Lectures on Lie Groups

Dragan Miličić

Contents

CHAPTER 1

Basic differential geometry

1. Differentiable manifolds

1.1. Differentiable manifolds and differentiable maps. Let M be a topological space. A *chart* on M is a triple $c = (U, \varphi, p)$ consisting of an open subset $U \subset M$, an integer $p \in \mathbb{Z}_+$ and a homeomorphism φ of U onto an open set in $\mathbb{R}^p$. The open set U is called the *domain* of the chart c, and the integer p is the *dimension* of the chart c.

The charts $c = (U, \varphi, p)$ and $c' = (U', \varphi', p')$ on M are *compatible* if either $U \cap U' = \emptyset$ or $U \cap U' \neq \emptyset$ and $\varphi' \circ \varphi^{-1} : \varphi(U \cap U') \longrightarrow \varphi'(U \cap U')$ is a C^∞-diffeomorphism.

A family $\mathcal{A}$ of charts on M is an *atlas* of M if the domains of charts form a covering of M and any two charts in $\mathcal{A}$ are compatible.

Atlases $\mathcal{A}$ and $\mathcal{B}$ of M are *compatible* if their union is an atlas on M. This is obviously an equivalence relation on the set of all atlases on M. Each equivalence class of atlases contains the largest element which is equal to the union of all atlases in this class. Such atlas is called *saturated*.

A differentiable manifold M is a hausdorff topological space with a saturated atlas.

Clearly, a differentiable manifold is a locally compact space. It is also locally connected. Therefore, its connected components are open and closed subsets.

Let M be a differentiable manifold. A chart $c = (U, \varphi, p)$ is a chart *around* $m \in M$ if $m \in U$. We say that it is *centered* at m if $\varphi(m) = 0$.

If $c = (U, \varphi, p)$ and $c' = (U', \varphi', p')$ are two charts around m, then $p = p'$. Therefore, all charts around m have the same dimension. Therefore, we call p the *dimension of M at the point m* and denote it by $\dim_m M$. The function $m \longmapsto \dim_m M$ is locally constant on M. Therefore, it is constant on connected components of M.

If $\dim_m M = p$ for all $m \in M$, we say that M is an *p-dimensional* manifold.

Let M and N be two differentiable manifolds. A continuous map $F : M \longrightarrow N$ is a *differentiable map* if for any two pairs of charts $c = (U, \varphi, p)$ on M and $d = (V, \psi, q)$ on N such that $F(U) \subset V$, the mapping

$$\psi \circ F \circ \varphi^{-1} : \varphi(U) \longrightarrow \varphi(V)$$

is a C^∞-differentiable map. We denote by $\mathrm{Mor}(M, N)$ the set of all differentiable maps from M into N.

If N is the real line $\mathbb{R}$ with obvious manifold structure, we call a differentiable map $f : M \longrightarrow \mathbb{R}$ a differentiable function on M. The set of all differentiable functions on M forms an algebra $C^\infty(M)$ over $\mathbb{R}$ with respect to pointwise operations.

Clearly, differentiable manifolds as objects and differentiable maps as morphisms form a category. Isomorphisms in this category are called *diffeomorphisms*.

1.2. Tangent spaces. Let M be a differentiable manifold and m a point in M. A linear form ξ on $C^\infty(M)$ is called a *tangent vector* at m if it satisfies

$$\xi(fg) = \xi(f)g(m) + f(m)\xi(g)$$

for any $f, g \in C^\infty(M)$. Clearly, all tangent vectors at m form a linear space which we denote by $T_m(M)$ and call the *tangent space* to M at m.

Let $m \in M$ and $c = (U, \varphi, p)$ a chart centered at m. Then, for any $1 \leq i \leq p$, we can define the linear form

$$\partial_i(f) = \frac{\partial(f \circ \varphi^{-1})}{\partial x_i}(0).$$

Clearly, ∂_i are tangent vectors in $T_m(M)$.

1.2.1. Lemma. *The vectors $\partial_1, \partial_2, \dots, \partial_p$ for a basis of the linear space $T_m(M)$. In particular, $\dim T_m(M) = \dim_m M$.*

Let $F : M \longrightarrow N$ be a morphism of differentiable manifolds. Let $m \in M$. Then, for any $\xi \in T_m(M)$, the linear form $T_m(F)\xi : g \longmapsto \xi(g \circ F)$ for $g \in C^\infty(N)$, is a tangent vector in $T_{F(m)}(N)$. Clearly, $T_m(F) : T_m(M) \longrightarrow T_{F(m)}(N)$ is a linear map. It is called the *differential* of F at m.

The *rank* $\operatorname{rank}_m F$ of a morphism $F : M \longrightarrow N$ at m is the rank of the linear map $T_m(F)$.

1.2.2. Lemma. *The function $m \longmapsto \operatorname{rank}_m F$ is lower semicontinuous on M.*

1.3. Local diffeomorphisms, immersions, submersions and subimmersions. Let $F : M \longrightarrow N$ be a morphism of differentiable manifolds. The map F is a *local diffeomorphism at m* if there is an open neighborhood U of m such that $F(U)$ is an open set in N and $F : U \longrightarrow F(U)$ is a diffeomorphism.

