understood to be zero. The linear relation between the elements of S may now be expressed as $\epsilon_+ - \epsilon_- = 0$ or

$$\epsilon_+ = \epsilon_-.$$

From the fact that $\langle \alpha\, , \, \beta \rangle \leq 0$ for all $\alpha \in S_+$ and all $\beta \in S_-$ it now follows that $\langle \epsilon_+\, , \, \epsilon_+ \rangle = \langle \epsilon_+\, , \, \epsilon_- \rangle \leq 0$. Hence, $\langle \epsilon_+\, , \, \epsilon_+ \rangle = 0$ and we conclude that $\epsilon_+ = \epsilon_- = 0$. We now observe that $0 = \langle \xi\, , \, \epsilon_+ \rangle = \sum_{\alpha \in S_+} |\lambda_\alpha| \langle \xi\, , \, \alpha \rangle$. Since each of the inner products $\langle \xi\, , \, \alpha \rangle$ is strictly positive, it follows that $S_+ = \emptyset$. Similarly, $S_- = \emptyset$ and we conclude that $\lambda_\alpha = 0$ for all $\alpha \in S$. This establishes the linear independence. $\square$

For each $\alpha \in R$, the set $P_\alpha := \ker(I - s_\alpha)$ is called the root hyperplane associated with α. Relative to the given W-invariant inner product, $P_\alpha = \alpha^\perp$. We define the set of regular points in E by

$$E^{\text{reg}} := E \setminus \cup_{\alpha \in R} P_\alpha. \tag{58}$$

Since R is a finite set, it is easy to show that E^{reg} is an open dense subset of E; in particular, the set of regular points is non-empty.

We can now establish the existence of positive systems, hence also of fundamental systems. For $\gamma \in E^{\text{reg}}$ we define

$$R^+(\gamma) = R \cap E^+(\gamma) = \{\alpha \in R \mid \langle \gamma\, , \, \alpha \rangle > 0\}.$$

Lemma 38.11 *For every $\gamma \in E^{\text{reg}}$ the set $R^+(\gamma)$ is a positive system for R. Moreover, every positive system arises in this way.*

Proof: That $R^+(\gamma)$ is a positive system is immediate from the definitions. Conversely, let P be a positive system for R, and let $S = S(P)$ be the associated fundamental system. The linear functionals $\langle \cdot\, , \alpha\rangle$, for $\alpha \in S$, form a basis for the dual space E^*. Hence there exists a $\gamma \in E$ such that $\langle \gamma\, , \alpha\rangle > 0$ for all $\alpha \in S$. From $P \subset \mathbb{N}S$ it now follows that $P \subset R^+(\gamma)$, hence also $[-P] \cap R^+(\gamma) = \emptyset$. Since P is a positive system, $R^+(\gamma) \subset P \cup (-P)$, so we must have that $P = R^+(\gamma)$. □

Definition 38.12 Let S be a fundamental system for R. The integers $n_{\alpha\beta}$, for $\alpha, \beta \in S$, are called the *Cartan integers* associated with S. The square matrix $n : S \times S \to \mathbb{Z}$, $(\alpha, \beta) \mapsto n_{\alpha\beta}$ is called the *Cartan matrix* for S.

We will end this section with a result that asserts that every root system is completely determined by the Cartan matrix of a fundamental system. It depends crucially on the following lemma and Lemma 38.5.

Lemma 38.13 *Let S be a fundamental system for R and $R^+ = R \cap \mathbb{N}S$ the associated positive system. If $\beta \in R^+ \setminus S$, then there exists an $\alpha \in S$ such that $\beta - \alpha \in R^+$.*

Proof: Since β is not simple, it is of the form $\sum_{\gamma \in S} k_\gamma \gamma$, $k_\gamma \in \mathbb{N}$, with at least two coefficients non-zero. Thus, if $\alpha \in S$ and $\beta - \alpha \in R$, then at least one of the coefficients of $\beta - \alpha$ is still positive, and it follows that $\beta - \alpha \in R^+$.

Assume that $\beta - \alpha \notin R^+$ for all $\alpha \in S$. Then it follows that $\beta - \alpha \notin R$ for all $\alpha \in S$. By Lemma 38.4 this implies that $\langle \beta\, , \alpha\rangle \le 0$ for all $\alpha \in S$, hence $\langle \beta\, , \beta\rangle \le 0$, hence $\beta = 0$, contradiction. □

Given any finite set S we write E_S for the real linear space with basis S. As a concrete model we may take the space $\mathbb{R}^S$ of functions $S \to \mathbb{R}$; here S is embedded in $\mathbb{R}^S$ by identifying an element $\alpha \in S$ with the function $\delta_\alpha : S \to \mathbb{R}$ given by $\beta \mapsto \delta_{\alpha\beta}$. If $v \in E_S$, we put $v = \sum_{\alpha \in S} v_\alpha \alpha$. With the above identification, as an element of $\mathbb{R}^S$, the vector v is given by $\alpha \mapsto v_\alpha$.

Let E be a real linear space and $f : S \to E$ a map, then f has a unique extension to a linear map $E_S \to E$, again denoted by f. Moreover, if $f : S \to S'$ is a map, then f may be viewed as a map $S \to E_{S'}$ which in turn has a unique linear extension to a map $f : E_S \to E_{S'}$.

Theorem 38.14 *There exists a map $\mathcal{R}$ assigning to every pair consisting of a finite set S and a function $n : S \times S \to \mathbb{Z}$ a finite subset $\mathcal{R}(S, n) \subset E_S$ with the following properties.*

(a) *If $\varphi : S' \to S$ is a bijection of finite sets, and $n : S \times S \to \mathbb{Z}$ a function, then the induced map $\varphi : E_{S'} \to E_S$ maps $\mathcal{R}(S', \varphi^* n)$ bijectively onto $\mathcal{R}(S, n)$.*

(b) *If (E, R) is a root system with fundamental system S and Cartan matrix $n : S \times S \to \mathbb{Z}$, then the natural map $E_S \to E$ maps $\mathcal{R}(S, n)$ bijectively onto R. In particular, $(\mathbb{R}^S, \mathcal{R}(S, n))$ is a root system isomorphic to (E, R).*

Remark 38.15 The above result guarantees that the isomorphism class of a root system can be retrieved from the Cartan matrix of a fundamental system. Later we will see that all fundamental systems are conjugate under the Weyl group, so that all Cartan matrices of a given root system are essentially equal, cf. Lemma 38.1.

In the proof of the above result the set $\mathcal{R}$ will be defined by means of a recursive algorithm with input data S, n. This algorithm will provide us with a finite procedure for finding all root systems of a given rank. Let such a rank r be fixed. Let S be a given set with r elements. Each root system R of rank r can be realized in the linear space E_S, having the standard basis as fundamental system.

The possible Cartan matrices run over the finite set of maps $S \times S \to \{0, \pm 1, \pm 2, \pm 3\}$. For each such map n it can be checked whether or not $(E_S, \mathcal{R}(S, n))$ is a root system with fundamental system S. Condition (b) guarantees that all root systems of rank r are obtained in this way.

Proof: We shall describe the map $\mathcal{R}$ and then show that it satisfies the requirements. Requirement (b) is motivational for the definition.