1.3.1. Theorem. *Let $F : M \longrightarrow N$ be a morphism of differentiable manifolds. Let $m \in M$. Then the following conditions are equivalent:*

(i) *F is a local diffeomorphism at m;*
(ii) *$T_m(F) : T_m(M) \longrightarrow T_{F(m)}(N)$ is an isomorphism.*

A morphism $F : M \longrightarrow N$ is an *immersion at m* if $T_m(F) : T_m(M) \longrightarrow T_{F(m)}(N)$ is injective. A morphism $F : M \longrightarrow N$ is an *submersion at m* if $T_m(F) : T_m(M) \longrightarrow T_{F(m)}(N)$ is surjective.

If F is an immersion at m, $\operatorname{rank}_m F = \dim_m M$, and by 1.2.2, this condition holds in an open neighborhood of m. Therefore, F is an immersion in a neighborhood of m.

Analogously, if F is an submersion at m, $\operatorname{rank}_m F = \dim_{F(m)} N$, and by 1.2.2, this condition holds in an open neighborhood of m. Therefore, F is an submersion in a neighborhood of m.

A morphism $F : M \longrightarrow N$ is an *subimmerson at m* if there exists a neighborhood U of m such that the rank of F is constant on U. By the above discussion, immersions and submersions at m are subimmersions at p.

A differentiable map $F : M \longrightarrow N$ is an *local diffeomorphism* if it is a local diffeomorphism at each point of M. A differentiable map $F : M \longrightarrow N$ is an *immersion* if it is an immersion at each point of M. A differentiable map $F : M \longrightarrow N$ is an *submersion* if it is an submersion ant each point of M. A differentiable map $F : M \longrightarrow N$ is an *subimmersion* if it is an subimmersion at each point of M. The rank of a subimmersion is constant on connected components of M.

1.3.2. Theorem. *Let $F : M \longrightarrow N$ be a subimmersion at $p \in M$. Assume that $\operatorname{rank}_m F = r$. Then there exists charts $c = (U, \varphi, m)$ and $d = (V, \psi, n)$ centered at p and $F(p)$ respectively, such that $F(U) \subset V$ and*

$$(\psi \circ F \circ \varphi^{-1})(x_1, \ldots, x_n) = (x_1, \ldots, x_r, 0, \ldots, 0)$$

for any $(x_1, \ldots, x_n) \in \varphi(U)$.

1.3.3. Corollary. *Let $i : M \longrightarrow N$ be an immersion. Let $F : P \longrightarrow M$ be a continuous map. Then the following conditions are equivalent:*

(i) *F is differentiable;*
(ii) *$i \circ F$ is differentiable.*

1.3.4. Corollary. *Let $p : M \longrightarrow N$ be a surjective submersion. Let $F : N \longrightarrow P$ be a map. Then the following conditions are equivalent:*

(i) *F is differentiable;*
(ii) *$F \circ p$ is differentiable.*

1.3.5. Corollary. *A submersion $F : M \longrightarrow N$ is an open map.*

1.4. Submanifolds. Let N be a subset of a differentiable manifold M. Assume that any point $n \in N$ has an open neighborhood U in M and a chart (U, φ, p) centered at n such that $\varphi(N \cap U) = \varphi(U) \cap \mathbb{R}^q \times \{0\}$. If we equip N with the induced topology and define its atlas consisting of charts on open sets $N \cap U$ given by the maps $\varphi : N \cap U \longrightarrow \mathbb{R}^q$, N becomes a differentiable manifold. With this differentiable structure, the natural inclusion $i : N \longrightarrow M$ is an immersion. The manifold N is called a *submanifold* of M.

1.4.1. Lemma. *A submanifold N of a manifold M is locally closed.*

1.4.2. Lemma. *Let $f : M \longrightarrow N$ be an injective immersion. If f is a homeomorphism of M onto $f(M) \subset N$, $f(M)$ is a submanifold in N and $f : M \longrightarrow f(M)$ is a diffeomorphism.*

Let $f : M \longrightarrow N$ is a differentiable map. Denote by Γ_f the *graph* of f, i.e., the subset $\{(m, f(m)) \in M \times N \mid m \in M\}$. Then, $\alpha : m \longmapsto (m, f(m))$ is a continuous bijection of M onto Γ_f. The inverse of α is the restriction of the canonical projection $p : M \times N \longrightarrow M$ to the graph Γ_f. Therefore, $\alpha : M \longrightarrow \Gamma_f$ is a homeomorphism. On the other hand, the differential of α is given by $T_m(\alpha)(\xi) = (\xi, T_m(f)(\xi))$ for any $\xi \in T_m(M)$, hence α is an immersion. By 1.4.2, we get the following result.

1.4.3. Lemma. *Let $f : M \longrightarrow N$ be a differentiable map. Then the graph Γ_f of f is a closed submanifold of $M \times N$.*

1.4.4. Lemma. *Let M and N be differentiable manifolds and $F : M \longrightarrow N$ a differentiable map. Assume that F is a subimmersion. Then, for any $n \in N$, $F^{-1}(n)$ is a closed submanifold of M and*

$$T_m(F^{-1}(n)) = \ker T_m(F).$$

for any $m \in F^{-1}(n)$.

In the case of submersions we have a stronger result.

1.4.5. Lemma. *Let $F : M \longrightarrow N$ be a submersion and P a submanifold of N. Then $F^{-1}(P)$ is a submanifold of M and the restriction $f|_{F^{-1}(P)} : F^{-1}(P) \longrightarrow P$ is a submersion. For any $m \in F^{-1}(P)$ we also have*

$$T_m(F^{-1}(P)) = T_m(F)^{-1}(T_{F(m)}(P)).$$

1.5. Products and fibered products. Let M and N be two topological spaces and $c = (U, \varphi, p)$ and $d = (V, \psi, q)$ two charts on M, resp. N. Then $(U \times V, \varphi \times \psi, p+q)$ is a chart on the product space $M \times N$. We denote this chart by $c \times d$.

Let M and N be two differentiable manifolds with atlases $\mathcal{A}$ and $\mathcal{B}$. Then $\{c \times d \mid c \in \mathcal{A}, d \in \mathcal{B}\}$ is an atlas on $M \times N$. The corresponding saturated atlas defines a structure of differentiable manifold on $M \times N$. This manifold is called the *product manifold* $M \times N$ of M and N.

Clearly $\dim_{(m,n)}(M \times N) = \dim_m M + \dim_n N$ for any $m \in M$ and $n \in N$..

The canonical projections to $pr_1 : M \times N \longrightarrow M$ and $pr_2 : M \times N \longrightarrow N$ are submersions. Moreover,

$$(T_{(m,n)}(pr_1), T_{(m,n)}(pr_2)) : T_{(m,n)}(M \times N) \longrightarrow T_m(M) \times T_n(N)$$

is an isomorphism of linear spaces for any $m \in M$ and $n \in N$.

Let M, N and P be differentiable manifolds and $F : M \longrightarrow P$ and $G : N \longrightarrow P$ differentiable maps. Then we put

$$M \times_P N = \{(m,n) \in M \times N \mid f(m) = g(n)\}.$$

This set is called the *fibered product* of M and N with respect to maps F and G.

1.5.1. LEMMA. *If $F : M \longrightarrow P$ and $G : N \longrightarrow P$ are submersions, the fibered product $M \times_P N$ is a closed submanifold of $M \times N$.*

The projections $p : M \times_P N \longrightarrow M$ and $q : M \times_P N \longrightarrow N$ are submersions.

For any $(m,n) \in M \times_P N$,

$$T_{(m,n)}(M \times_P N) = \{(X,Y) \in T_{(m,n)}(M \times N) \mid T_m(f)(X) = T_n(G)(Y)\}.$$

PROOF. Since F and G are submersions, the product map $F \times G : M \times N \longrightarrow P \times P$ is also a submersion. Since the diagonal Δ is a closed submanifold in $P \times P$, from 1.4.5 we conclude that the fiber product $M \times_P N = (F \times G)^{-1}(\Delta)$ is a closed submanifold of $M \times N$. Moreover, we have

$$T_{(m,n)}(M \times_P N) = \{(X,Y) \in T_{(m,n)}(M \times N) \mid T_m(F)(X) = T_n(G)(Y)\}.$$

Assume that $(m,n) \in M \times_P N$. Then $p = f(m) = g(n)$. Let $X \in T_m(M)$. Then, since G is a submersion, there exists $Y \in T_n(N)$ such that $T_n(G)(Y) = T_m(F)(X)$. Therefore, $(X,Y) \in T_{(m,n)}(M \times_P N)$. It follows that $p : M \times_P N \longrightarrow M$ is a submersion. Analogously, $q : M \times_P N \longrightarrow N$ is also a submersion. □

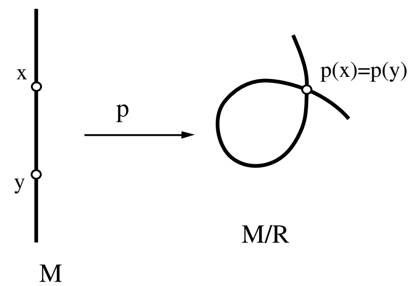
2. Quotients

2.1. Quotient manifolds. Let M be a differentiable manifold and $R \subset M \times M$ an equivalence relation on M. Let M/R be the set of equivalence classes of M with respect to R and $p : M \longrightarrow M/R$ the corresponding natural projection which attaches to any $m \in M$ its equivalence class $p(m)$ in M/R.

We define on M/R the quotient topology, i.e., we declare $U \subset M/R$ open if and only if $p^{-1}(U)$ is open in M. Then $p : M \longrightarrow M/R$ is a continuous map, and for any continuous map $F : M \longrightarrow N$, constant on the equivalence classes of R, there exists a unique continuous map $\bar{F} : M/R \longrightarrow N$ such that $F = \bar{F} \circ p$. Therefore, we have the commutative diagram

$$\begin{array}{ccc} M & \xrightarrow{F} & N . \\ {\scriptstyle p}\downarrow & \nearrow_{\bar{F}} & \\ M/R & & \end{array}$$

In general, M/R is not a manifold. For example, assume that $M = (0,1) \subset \mathbb{R}$, and R the union of the diagonal in $(0,1) \times (0,1)$ and $\{(x,y),(y,x)\}$ for $x, y \in (0,1)$, $x \neq y$. Then M/R is obtained from M by identifying x and y. Clearly this topological space doesn't allow a manifold structure.