For each $\alpha \in S$ we define the map $n_\alpha : S \to \mathbb{Z}$ by $n_\alpha(\beta) = n(\alpha, \beta)$. As said above this map induces a linear map $n_\alpha : E_S \to \mathbb{R}$. If the linear maps n_α, for $\alpha \in S$, are linearly dependent, we define $\mathcal{R}(S, n) = \emptyset$ (we need not proceed, since n can impossibly be a Cartan matrix of a root system). Thus, assume that the n_α are linearly independent linear functionals.

We consider the semi-lattice $\Lambda = \mathbb{N}S \subset E_S$. Then for each $\alpha \in S$ the map n_α has integral values on Λ. We define a height function on Λ in an obvious manner,

$$\mathrm{ht}(\lambda) = \sum_{\alpha \in S} \lambda_\alpha.$$

Let Λ_k be the finite set of $\lambda \in \Lambda$ with $\mathrm{ht}(\lambda) = k$. We put $P_1 = S$ and more generally will define sets $P_k \subset \Lambda_k$ by induction on k.

Let $P_1, \ldots, P_k$ be given, then P_{k+1} is defined as the subset of Λ_{k+1} consisting of elements that can be expressed in the form $\beta + \alpha$ with $(\alpha, \beta) \in S \times P_k$ satisfying the following conditions.

(i) α and β are not proportional.

(ii) $|n_\alpha(\beta + \alpha)| \leq 3$.

(iii) Let p be the smallest integer such that $\beta + p\alpha \in P_1 \cup \cdots \cup P_k$; then $p - n_\alpha(\beta) > 0$.

We define $\mathcal{P}(S, n)$ to be the union of the sets P_k, for $k \geq 1$ and put $\mathcal{R}(S, n) = \mathcal{P}(S, n) \cup (-\mathcal{P}(S, n))$.

The set F of $\beta \in E_S$ with $n_\alpha(\beta) \in \{0, \pm 1, \pm 2, \pm 3\}$ for all $\alpha \in S$ is finite, because the n_α are linearly independent functionals. In fact, $\#F \leq (\#S)^7$. From the above construction it follows that $\mathcal{R}(S, n) \subset F$, hence is finite. In particular, we see that the above inductive definition starts producing empty sets at some level. In fact, let N be an upper bound for the height function on F, then $\mathcal{P}(S, n) = P_1 \cup \cdots \cup P_N$.

From the definition it is readily seen that the map $\mathcal{R}$ defined above satisfies condition (a) of the theorem. We will finish the proof by showing that condition (b) holds. Sssume that S is a fundamental system for a root system (E, R). Let $R^+ = R \cap \mathbb{N}S$ be the associated positive system and $n : S \times S \to \mathbb{Z}$ the associated Cartan matrix. The inclusion map $S \subset E$ induces a linear isomorphism $E_S \to E$ via which we shall identify. Then it suffices to show that $\mathcal{R}(S, n) = R$. Since n is a genuine Cartan matrix, the functionals n_α, for $\alpha \in S$ are linearly independent. Thus it suffices to show that $P_k = R \cap \Lambda_k$, for every $k \in \mathbb{N}$. We will do this by induction on k. For $k = 1$ we have $R \cap \Lambda_k = S = P_1$, and the statement holds. Let $k \geq 1$ and assume that $P_j = R \cap \Lambda_j$ for all $j \leq k$. We will show that $P_{k+1} = R \cap \Lambda_{k+1}$.

First, consider an element of P_{k+1}. It may be written as $\beta + \alpha$ with $(\alpha, \beta) \in S \times P_k$ satisfying the conditions (i)-(iii). By the inductive hypothesis, $\beta \in R^+$. Moreover, there exists a smallest integer $p' \leq 0$ such that $\beta + p'\alpha \in R^+$. By the inductive hypothesis it follows that $p' = p$. The α-string through β now takes the form $L_\alpha(\beta) = \{\beta + k\alpha \mid p \leq k \leq q\}$ with q the non-negative integer determined by $p + q = -n_\alpha(\beta)$. From condition (iii) it follows that $q > 0$, hence $\beta + \alpha \in R^+$. It follows that $P_{k+1} \subset R \cap \Lambda_{k+1}$. For the converse inclusion, consider an element $\beta_1 \in R^+$ of weight $k + 1$. Since $k + 1 \geq 2$, the root β_1 does not belong to S. By Lemma 38.13 there exists a $\alpha \in S$ such that $\beta := \beta_1 - \alpha \in R^+$. Clearly, $\mathrm{ht}(\beta) = k$, so $\beta \in P_k$ by the inductive hypothesis. We will proceed to show that the pair (α, β) satisfies conditions (i)-(iii). This will imply that $\beta_1 \in P_{k+1}$, completing the proof.

Since β_1 is a root, $\alpha \neq \beta$, hence (i). Since $n_\alpha(\beta_1) = n_{\alpha\beta_1}$, condition (ii) holds by Lemma 38.2. The α-root string through β has the form $L_\alpha(\beta) = \{\beta + k\alpha \mid p \leq k \leq q\}$, with p the smallest integer such that $\beta + p\alpha$ is a root and with q the largest integer such that $\beta + q\alpha$ is a root. We note that $p \leq 0$ and $q \geq 1$. By the inductive hypothesis, p is the smallest integer such that $\beta \in P_1 \cup \cdots \cup P_k$. Moreover, by Lemma 38.5, $n_\alpha(\beta) = n_{\alpha\beta} = -(p + q)$ and (iii) follows. $\square$

38.3 The rank two root systems

We can use the method of the proof of Theorem 38.14 to classify the (isomorphism classes of) rank two root systems. Let (E, R) be a rank two root system. Then R has a fundamental system S consisting of two elements, α and β. Without loss of generality we may assume that $|\alpha| \leq |\beta|$. Moreover, changing the inner product on E by a positive scalar we may as well assume that $|\alpha| = 1$. From Lemma 38.7 it follows that there are 4 possible values for $n_{\alpha\beta}$, namely $0, -1, -2, -3$, with corresponding angles $\varphi_{\alpha\beta}$ equal to $\pi/2, 2\pi/3, 3\pi/4, 5\pi/6$. If $n_{\alpha\beta} = 0$ then the length of β is undetermined. In the remaining cases, the length of β equals 1, $\sqrt{2}$ and $\sqrt{3}$, respectively. It follows from Theorem 38.14 that for each of these cases there exists at most one isomorphism class of root spaces. We shall discuss these cases separately.

Case $n_{\alpha\beta} = 0$. In the notation of the proof of Theorem 38.14, $P_1 = \{\alpha, \beta\}$ It follows that P_2 can only contain the element $\beta + \alpha$. In the notation of condition (iii) of the mentioned proof, we have $p = 0$ and $n_\alpha(\beta) = 0$, hence $\beta + \alpha \notin P_2$. It follows that $P_j = \emptyset$ for $j \geq 2$. Therefore, $\mathcal{R} = \{\pm\alpha, \pm\beta\}$ is the only possible root system with the given Cartan matrix. We leave it to the reader to check that this is indeed a root system. It is called $A_1 \times A_1$.

Case $n_{\alpha\beta} = -1$. In this case $P_1 = \{\alpha, \beta\}$. There is only one possible element in P_2, namely $\beta + \alpha$. Here $p = 0$ and $n_\alpha(\beta) = -1$ whence $-p - n_\alpha(\beta) > 0$ and it follows that $\beta + \alpha \in P_2$. The possible elements in P_3 are $(\alpha + \beta) + \alpha$ or $(\alpha + \beta) + \beta$. For the first element, Now $p = -1$ and $n_\alpha(\alpha + \beta) = 1$, whence $2\alpha + \beta \notin P_3$. Similarly, $(\alpha + \beta) + \beta \notin P_3$. It follows that $P_j = \emptyset$ for $j \geq 3$. Hence, $R = \{\pm\alpha, \pm\beta, \pm(\alpha + \beta)\}$ is the only possible root system. We leave it to the reader to check that it is indeed a root system. It is called A_2.

Case $n_{\alpha\beta} = -2$. We have $P_1 = \{\alpha, \beta\}$ and $P_2 = \{\alpha + \beta\}$. The only possible elements in P_3 are $(\alpha+\beta)+\alpha$ and $(\alpha+\beta)+\beta$. For the first of these we have From $p = -1$ and $n_\alpha(\alpha+\beta) = 0$, so that $-p - n_\alpha(\alpha_+\mathfrak{g}) > 0$ and $\beta + 2\alpha \in P_3$. For the second element we have From $p = -1$ and $n_\beta(\alpha + \beta) = 1$, whence $-p - n_\beta(\alpha_+\beta) = 0$, from which we infer that $2\beta + \alpha \notin P_3$. Thus, $P_3 = \{\beta + 2\alpha\}$.

The possible elements of P_4 are $(\beta+2\alpha)+\alpha$ and $(\beta+2\alpha)+\beta$. For the first element, $p = -2$ and $n_\alpha(\beta + 2\alpha) = 2$, hence $\beta + 3\alpha \notin P_4$. For the second element, $p = 0$ and $n_\beta(\beta + 2\alpha) = 0$, hence $2\beta + 2\alpha \notin P_4$. We conclude that $P_j = \emptyset$ for $j \geq 4$.

Thus, in the present case the only possible root system is $R = \{\pm\alpha, \pm\beta, \pm(\alpha + \beta), \pm(\beta + 2\alpha)\}$. Again we leave it to the reader to check that this is a root system. It is called B_2.

Case $n_{\alpha\beta} = -3$. We have $P_1 = \{\alpha, \beta\}$ and $P_2 = \{\alpha + \beta\}$. The possible elements of P_3 are $\beta + 2\alpha$ and $2\beta + \alpha$. For the first element we have $p_{\alpha,\alpha+\beta} = -1$ and $n_\alpha(\alpha + \beta) = -1$, hence $\beta + 2\alpha \in P_3$. For the second we have $p_{\beta,\alpha+\beta} = -1$ and $n_\beta(\alpha + \beta) = 1$, hence $2\beta + \alpha \notin P_3$. Thus, $P_3 = \{\beta + 2\alpha\}$.

The possible elements of P_4 are $\beta+3\alpha$ and $2\beta+2\alpha$. For the first element we have $p_{\alpha,2\alpha+\beta} = -2$ and $n_\alpha(2\alpha+\beta) = 1$, hence $\beta+3\alpha \in P_3$. For the second, $p_{\beta,2\alpha+\beta} = 0$ and $n_\beta(2\alpha+\beta) = 0$, hence $2\beta + 2\alpha \notin P_4$. Thus, $P_4 = \{\beta + 3\alpha\}$.

The possible elements of P_5 are $\beta + 3\alpha + \alpha$ and $\beta + 3\alpha + \beta$. For the first element we have $p = -3$ and $n_\alpha(\beta + 3\alpha) = 3$, whence $\beta + 4\alpha \notin P_5$. For the second element we have $p = -1$ and $n_\beta(\beta + 3\alpha) = -1$, whence $2\beta + 3\alpha \in P_5$, and we conclude that $P_5 = \{2\beta + 3\alpha\}$.

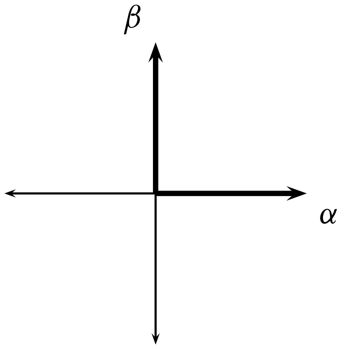
The possible elements of P_6 are $2\beta + 3\alpha + \alpha$ and $2\beta + 3\alpha + \beta$. For the first element we have $p = 0$ and $n_\alpha(2\beta + 3\alpha) = 0$, and for the second $p = -1$ and $n_\beta(2\beta + 3\alpha) = 1$. Hence $P_j = \emptyset$ for $j \geq 6$.

We conclude that the only possible root system is $R = \pm\{\alpha, \beta, \alpha + \beta, 2\alpha + \beta, 3\alpha + \beta, 3\alpha + 2\beta\}$. We leave it to the reader to check that this is indeed a root system, called G_2.

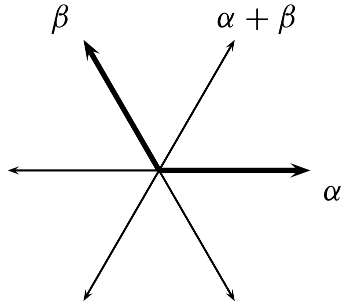
Lemma 38.16 *Up to isomorphism, the rank two root systems are completely classified by the integer $n_{\alpha\beta}n_{\beta\alpha}$, for $\{\alpha, \beta\}$ a fundamental system. The integer takes the values $\{0, 1, 2, 3\}$, giving the root systems $A_1 \times A_1$, A_2, B_2 and G_2, respectively.*

Proof: This has been established above. $\square$

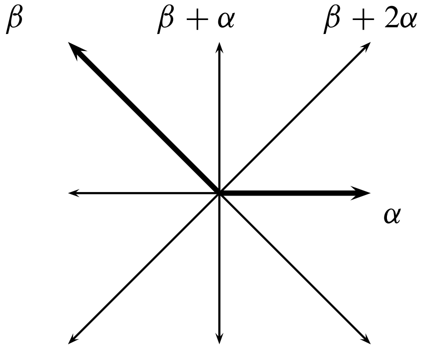
The rank 2 root systems are depicted below.



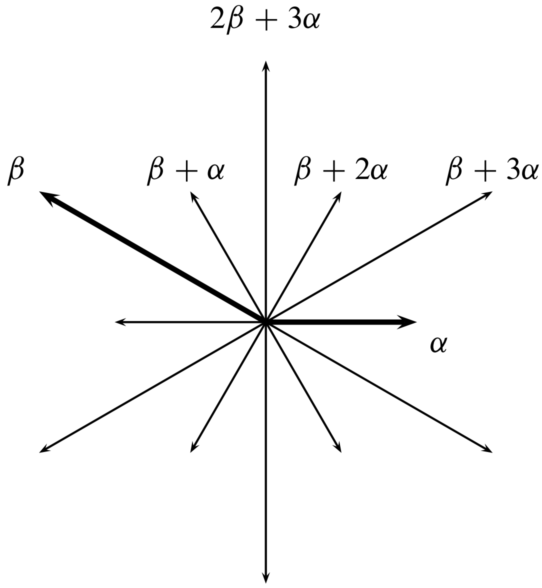
$A_1 \times A_1$



A_2



B_2



G_2

38.4 Weyl chambers

We proceed to investigate the collection of fundamental systems of the root system (E, R). An important role is played by the connected components of E^{reg}, see (58), called the *Weyl chambers* of R.