Assume that M/R has a differentiable structure such that $p : M \longrightarrow M/R$ is a submersion. Since p is continuous, for any open set U in M/R, $p^{-1}(U)$ is open in M. Moreover, p is an open map by 1.3.5. Hence, for any subset $U \in M/R$ such that $p^{-1}(U)$ is open in M, the set $U = p(p^{-1}(U))$ is open in M/R. Therefore, a subset U in M/R is open if and only if $p^{-1}(U)$ is open in M, i.e., the topology on M/R is the quotient topology. Moreover, by 1.3.4, if the map F from M into a differentiable manifold N is differentiable, the map $\bar{F} : M/R \longrightarrow N$ is also differentiable.

We claim that such differentiable structure is unique. Assume the contrary and denote $(M/R)_1$ and $(M/R)_2$ two manifolds with these properties. Then, by the above remark, the identity maps $(M/R)_1 \longrightarrow (M/R)_2$ and $(M/R)_2 \longrightarrow (M/R)_1$ are differentiable. Therefore, the identity map is a diffeomorphism of $(M/R)_1$ and $(M/R)_2$, i.e., the differentiable structures on M/R are identical.

Therefore, we say that M/R is *the quotient manifold* of M with respect to R if it allows a differentiable structure such that $p : M \longrightarrow M/R$ is a submersion. In this case, the equivalence relation is called *regular*.

If the quotient manifold M/R exists, since $p : M \longrightarrow M/R$ is a submersion, it is also an open map.

2.1.1. Theorem. *Let M be a differentiable manifold and R an equivalence relation on M. Then the following conditions are equivalent:*

(i) *the relation R is regular;*
(ii) *R is closed submanifold of $M \times M$ and the restrictions $p_1, p_2 : R \longrightarrow M$ of the natural projections $pr_1, pr_2 : M \times M \longrightarrow M$ are submersions.*

The proof of this theorem follows from a long sequence of reductions. First we remark that it is enough to check the submersion condition in (ii) on only one map p_i, $i = 1, 2$. Let $s : M \times M \longrightarrow M \times M$ be given by $s(m,n) = (n,m)$ for $m, n \in M$. Then, $s(R) = R$ since R is symmetric. Since R is a closed submanifold and $s : M \times M \longrightarrow M \times M$ a diffeomorphism, $s : R \longrightarrow R$ is also a diffeomorphism. Moreover, $pr_1 = pr_2 \circ s$ and $pr_2 = pr_1 \circ s$, immediately implies that p_1 is a submersion if and only if p_2 is a submersion.

We first establish that (i) implies (ii). It is enough to remark that $R = M \times_{M/R} M$ with respect to the projections $p : M \longrightarrow M/R$. Then, by 1.5.1 we see that R is regular, i.e., it is a closed submanifold of $M \times M$ and $p_1, p_2 : R \longrightarrow M$ are submersions.

Now we want to prove the converse implication, i.e., that (ii) implies (i). This part is considerably harder. Assume that (ii) holds, i.e., R is a closed submanifold in $M \times M$ and $p_1, p_2 : R \longrightarrow M$ are submersions. We equip M/R with the quotient topology. Hence, $p : M \longrightarrow M/R$ is a continuous map. We first observe the following fact.

2.1.2. Lemma. *The map $p : M \longrightarrow M/R$ is open.*

Proof. Let $U \subset M$ be open. Then

$$
\begin{aligned}
p^{-1}(p(U)) &= \{m \in M \mid p(m) \in p(U)\} \\
&= \{m \in M \mid (m, n) \in R,\ n \in U\} = pr_1(R \cap (M \times U)) = p_1(R \cap (M \times U)).
\end{aligned}
$$

Clearly, $M \times U$ is open in $M \times M$, hence $R \cap (M \times U)$ is open in R. Since $p_1 : R \longrightarrow M$ is a submersion, it is an open map. Hence $p_1(R \cap (M \times U))$ is an open set in M. By the above formula it follows that $p^{-1}(p(U))$ is an open set in M. Therefore, $p(U)$ is open in M/R. □

Moreover, we have the following fact.

2.1.3. Lemma. *The quotient topology on M/R is hausdorff.*

Proof. Let $x = p(m)$ and $y = p(n)$, $x \neq y$. Then, $(m, n) \notin R$. Since R is closed in $M \times M$, there exist open neighborhoods U and V of m and n in M respectively, such that $U \times V$ is disjoint from R. Clearly, by 2.1.2, $p(U)$ and $p(V)$ are open neighborhoods of x and y respectively. Assume that $p(U) \cap p(V) \neq \emptyset$. Then there exists $r \in M$ such that $p(r) \in p(U) \cap p(V)$. It follows that we can find $u \in U$ and $v \in V$ such that $p(u) = p(r) = p(v)$. Therefore, $(u, v) \in R$, contrary to our assumption. Hence, $p(U)$ and $p(V)$ must be disjoint. Therefore, M/R is hausdorff. □

Now we are going to reduce the proof to a "local situation".