For every $\alpha \in R$, the complement $E \setminus P_\alpha$ is the disjoint union of the open half spaces $E^+(\alpha)$ and $E^+(-\alpha)$. Since E^{reg} is the intersection of the complements $E \setminus P_\alpha$, each Weyl chamber can be written in the form $\cap_{\alpha \in F} \; E^+(\alpha)$, with $\alpha \in F$, $F \subset R$. It follows that each Weyl chamber is an open polyhedral cone. We denote the set of Weyl chambers by $\mathcal{C}$. If $C \in \mathcal{C}$ then for every $\alpha \in R$ the functional $\langle \alpha, \cdot \rangle$ is nowhere zero on C, hence either everywhere positive or everywhere negative. We define

$$R^+(C) = \{\alpha \in R \mid \langle \alpha, \cdot \rangle > 0 \text{ on } C\}.$$

Note that for every $\gamma \in C$ we have $R^+(C) = R^+(\gamma)$. Thus, by Lemma 38.11 the set $R^+(C)$ is a positive system for R and every positive system arises in this way.

If C is a Weyl chamber, then by $S(C)$ we denote the collection of simple roots in the positive system $R^+(C)$. According to Lemma 38.9 this is a fundamental system for R.

Proposition 38.17

(a) *The map $C \mapsto R^+(C)$ defines a bijection between the collection of Weyl chambers and the collection of positive systems for R.*

(b) *The map $C \mapsto S(C)$ defines a bijection between the collection of Weyl chambers and the collection of fundamental systems for R.*

(c) *Is C is a Weyl chamber, then*

$$C = \{x \in E \mid \forall \alpha \in R^+(C) : \langle x, \alpha \rangle > 0\} = \{x \in E \mid \forall \alpha \in S(C) : \langle x, \alpha \rangle > 0\}.$$

Proof: Recall that we denote the collections of Weyl chambers, positive systems and fundamental systems by $\mathcal{C}$, $\mathcal{P}$ and $\mathcal{S}$, respectively.

If $P \in \mathcal{P}$ we define $C(P) := \{x \in E \mid \forall \alpha \in P : \langle x, \alpha \rangle > 0\}$, and if $S \in \mathcal{S}$ we put $C(S) := \{x \in E \mid \forall \alpha \in S : \langle x, \alpha \rangle > 0\}$. With this notation, assertion (c) becomes $C = C(R^+(C)) = C(S(C))$ for every $C \in \mathcal{C}$.

Let $S \in \mathcal{S}$. Then the set $C(S)$ is non-empty and convex, hence connected. Since $R \subset \mathbb{N}S \cup [-\mathbb{N}S]$, it follows that $C(S) \subset E^{\mathrm{reg}}$. We conclude that there exists a connected component $C \in \mathcal{C}$ such that $C(S) \subset C$. Every root from R has the same sign on C as on $C(S)$; hence, $C \subset C(S)$. We conclude that $C(S) = C$. In particular, $S \mapsto C(S)$ maps $\mathcal{S}$ into $\mathcal{C}$.

Let $P \in \mathcal{P}$ and let S be the collection of simple roots in P. From $S \subset P \subset \mathbb{N}S$ it readily follows that $C(S) = C(P)$. In particular, $C(P) \in \mathcal{C}$.

From Lemma 38.11 it follows that the map $C \mapsto R^+(C)$ is surjective. If $C \in \mathcal{C}$ then from the definitions it is obvious that $C \subset C(R^+(C)) \subset C(S(C))$. The extreme members in this chain of inclusions are Weyl chambers, i.e., connected components of E^{reg}, hence equal. Thus (c) follows. Moreover, $C(R^+(C)) = C$, from which it follows that $C \mapsto R^+(C)$ is injective, whence (a). Finally, (b) follows from (a) and (c) combined with Lemma 38.9. $\square$

The following result gives a useful characterization of the simple roots in terms of the associated Weyl chamber.

Lemma 38.18 *Let C be an open Weyl chamber. A root $\alpha \in R$ belongs to the associated fundamental system $S(C)$ if and only if the following two conditions are fulfilled.*

(a) $\langle \alpha, \cdot \rangle > 0$ *on* C;

(b) $\overline{C} \cap \alpha^\perp$ *has non-empty interior in* $\alpha^\perp$.

Proof: Put $S = S(C)$ and assume that $\alpha \in S$. Then (a) follows by definition. From Proposition 38.17 we know that C consists of the points $x \in E$ with $\langle x\,,\,\beta\rangle > 0$ for all $\beta \in S$. Since S is a basis of the linear space E, it is readily seen that $\overline{C}$ consists of the points $x \in E$ with $\langle x\,,\,\beta\rangle \geq 0$ for all $\beta \in S$. The functionals $\langle \beta\,,\,\cdot\rangle|_{\alpha^\perp}$, for $\beta \in S \setminus \{\alpha\}$, form a basis of $\alpha^\perp$, hence the set $\bar{C} \cap \alpha^\perp$ contains the non-empty open subset of $\alpha^\perp$ consisting of the points $x \in \alpha^\perp$ with $\langle x\,,\,\beta\rangle > 0$ for all $\beta \in S \setminus \{\alpha\}$. This implies (b).

Conversely, assume that α is a root and that (a) and (b) are fulfilled. From (a) it follows that $\alpha \in R^+(C)$. It remains to be shown that α is indecomposable. Assume the latter were not true. Then $\alpha = \beta + \gamma$, for $\beta, \gamma \in R^+(C)$. From (b) it follows that $\langle \beta\,,\,\cdot\rangle \geq 0$ and $\langle \gamma\,,\,\cdot\rangle \geq 0$ on an open subset U of $\alpha^\perp$ On the other hand, $\langle \beta+\gamma\,,\,\cdot\rangle = 0$ on U. It follows that $\langle \beta\,,\,\cdot\rangle$ and $\langle \gamma\,,\,\cdot\rangle$ are zero on U, hence on $\alpha^\perp$ by linearity. From this it follows in turn that $\beta^\perp = \alpha^\perp = \gamma^\perp$. Hence β and γ are proportional to α, contradiction. $\square$

The Weyl group leaves R, hence E^{reg}, invariant. It follows that W acts on the set of connected components on E^{reg}, i.e., on the set $\mathcal{C}$ of Weyl chambers. Clearly, W acts on the set of positive systems and on the set of fundamental systems, and the actions are compatible with the maps of Proposition 38.17. More precisely, if $w \in W$ and $C \in \mathcal{C}$, then $R^+(wC) = wR^+(C)$ and $S(wC) = wS(C)$.