Let U be an open set in M. Since p is an open map, $p(U)$ is open in M/R. Then we put $R_U = R \cap (U \times U)$. Clearly, R_U is an equivalence relation on U. Let $p_U : U \longrightarrow U/R_U$ be the corresponding quotient map. Clearly, $(u, v) \in R_U$ implies $(u, v) \in R$ and $p(u) = p(v)$. Hence, the restriction $p|_U : U \longrightarrow M/R$ is constant on equivalence classes. This implies that we have a natural continuous map $i_U : U/R_U \longrightarrow M/R$ such that $p|_U = i_U \circ p_U$. Moreover, $i_U(U/R_U) = p(U)$. We claim that i_U is an injection. Assume that $i_U(x) = i_U(y)$ for some $x, y \in U/R_U$. Then $x = p_U(u)$ and $y = p_U(v)$ for some $u, v \in U$. Therefore,

$$p(u) = i_U(p_U(u)) = i_U(x) = i_U(y) = i_U(p_U(v)) = p(v)$$

and $(u, v) \in R$. Hence, $(u, v) \in R_U$ and $x = p_U(u) = p_U(v) = y$. This implies our assertion. Therefore, $i_U : U/R_U \longrightarrow p(U)$ is a continuous bijection. We claim that it is a homeomorphism. To prove this we have to show that it is open. Let V be an open subset of U/R_U. Then $p_U^{-1}(V)$ is open in U. On the other hand,

$$p_U^{-1}(V) = p_U^{-1}(i_U^{-1}(i_U(V))) = (p|_U)^{-1}(i_U(V)) = p^{-1}(i_U(V)) \cap U$$

is open in M. Since p is open, $p(p^{-1}(i_U(V)) \cap U)$ is open in M/R. Clearly, $p(p^{-1}(i_U(V)) \cap U) \subset i_U(V)$. On the other hand, if $y \in i_U(V)$, it is an equivalence class of an element $u \in U$. So, $u \in p^{-1}(i_U(V)) \cap U$. Therefore, $y \in p(p^{-1}(i_U(V)) \cap U)$. It follows that $p(p^{-1}(i_U(V)) \cap U) = i_U(V)$. Therefore, $i_U(V)$ is open in M/R and i_U is an open map. Therefore, $i_U : U/R_U \longrightarrow M/R$ is a homeomorphism of U/R_U onto the open set $p(U)$. To summarize, we have the following commutative diagram

$$\begin{array}{ccc} U & \longrightarrow & M \\ {\scriptstyle p_U}\downarrow & & \downarrow{\scriptstyle p} \\ U/R_U & \xrightarrow[i_U]{} & M/R \end{array}$$

where i_U is a homeomorphism onto the open set $p(U) \subset M/R$.

2.1.4. Remark. Since we already established that (i) implies (ii) in 2.1.1, we see that if R is regular, the present conditions are satisfied. In addition, M/R has a structure of a differentiable manifold and $p : M \longrightarrow M/R$ is a submersion. Since U/R_U is an open in M/R, it inherits a natural differentiable structure, and from the above diagram we see that p_U is a submersion. Therefore, R_U is also regular.

Assume now again that only (ii) holds for M. Then R_U is a closed submanifold of $U \times U$ and open submanifold of R. Therefore, the restrictions $p_i|_{R_U} : R_U \longrightarrow U$ are submersions. It follows that R_U satisfies the conditions of (ii).

We say that the subset U in M is *saturated* if it is a union of equivalence classes, i.e., if $p^{-1}(p(U)) = U$.

First we reduce the proof of the implication to the case local with respect to M/R.

2.1.5. Lemma. *Let $(U_i \mid i \in I)$ be an open cover of M consisting of saturated sets. Assume that all R_{U_i}, $i \in I$, are regular. Then R is regular.*

Proof. We proved that M/R is hausdorff. By the above discussion, for any $j \in I$, the maps $i_{U_j} : U_j/R_{U_j} \longrightarrow M/R$ are homeomorphisms of manifolds U_j/R_{U_j} onto open sets $p(U_j)$ in M/R. Clearly, $(p(U_j) \mid j \in I)$ is an open cover of M/R. Therefore, to construct a differentiable structure on M/R, it is enough to show that for any pair $(j,k) \in J \times J$, the differentiable structures on the open set $p(U_j) \cap p(U_k)$ induced by differentiable structures on $p(U_j)$ and $p(U_k)$ respectively, agree. Since U_j and U_k are saturated, $U_j \cap U_k$ is also saturated, and $p(U_j \cap U_k) = p(U_j) \cap p(U_k)$. From 2.1.4, we see that differentiable structures on $p(U_j)$ and $p(U_k)$ induce the quotient differentiable structure on $p(U_j \cap U_k)$ for the quotient of $U_j \cap U_k$ with respect to $R_{U_j \cap U_k}$. By the uniqueness of the quotient manifold structure, it follows that these induced structures agree. Therefore, by gluing these structures we get a differentiable structure on M/R. Since $p_{U_j} : U_j \longrightarrow U_j/R_{U_j}$ are submersions for all $j \in I$, we conclude that $p : M \longrightarrow M/R$ is a submersion. Therefore, R is regular. □

The next result will be used to reduce the proof to the saturated case.