Lemma 38.19 *Let R^+ be a positive system for R and let α be an associated simple root. Then s_α maps $R^+ \setminus \{\alpha\}$ onto itself.*

Proof: Let S be the set of simple roots in R^+ and let $\beta \in R^+$, $\beta \neq \alpha$. Then $\beta = \sum_{\gamma \in S} k_\gamma \gamma$, with $k_\gamma \in \mathbb{N}$ and $k_{\gamma_0} > 0$ for at least one γ_0 different from α. Now $s_\alpha(\beta) = \sum_{\gamma \in S \setminus \{\alpha\}} k_\gamma \gamma + l_\alpha \alpha$ for some $l_\alpha \in \mathbb{Z}$. Since $s_\alpha \beta$ is a root, it either belongs to $\mathbb{N}S$ or to $-\mathbb{N}S$. The latter possibility is excluded by $k_{\gamma_0} > 0$. Hence $s_\alpha \beta \in \mathbb{N}S \cap R = R^+$. $\square$

If R^+ is a positive system for R, we define $\delta(R^+) = \delta$ to be half the sum of the positive roots, i.e.,

$$\delta = \frac{1}{2} \sum_{\gamma \in R^+} \gamma.$$

Corollary 38.20 *If α is simple in R^+, then $s_\alpha \delta = \delta - \alpha$.*

Proof: Write $\delta = \frac{1}{2} \sum_{\gamma \in R^+ \setminus \{\alpha\}} \gamma + \frac{1}{2}\alpha$. The sum in the first term is fixed by s_α, whereas the term $\frac{1}{2}\alpha$ is mapped onto $-\frac{1}{2}\alpha$. $\square$

Two Weyl chambers C, C' are said to be separated by the root hyperplane $\alpha^\perp$ if the linear functional $\langle \alpha\,,\,\cdot\rangle$ has different signs on C and C'. We will write $d(C, C')$ for the number of root hyperplanes separating C and C'. If P is any positive system for R then $d(C, C')$ is the number of $\alpha \in P$ such that $\langle \alpha\,,\,\cdot\rangle$ has different signs on C and C' (use that R is the disjoint union of P and $-P$ and that roots define the same hyperplane if and only if they are proportional). In particular,

$$d(C, C') = \#[\, R^+(C) \setminus R^+(C')\,].$$

Definition 38.21 Two Weyl chambers C and C' are called *adjacent* if $d(C, C') = 1$, i.e., the chambers are separated by precisely one root hyperplane.

Lemma 38.22 *Let C, C' be Weyl chambers. Then C, C' are adjacent if and only if $C' = s_\alpha(C)$ for some $\alpha \in S(C)$. If the latter holds, then $-\alpha \in S(C')$.*

Proof: Let C and C' be adjacent. Then $R^+(C) \setminus R^+(C') = \{\alpha\}$ for a unique root α. From $S(C) \setminus R^+(C') = \emptyset$ it would follow that $S(C) \subset R^+(C')$, whence $R^+(C) \subset R^+(C')$. Since both members of this inclusion have half the cardinality of R, they must be equal, contradiction. Hence $S(C) \setminus R^+(C')$ contains a root, which must be α. Similarly, $S(C')$ contains the root $-\alpha$. Since $R^+(C')$ and $R^+(C)$ have the same cardinality, we infer that $R^+(C') = [R^+(C) \setminus \{\alpha\}] \cup \{-\alpha\} = s_\alpha(R^+(C))$, by Lemma 38.19. It follows that $R^+(C') = R^+(s_\alpha(C))$, hence $C' = s_\alpha(C)$.

Conversely, assume that $\alpha \in S(C)$ and $s_\alpha(C) = C'$. Then $R^+(C') = s_\alpha R^+(C) = [R^+(C) \setminus \{\alpha\}] \cup \{-\alpha\}$ from which one sees that $\# R^+(C) \setminus R^+(C') = 1$. Hence, C and C' are adjacent. □

Lemma 38.23 *Let C, C' be distinct Weyl chambers. Then there exists a chamber C'' that is adjacent to C' and such that $d(C, C'') = d(C, C') - 1$.*

Proof: There must be a root $\alpha \in S(C') \setminus R^+(C)$, for otherwise $S(C') \subset R^+(C)$, hence $R^+(C') \subset R^+(C)$, contradiction. Let $C'' = s_\alpha(C')$. Then C' and C'' are adjacent by the previous lemma. Also, by Lemma 38.19, $R^+(C'') = s_\alpha R^+(C') = [R^+(C') \setminus \{\alpha\}] \cup \{-\alpha\}$. From this we see that $R^+(C') \setminus R^+(C)$ is the disjoint union of $R^+(C'') \setminus R^+(C)$ and $\{\alpha\}$. It follows that $d(C, C'') = d(C, C') - 1$. □

Lemma 38.24 *Let C be a Weyl chamber and $S = S(C)$ the associated fundamental system. Then for every Weyl chamber $C' \neq C$ there exists a sequence $s_1, \dots s_n$ of reflections in roots from S such that $C' = s_1 \cdots s_n(C)$.*

Proof: We give the proof by induction on $d = d(C, C')$. If $d = 1$, then the result follows from Lemma 38.22. Thus, let $d > 1$ and assume the result has been established for C' with $d(C, C') < d$. By the previous lemma, there exists a chamber C'', adjacent to C' and such that $d(C, C'') = d(C, C') - 1$. By Lemma 38.22, $C'' = s_\alpha(C')$ for a simple root $\alpha \in S(C')$.

By the induction hypothesis there exists a $w \in W$ that can be expressed as a product of reflections in roots from $S(C)$ such that $w(C) = C''$. Thus, $s_\alpha w(C) = s_\alpha(C'') = C'$. Moreover, $s_\alpha w = w s_{w^{-1}\alpha} = w s_{-w^{-1}\alpha}$, and since $-\alpha \in S(C'')$, it follows that $\beta := -w^{-1}\alpha$ belongs to $S(C) = w^{-1}S(C'')$. We conclude that $C' = ws(C)$ with w a product of reflections from roots in $S(C)$ and with $s = s_\beta$, reflection in a root from $S(C)$. □

Lemma 38.25 *Let S be a fundamental system for R. Then every root from R is conjugate to a root from S by an element of W that can be written as a product of simple reflections, i.e., reflections in roots from S.*

Proof: Let $\alpha \in R$. There exists a Weyl chamber C such that $\alpha^\perp \cap \overline{C}$ has non-empty interior in $\alpha^\perp$. By Lemma 38.18 it follows that either α or $-\alpha$ belongs to $S(C)$. Replacing C by $s_\alpha(C)$ if necessary, we may assume that $\alpha \in S(C)$. Let C^+ be the unique Weyl chamber with $S(C^+) = S$. Then there exists a Weyl group element of the form stated such that $w^{-1}(C) = C^+$. It follows that $w\alpha \in S(C^+) = S$. $\square$

Corollary 38.26 *Let S be a fundamental system for R. Then W is already generated by the associated collection of simple reflections.*

Proof: Let W_0 be the subgroup of W generated by reflections in roots from S. Let $\alpha \in R$. Then by the previous lemma there exists a $w \in W_0$ such that $\alpha = w\beta$, with $\beta \in S$. It follows that $s_\alpha = w s_\beta w^{-1} \in W_0$. Since W is generated by the s_α, for $\alpha \in R$, it follows that $W = W_0$. $\square$

Definition 38.27 Let S be a fundamental system for R. If $w \in W$ then an expression $w = s_1 \cdots s_n$ of w in terms of simple reflections is called a *reduced expression* if it is not possible to extract a non-empty collection of factors without changing the product.

Lemma 38.28 *Let $\alpha_1, \ldots, \alpha_n \in S$ be simple roots (possibly with repetitions), and let $s_j = s_{\alpha_j}$ be the associated simple reflections. Assume that $s_1 \cdots s_n(\alpha_n)$ is positive relative to S. Then $s_1 \cdots s_n$ is not a reduced expression. More precisely, there exists a $1 \leq k < n$ such that*

$$s_1 \cdots s_n = s_1 \cdots s_{k-1} s_{k+1} \cdots s_{n-1}.$$

Proof: Write $\beta_j = s_{j+1} \ldots s_{n-1}(\alpha_n)$, for $0 \leq j < n$. Let P be the positive system determined by S. Then $\beta_0 \in -P$ and $\beta_{n-1} = \alpha_n \in P$, hence there exists a smallest index $1 \leq k \leq n-1$ such that $\beta_k \in P$. We have that $s_k(\beta_k) = \beta_{k-1} \in -P$, hence, by Lemma 38.19, $\beta_k = \alpha_k$. We now observe that for every $w \in W$ we have $w s_n = s_{w\alpha_n} w$. Applying this with $w = s_{k+1} \ldots s_{n-1}$ we obtain $s_{k+1} \cdots s_{n-1} s_n = s_{\beta_k} s_{k+1} \cdots s_{n-1} = s_k \cdots s_{n-1}$. This implies that

$$s_1 \cdots s_n = s_1 \cdots s_k s_k \cdots s_{n-1} = s_1 \cdots s_{k-1} s_{k+1} \cdots s_{n-1}.$$

$\square$

Lemma 38.29 *The Weyl group acts simply transitively on the set of Weyl chambers.*

Proof: Let $\mathcal{C}$ denote the collection of Weyl chambers. The transitivity of the action of W on $\mathcal{C}$ follows from Lemma 38.24. To establish that the action is simple, we must show that for all $C \in \mathcal{C}$ and $w \in W$, $wC = C \Rightarrow w = 1$.

Fix $C \in \mathcal{C}$ and let $S = S(C)$ be the associated fundamental system for R. Let $w \in W \setminus \{1\}$. Then w^{-1} has a reduced expression of the form $w^{-1} = s_1 \cdots s_n$, with $n \geq 1$, $s_j = s_{\alpha_j}$, $\alpha_j \in S(C)$. From Lemma 38.28 it follows that $w^{-1}\alpha_n < 0$ on C, hence $\alpha_n < 0$ on $w(C)$. It follows that $w(C) \neq C$. $\square$

Remark 38.30 It follows from the above result, combined with Proposition 38.17, that the Weyl group acts simply transitively on the collection of fundamental systems for R as well as on the collection of positive systems.

Let S, S' be two fundamental systems, and let w be the unique Weyl group element such that $w(S) = S'$. Let $n : S \times S \to Z$ and $n' : S' \times S' \to \mathbb{Z}$ be the associated Cartan matrices. Then it follows from Lemma 38.1 that $n'(w\alpha, w\beta) = n(\alpha, \beta)$ for all $\alpha, \beta \in S$, or more briefly, $w^*n' = n$. Thus, the Cartan matrices are essentially equal.

Let S be a fixed fundamental system for R. From now on we denote the associated positive system by R^+. The elements of S are called the simple roots, those of R^+ are called the positive roots. The associated Weyl chamber

$$E^+ = \{x \in E \mid \forall\alpha \in R^+ : \ \langle \alpha\, , \, x \rangle > 0\}$$

is called the associated positive chamber. Given a root α, we will use the notation $\alpha > 0$ to indicate that $\alpha \in R^+$; this is equivalent to $\langle \alpha\, , \, \cdot \rangle > 0$ on E^+.

We define numbers $l_S(w) = l(w)$ and $n_S(w) = n(w)$ for a Weyl group element $w \in W$. Firstly, $l(w)$, the length of w, is by definition the shortest length of a reduced expression for w. Secondly, $n(W)$ is the number of positive roots $\alpha \in R^+$ such that $w\alpha$ is negative, i.e., $w\alpha \in -R^+$.

Remark 38.31 In general, the numbers $l_S(w)$ and $n_S(w)$ do depend on the particular choice of fundamental system. This can already be verified for the root system A_2.

Lemma 38.32 *For every $w \in W$,*

$$n(w) = l(w) = d(E^+, w^{-1}(E^+)) = d(E^+, w(E^+)).$$

Moreover, any reduced expression for w, relative to S, has length $l(w)$.

Proof: $d(E^+, w^{-1}(E^+))$ equals the number of positive roots $\alpha \in R^+$ such that $\alpha < 0$ on $w^{-1}(E^+)$. The latter condition is equivalent with $w\alpha < 0$ on E^+ or $w\alpha \in -R^+$. Thus, $n(w) = d(E^+, w^{-1}(E^+))$. On the other hand, clearly

$$d(E^+, w^{-1}(E^+)) = d(wE^+, ww^{-1}E^+) = d(E^+, wE^+).$$

It follows from the proof of Lemma 38.24 that any reduced expression has length at most $d(E, wE^+)$. In particular, $l(w) \leq d(E^+, wE^+)$.

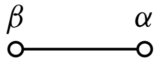
We will finish the proof by showing that $n(w) \leq l(w)$, by induction on $l(w)$. If $l(w) = 1$, then w is a simple reflection, and the inequality is obvious. Thus, let $n > 1$ and assume the estimate has been established for all w with $l(w) < n$. Let $w \in W$ with $l(w) = n$. Then w has a reduced expression of the form $w = s_1 \cdots s_{n-1}s_\alpha$, with $\alpha \in S(C)$. Put $v = s_1 \ldots s_{n-1}$; this expression must be reduced, hence $l(v) < n$ and it follows that $n(v) \leq n-1$ by the inductive hypothesis. On the other hand, from Lemma 38.28 it follows that $w\alpha \in -R^+$, hence $\beta := v\alpha > 0$. The root β belongs to $S(vE^+)$, hence $R^+(wE^+) = R^+(s_\beta vE^+) = [R^+(vE^+) \setminus \{\beta\}] \cup \{-\beta\}$. It follows that $R^+ \setminus R^+(wE^+)$ is the disjoint union of $R^+ \setminus R^+(vE^+)$ and $\{\beta\}$. Hence $n(w) = d(E^+, wE^+) = d(E^+, vE^+) + 1 = n(v) + 1 \leq l(v) + 1 \leq l(w)$. $\square$

38.5 Dynkin diagrams

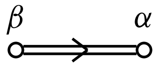
Let (E, R) be a root system, S a fundamental system for R. The *Coxeter graph* attached to S is defined as follows. The vertices of the graph are in bijective correspondence with the roots of S; two vertices α, β are connected by $n_{\alpha\beta} \cdot n_{\beta\alpha}$ edges. Thus, every pair is connected by 0, 1, 2 or 3 edges, see the table in Lemma 38.2.

The *Dynkin diagram* of S consists of the Coxeter graph together with the symbol $>$ or $<$ attached to each multiple edge, pointing towards the shorter root. From Lemma 38.16 it follows that (up to isomorphism) the Dynkin diagrams of the rank-2 root systems are given by the following list:

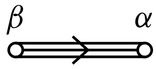
$A_1 \times A_1$



A_2



B_2



G_2

It follows from Remark 38.30 that the Dynkin diagrams for two different choices of fundamental systems for R are isomorphic (in an obvious sense). We may thus speak of the Dynkin diagram of a root system. The following result expresses that the classification of root systems amounts to describing the list of all possible Dynkin diagrams.