2.1.6. Lemma. *Let U be an open subset of M such that $p^{-1}(p(U)) = M$. If R_U is regular, then R is also regular.*

Proof. As we already remarked, $i_U : U/R_U \longrightarrow M/R$ is a homeomorphism onto the open set $p(U)$. By our assumption, $p(U) = M/R$, so $i_U : U/R_U \longrightarrow M/R$ is a homeomorphism. Therefore, we can transfer the differentiable structure from U/R_U to M/R.

It remains to show that p is a submersion. Consider the following diagram

$$\begin{array}{ccc} (U \times M) \cap R & \xrightarrow{p_2} & M \\ {\scriptstyle p_1}\downarrow & & \downarrow{\scriptstyle p} \\ U & \xrightarrow[p_U]{} & U/R_U = M/R \end{array} .$$

It is clearly commutative. Since $p_i : R \longrightarrow M$, $i = 1, 2$, are submersions, their restrictions to the open submanifold $(U \times M) \cap R$ are also submersions. By our assumption, $p_U : U \longrightarrow U/R_U$ is also a submersion. Therefore, $p \circ p_2|_{(U \times M) \cap R} = p_U \circ p_1|_{(U \times M) \cap R} : (U \times M) \cap R \longrightarrow M$ is a submersion. By our assumption, $p_2|_{(U \times M) \cap R} : (U \times M) \cap R \longrightarrow M$ is also surjective. Therefore, $p : M \longrightarrow M/R$ is differentiable by 1.3.4. Moreover, since $p \circ p_2|_{(U \times M) \cap R}$ is a submersion, it also follows that p is a submersion. □

Now we can reduce the proof to a situation local in M.

2.1.7. Lemma. *Let $(U_i \mid i \in I)$ be an open cover of M such that R_{U_i} are regular for all $i \in I$. Then R is regular.*

Proof. Since p is open by 2.1.2, we see that $p(U_i)$ are all open. Therefore, $V_i = p^{-1}(p(U_i))$, $i \in I$, are open sets in M. They are clearly saturated. Moreover, since $U_i \subset V_i$ for $i \in I$, $(V_i \mid i \in I)$ is an open cover of M. Since R_{V_i} satisfy the conditions of (ii) and R_{U_i} are regular, by 2.1.6, we see that R_{V_i} are regular for $i \in I$. Therefore, by 2.1.5, we conclude that R is regular. □

It remains to treat the local case. Assume, for a moment, that R is regular. Let $m_0 \in M$. Then $N = p^{-1}(p(m_0))$ is the equivalence class of m_0, and it is a closed submanifold of M by 1.4.5. Also, the tangent space $T_{m_0}(N)$ to N at m_0 is equal to $\ker T_{m_0}(p) : T_{m_0}(M) \longrightarrow T_{p(m_0)}(M/R)$. On the other hand, since $R = M \times_{M/R} M$, by 1.5.1, we see that

$$T_{m_0,m_0}(R) = \{(X, Y) \in T_{m_0}(M) \times T_{m_0}(M) \mid T_{m_0}(p)(X) = T_{m_0}(p)(Y)\}.$$

Therefore, we have

$$T_{m_0}(N) = \{X \in T_{m_0}(M) \mid (X, 0) \in T_{(m_0,m_0)}(R)\}.$$

This explains the construction in the next lemma.

2.1.8. Lemma. *Let $m_0 \in M$. Then there exists an open neighborhood U of m_0 in M, a submanifold W of U containing m_0, and a differentiable map $r : U \longrightarrow W$ such that for any $m \in U$ the point $r(m)$ is the unique point in W equivalent to m.*

Proof. Let

$$E = \{X \in T_{m_0}(M) \mid (X, 0) \in T_{(m_0,m_0)}(R)\}.$$

Let F be a direct complement of the linear subspace E in $T_{m_0}(M)$. Denote by W' a submanifold of M such that $m_0 \in W'$ and $F = T_{m_0}(W')$. Put $\Sigma = (W' \times M) \cap R$.

Since $p_1 : R \longrightarrow M$ is a submersion, by 1.4.5 we see that $\Sigma = p_1^{-1}(W')$ is a submanifold of R. Moreover, we have

$$T_{(m_0,m_0)}(\Sigma) = \{(X,Y) \in T_{(m_0,m_0)}(R) \mid X \in T_{m_0}(W')\}$$
$$= \{(X,Y) \in T_{(m_0,m_0)}(R) \mid X \in F\}.$$

Let $\phi = p_2|\Sigma$, then $\phi : \Sigma \longrightarrow M$ is a differentiable map. In addition, we have

$$\ker T_{(m_0,m_0)}(\phi) = \{(X,0) \in T_{(m_0,m_0)}(\Sigma)\} = \{(X,0) \in T_{(m_0,m_0)}(R) \mid X \in F\}.$$

On the other hand, $(X,0) \in T_{(m_0,m_0)}(R)$ implies that $X \in E$, hence for any X in the above formula we have $X \in E \cap F = \{0\}$. Therefore, $\ker T_{(m_0,m_0)}(\phi) = 0$ and ϕ is an immersion at m_0.

Let $Y \in T_{m_0}(M)$. Then, since $p_2 : R \longrightarrow M$ is a submersion, there exists $X \in T_{m_0}(M)$ such that $(X,Y) \in T_{(m_0,m_0)}(R)$. Put $X = X_1 + X_2$, $X_1 \in E$, $X_2 \in F$. Then, since $(X_1, 0) \in T_{(m_0,m_0)}(R)$, we have

$$(X_2, Y) = (X,Y) - (X_1, 0) \in T_{(m_0,m_0)}(R).$$

Therefore, $(X_2, Y) \in T_{(m_0,m_0)}(\Sigma)$ and ϕ is also a submersion at (m_0, m_0). It follows that ϕ is a local diffeomorphism at (m_0, m_0). Hence, there exist open neighborhoods U_1 and U_2 of m_0 in M such that $\phi : \Sigma \cap (U_1 \times U_1) \longrightarrow U_2$ is a diffeomorphism. Let $f : U_2 \longrightarrow \Sigma \cap (U_1 \times U_1)$ be the inverse map. Then $f(m) = (r(m), m)$ for any $m \in U_2$, where $r : U_2 \longrightarrow U_1$ is a differentiable map. Since $\phi : \Sigma \cap (U_1 \times U_1) \longrightarrow U_2$ is surjective, we have $U_2 \subset U_1$. Let $m \in U_2 \cap W'$. Then we have $(m,m) \in (W' \times M) \cap R = \Sigma$. Hence, it follows that $(m,m) \in \Sigma \cap (U_1 \times U_1)$. Also, since $m \in U_2$, $(r(m), m) = f(m) \in \Sigma \cap (U_1 \times U_1)$. Clearly,

$$\phi(m,m) = p_2(m,m) = m = p_2(r(m), m) = \phi(r(m), m)$$

and since $\phi : \Sigma \cap (U_1 \times U_1) \longrightarrow U_2$ is an injection, we conclude that $r(m) = m$. Therefore, $r(m) = m$ for any $m \in U_2 \cap W'$.

Finally, since r is a differentiable map from U_2 into W', we can define open sets

$$U = \{m \in U_2 \mid r(m) \in U_2 \cap W'\} \text{ and } W = U \cap W'.$$

We have to check that U, W and r satisfy the assertions of the lemma. First, $m_0 \in U_2 \cap W'$ and $r(m_0) = m_0 \in U_2 \cap W'$. Hence, $m_0 \in U$, i.e., U is an open neighborhood of m_0. Now, we show that $r(U) \subset W$. By definition of U, for $m \in U$ we have $r(m) \in U_2 \cap W'$. Hence $r(r(m)) = r(m) \in U_2 \cap W'$. This implies that $r(m) \in U$. Hence, $r(m) \in W$. Since W is an open submanifold of W', $r : U \longrightarrow W$ is differentiable.

Let $m \in U$. Then $(r(m), m) = f(m) \in R$, i.e., $r(m)$ is in the same equivalence class as m. Assume that $n \in W$ is in the same class as m. Then

$$(n,m) \in (W \times U) \cap R \subset \Sigma \cap (U \times U)$$

and $\phi(n,m) = p_2(n,m) = m = \phi(r(m), m)$. Since $\phi : \Sigma \cap (U_1 \times U_1) \longrightarrow U_2$ is an injection, we see that $n = r(m)$. Therefore, $r(m)$ is the only point in W equivalent to m. □

Now we can complete the proof of the theorem. Let $m_0 \in M$ and (U, W, r) the triple satisfying 2.1.8. Let $i : W \longrightarrow U$ be the natural inclusion. Then $r \circ i = id$. Therefore, $T_{m_0}(r) \circ T_{m_0}(i) = 1_{T_{m_0}(W)}$ and r is a submersion at m_0. Therefore, there exists an open neighborhood V of m_0 contained in U such that $r : V \longrightarrow W$

is a submersion. Let $W_1 = r(V)$. Then W_1 is open in W. We have the following commutative diagram

$$\begin{array}{ccc} V & \xrightarrow{r} & W_1 \\ {\scriptstyle p_V}\downarrow & \nearrow{\scriptstyle \beta} & \\ V/R_V & & \end{array}.$$

Clearly, β is a continuous bijection. We claim that β is a homeomorphism. Let O be an open set in V/R_V. Then $p_V^{-1}(O)$ is open in V. Since r is a submersion, it is an open map. Hence, $r(p_V^{-1}(O)) = \beta(p_V(p_V^{-1}(O))) = \beta(O)$ is open. It follows that β is also an open map, i.e, a homeomorphism. Hence, we can pull the differentiable structure from W_1 to V/R_V. Under this identification, p_V corresponds to r, i.e., it is a submersion. Therefore, R_V is regular. This shows that any point in M has an open neighborhood V such that R_V is regular. By 2.1.7, it follows that R is regular. This completes the proof of the theorem.

2.1.9. PROPOSITION. *Let M be a differentiable manifold and R a regular equivalence relation on M. Denote by $p : M \longrightarrow M/R$ the natural projection of M onto M/R. Let $m \in M$ and N the equivalence class of m. Then N is a closed submanifold of M and*

$$\dim_m N = \dim_m M - \dim_{p(m)} M/R.$$

PROOF. Clearly, $N = p^{-1}(p(m))$ and the assertion follows from 1.4.5 and the fact that $p : M \longrightarrow M/R$ is a submersion. □

In particular, if M is connected, M/R is also connected and all equivalence classes have the same dimension equal to $\dim M - \dim M/R$.