Theorem 38.33 *Let R_1, R_2 be two root systems. If the Dynkin diagrams associated with R_1 and R_2 are isomorphic, then R_1 and R_2 are isomorphic as well.*

Proof: Let S_1 and S_2 be fundamental systems for R_1 and R_2, respectively. It follows from Lemma 38.2 that the Cartan matrices n_1 and n_2 of S_1 and S_2 are completely determined by their Dynkin diagrams. An isomorphism between these Dynkin diagrams gives rise to a bijection $\varphi : S_1 \to S_2$ such that, $n_1 = \varphi^* n_2$. By Theorem 38.14 it follows that R_1 and R_2 are isomorphic. $\square$

Remark 38.34 It follows from the above result combined with Theorem 36.12 that the (isomorphism classes of) Dynkin diagrams are in bijective correspondence with the isomorphism classes of semisimple compact Lie algebras.

Let S be a fundamental system. The decomposition of its Dynkin diagram D into connected components D_j, $(1 \leq j \leq p)$, determines a decomposition of S into a disjoint union of subsets $S_j, (1 \leq j \leq p)$. Here S_j consists of the roots labelling the vertices in D_j. The decomposition of S is uniquely determined by the conditions that $S_i \perp S_j$ if $i \neq j$, and that every S_j cannot be written as a disjoint union of proper subsets S_{j1}, S_{j2} with $S_{j1} \perp S_{j2}$. We will investigate what this means for the root system R.

If (E_j, R_j), with $j = 1, 2$, are two root systems, we define their direct sum (E, R) as follows. First, $E := E_1 \oplus E_2$. Via the natural embeddings $E_j \to E$, the sets R_1 and R_2 may be viewed as subsets of E; accordingly we define R to be their union. If $\alpha \in R_1$, the map $s_\alpha \oplus I$ is a reflection in $(\alpha, 0)$ preserving R. By a similar remark for R_2, we see that R is a root system. Moreover, for all $\alpha \in R_1$ and $\beta \in R_2$, $n_{\alpha\beta} = 0$. From this we see that $E_1 \perp E_2$ for every W-invariant inner product on E. Every reflection preserves both R_1 and R_2, hence E_1 and E_2 are invariant subspaces for the Weyl group. Moreover, the maps $v \mapsto v \otimes I$ and $w \mapsto I \otimes w$ define embeddings $W_1 \hookrightarrow W$ and $W_2 \hookrightarrow W$ via which we shall identify. Accordingly we have $W = W_1 \times W_2$. Similar remarks hold for the direct sum of finitely many root systems.

Definition 38.35 A root system (E, R) is called *reducible* if R is the union of two non-empty subsets R_1 and R_2 such that $E = \mathrm{span}(R_1) \oplus \mathrm{span}(R_2)$. It is called *irreducible* if it is not reducible.

The following result expresses that every root system allows a decomposition as a direct sum of irreducibles, which is essentially unique.

Proposition 38.36 *Let (E, R) be a root system. Then there exist finitely many linear subspaces E_j, $1 \le j \le n$, such that $R_j := E_j \cap R$ is an irreducible root system for every j, and such that $R = \cup_j R_j$. The E_j are uniquely determined up to order.*

If S_j is a fundamental system of R_j, for $j = 1 \cdots n$, then $S = S_1 \cup \cdots \cup S_n$ is a fundamental system for R. Every fundamental system for R arises in this way.

If P_j is a positive system of P_j, for $j = 1 \cdots n$, then $P = P_1 \cup \cdots \cup P_n$ is a positive system for R. Every positive system of R arises in this way.

Proof: From the definition of irreducibility, it follows that (E, R) has a decomposition as stated. We will establish its uniqueness at the end of the proof.

If the S_j are fundamental systems as stated, then it is readily checked from the definition that their union S is a fundamental system for R. Let P_j be positive systems as stated, then again from the definition it is readily verified that their union P is a positive system for R.

Conversely, let P be a positive system for R. Then it is readily verified that every set $P_j := P \cap R_j$ is a positive system for R_j. Moreover, let S be a fundamental system for R. Since R is the disjoint union of the sets R_j, it follows that S is the disjoint union of the sets $S_j := S \cap R_j$. Each S_j is linearly independent, hence for dimensional reasons a basis of E_j. Now $R_j \subset (\mathbb{N}S \cup (-\mathbb{N}S))$ and $R_j \subset \mathbb{R}S_j$. By linear independence this implies that $R_j \subset \mathbb{N}S_j \cup (-\mathbb{N}S_j)$ for every j. Hence every S_j is a fundamental system.

We now turn to uniqueness of the decomposition as stated. Let $E = \oplus_{1 \le j \le m} E'_j$ be a decomposition with similar properties. Fix a fundamental system S'_j for $R'_j = R \cap E'_j$, for every j. The union S' is a fundamental system for R hence of the form $S = S_1 \cup \cdots \cup S_n$, with S_j a fundamental system for R_j, for each j. It follows that S'_1 is the disjoint union of the sets $S'_1 \cap S_j$, $1 \le j \le n$. Hence E'_1 is the direct sum of the spaces $E'_1 \cap E_j$ and R'_1 is the union of the sets $R'_1 \cap R_j = R'_1 \cap E_j$. From the irreducibility of E'_1 it follows that there exists a unique j such that $E'_1 = E_j$. The other components may be treated similarly. $\square$

In view of the above result we may now call the uniquely determined (E_j, R_j) the irreducible components of the root system (E, R).

Lemma 38.37

(a) *Let R be a root system. Then the Dynkin diagram of R is the disjoint union of the Dynkin diagrams of the irreducible components of R.*

(b) *A root system is irreducible if and only if the associated Dynkin diagram is connected.*

Proof: Let (E, R) be an root system, with irreducible components (E_j, R_j). Select a fundamental system S_j for each R_j and let S be their union. The inclusion $S_j \subset S$ induces an inclusion of $D_j \hookrightarrow D$ via which we may identify. For distinct indices i, j we have $n_{\alpha\beta} = 0$ for all $\alpha \in S_i$, $\beta \in S_j$. Hence no vertex of D_i is connected with any vertex of D_j. It follows that D is the disjoint union of the D_j, and (a) follows.

We turn to (b). If R is reducible, then by (a), the associated Dynkin diagram is not connected. Conversely, assume that the Dynkin diagram of R is not connected. Then it may be written as the disjoint union of two non-empty diagrams D_1 and D_2. Fix a fundamental system S of R. Then S decomposes into a disjoint union of two non-empty subsets S_1 and S_2 such that the elements of S_j label the vertices of D_j. It follows that for all $\alpha \in S_1$ and all $\beta \in S_2$, $n_{\alpha\beta} = 0$. Put $E_j = \mathrm{span}(S_j)$, then it follows that for each $\alpha \in S$ the reflection s_α leaves the decomposition $E = E_1 \oplus E_2$ invariant. Hence, the Weyl group W of R leaves the decomposition invariant. Let $\beta \in R$, then there exists a $w \in W$ such that $w\beta \in S = S_1 \cup S_2$. It follows that β lies either in E_1 or in E_2. Hence $R = R_1 \cup R_2$ with $R_j = E_j \cap R$, and we see that R is reducible. $\square$

The following result relates the notion of irreducibiliy of a root system with decomposability of a semisimple Lie algebra.

Proposition 38.38 *Let $\mathfrak{g}$ be a compact semisimple Lie algebra with Dynkin diagram D. Let $D = D_1 \cup \ldots \cup D_n$ be the decomposition of D into its connected components. Then every D_j is the Dynkin diagram of a compact simple Lie algebra $\mathfrak{g}_j$. Moreover,*

$$\mathfrak{g} \simeq \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n.$$

In particular, $\mathfrak{g}$ is simple if and only if D is connected.

Remark 38.39 Note that in view of Lemma 35.10 the above result implies that the connected components of D are in bijective correspondence with the simple ideals of $\mathfrak{g}$.

Proof: Let $\mathfrak{g} = \oplus_j \mathfrak{h}_j$ be the decomposition of $\mathfrak{g}$ into its simple ideals. For each j we fix a maximal torus $\mathfrak{t}_j \subset \mathfrak{h}_j$. Then $\mathfrak{t} := \mathfrak{t}_1 \oplus \cdots \oplus \mathfrak{t}_n$ is a maximal torus in $\mathfrak{g}$ (use that $\mathfrak{h}_i$ commutes with $\mathfrak{h}_j$ for every $i \neq j$). Via the direct sum decomposition of $\mathfrak{t}$, we view $\mathfrak{t}_j^*$ as the linear subspace of elements of $\mathfrak{t}^*$ that vanish on $\mathfrak{t}_k$ for every $k \neq j$. Accordingly, $\mathfrak{t}^* = \mathfrak{t}_1^* \oplus \cdots \oplus \mathfrak{t}_n^*$, and a similar decomposition of the complexification. Let R_j be the root system of $\mathfrak{t}_j$ in $\mathfrak{h}_j$. Since $\mathfrak{g}_\mathbb{C}$ is the direct sum of $\mathfrak{t}_\mathbb{C}$ and the root spaces $\mathfrak{g}_{\mathbb{C}\alpha}$, for $\alpha \in R_1 \cup \cdots \cup R_n$, it follows that the root system R of $\mathfrak{t}$ in $\mathfrak{g}$ equals the disjoint union of the R_j. Hence, R is the direct sum of the R_j.

The Dynkin diagram of R is the disjoint union of the Dynkin diagrams of the R_j. The proof will be finished if we can show that the Dynkin diagram of R_j, is connected, for each j. By Lemma 38.37 this is equivalent to the assertion that each R_j is irreducible.

Thus, we may assume $\mathfrak{g}$ is simple, $\mathfrak{t}$ a maximal torus in $\mathfrak{g}$, and then we must show that $R = R(\mathfrak{g}, \mathfrak{t})$ is irreducible. Assume not. Then we may decompose R as the disjoint union of two non-empty subsets R_1 and R_2 whose spans have zero intersection. Put $E = i\mathfrak{t}^*$, and for $j = 1, 2$, define $E_j = \text{span}(R_j)$. Then $E = E_1 \oplus E_2$. Let

$$\mathfrak{t}_1 := \cap_{\alpha \in R_2} \ker \alpha \quad \text{and} \quad \mathfrak{t}_2 := \cap_{\beta \in R_1} \ker \beta.$$

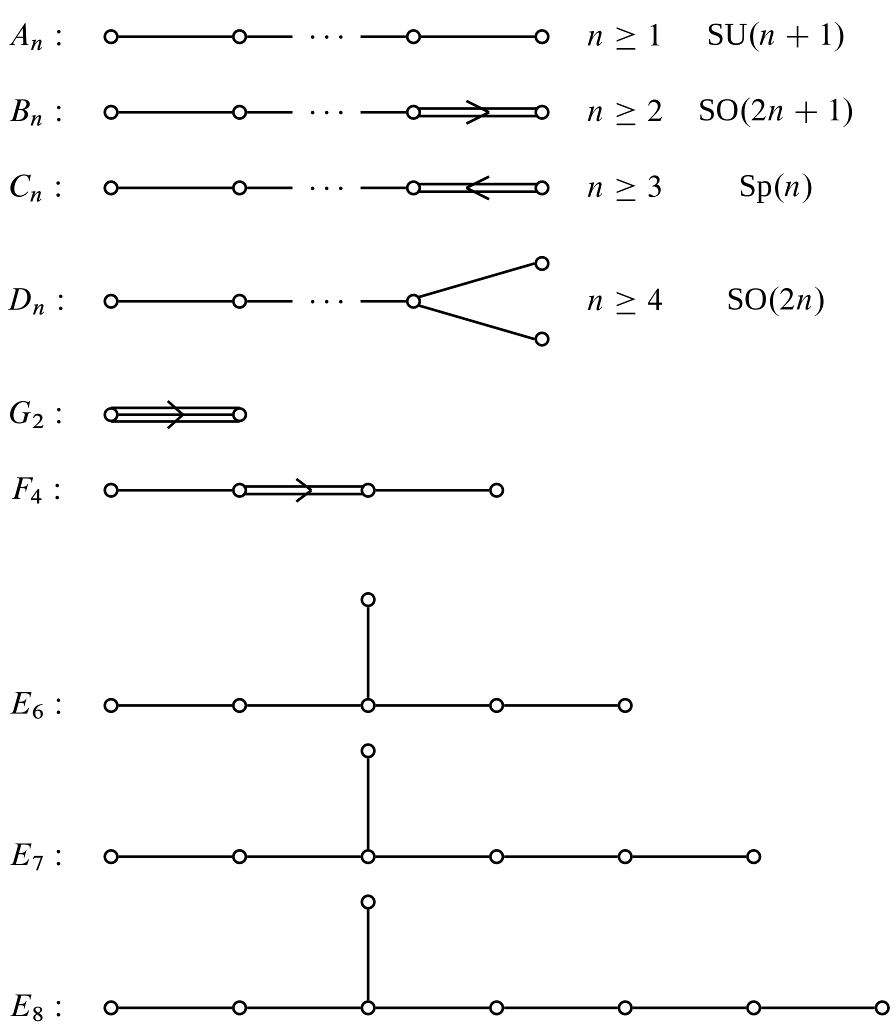
Then $\mathfrak{t} = \mathfrak{t}_1 \oplus \mathfrak{t}_2$ and, accordingly, $E_j \simeq i\mathfrak{t}_j^*$. For $j = 1, 2$, let

$$\mathfrak{g}_j = \mathfrak{t}_j \oplus (\, \mathfrak{g} \cap \sum_{\alpha \in R_j} \mathfrak{g}_{\mathbb{C}\alpha}\,).$$

Then $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ as a vector space. Moreover, $\text{ad}\,\mathfrak{t}$ normalizes this decomposition, $\mathfrak{t}_1$ centralizes $\mathfrak{g}_2$ and $\mathfrak{t}_2$ centralizes $\mathfrak{g}_1$. If $\alpha, \beta \in R$ and $\alpha + \beta \in R$, then we must have that $\{\alpha, \beta\}$ is a subset of either R_1 or R_2. From this we readily see that $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are subalgebras of $\mathfrak{g}$. Moreover, if $\alpha \in R_1$ and $\beta \in R_2$, then $\alpha + \beta \notin R$, hence $\mathfrak{g}_{\mathbb{C}(\alpha+\beta)} = 0$. It follows that $[\mathfrak{g}_1, \mathfrak{g}_2] = 0$. We conclude that $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ as a direct sum of ideals, contradicting the assumption that $\mathfrak{g}$ is simple. $\square$

In view of the above the following result amounts to the classification of all simple compact Lie algebras.

Theorem 38.40 *The following is a list of all connected Dynkin diagrams of root systems. These diagrams are in bijective correspondence with the (isomorphism classes of) the simple compact Lie algebras.*



Acknowledgement: I warmly thank Lotte Hollands for providing me with LaTeX files for these and all other pictures in the lecture notes.

Index