Let M and N be differentiable manifolds and R_M and R_N regular equivalence relations relation on M and N, respectively. Then we can define an equivalence relation R on $M \times N$ by putting $(m, n) \sim (m', n')$ if and only if $(m, m') \in R_M$ and $(n, n') \in R_N$. Consider the diffeomorphism $q : M \times M \times N \times N \longrightarrow M \times N \times M \times N$ given by $q(m, m', n, n') = (m, n, m', n')$ for $m, m' \in M$ and $n, n' \in N$. It clearly maps the closed submanifold $R_M \times R_N$ onto R. Therefore, R is a closed submanifold of $M \times N \times M \times N$. If we denote by $p_{M,i} : R_M \longrightarrow M$, $p_{N,i} : R_N \longrightarrow N$ and $p_i : R \longrightarrow M \times N$ the corresponding projections, we have the following commutative diagram

$$\begin{array}{ccc} M \times N & \xrightarrow{q} & R \\ {\scriptstyle p_{M,i} \times p_{N,i}}\downarrow & \swarrow{\scriptstyle p_i} & \\ M \times N & & \end{array}.$$

This implies that R is regular and $(M \times N)/R$ exists. Moreover, if we denote by $p_M : M \longrightarrow M/R_M$, $p_N : N \longrightarrow N/R_N$ and $p : M \times N \longrightarrow (M \times N)/R$, it clear that the following diagram is commutative

$$\begin{array}{ccc} M \times N & \xrightarrow{\;p\;} & (M \times N)/R \\ {\scriptstyle p_M \times p_N}\big\downarrow & \swarrow & \\ M/R_M \times N/R_N & & \end{array}$$

where all maps are differentiable and the horizontal maps are also submersions. Since $(M \times N)/R \longrightarrow M/R_M \times N/R_N$ is a bijection, it is also a diffeomorphism. Therefore, we established the following result.

2.1.10. Lemma. *Let M and N be differentiable manifolds and R_M and R_N regular equivalence relations on M and N respectively. Then the equivalence relation*

$$R = \{((m,n),(m',n')) \mid (m,m') \in R_M,\ (n,n') \in R_N\}$$

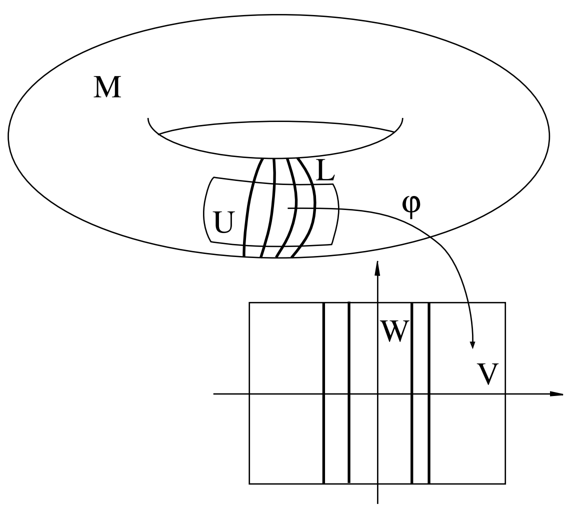
is regular. Moreover, the natural map $(M \times N)/R \longrightarrow M/R_M \times N/R_N$ is a diffeomorphism.

3. Foliations

3.1. Foliations. Let M be a differentiable manifold. Let $i : L \longrightarrow M$ be an immersion of a differentiable manifold L such that

(i) i is a bijection;
(ii) for any $m \in M$ there exist a chart (U, φ, n) at m; the integers $p, q \in \mathbb{Z}_+$ such that $p + q = n$; and connected open sets $V \subset \mathbb{R}^p$, $W \subset \mathbb{R}^q$ such that
  (a) $\varphi(U) = V \times W$;
  (b) $(\varphi \circ i)^{-1}(\{v\} \times W)$ is open in L for any $v \in V$;
  (c) $\varphi \circ i : (\varphi \circ i)^{-1}(\{v\} \times W) \longrightarrow \{v\} \times W$ is a diffeomorphism for any $v \in V$.

The pair (L, i) is called a *foliation* of M.



Let $m \in M$. Then the connected component of L containing $i^{-1}(m)$ is called the *leaf* of L through m. We denote it by L_m. The map $i|_{L_m} : L_m \longrightarrow M$ is an immersion since L_m is open in L. In general, L_m is not a submanifold of M.

Clearly, the function $m \longrightarrow \dim L_m$ is locally constant. Therefore, all leaves of L lying in the same connected component of M have the same dimension.

Let $T(M)$ be the tangent bundle of M. Let E be a vector subbundle of $T(M)$. We say that E is *involutive* if the submodule of the $C^\infty(M)$-module of all vector fields on M consisting of sections of E is closed under the Lie bracket $[X,Y] = X \circ Y - Y \circ X$, i.e., if for any two differentiable vector fields X and Y on M such that $X_m, Y_m \in E_m$ for all $m \in M$, we have $[X,Y]_m \in E_m$ for all $m \in M$.

3.1.1. Lemma. *Let (L,i) be a foliation of M. Then $T(i)T(L)$ is an involutive subbundle of $T(M)$.*

Proof. Let $m \in M$. Assume that $s \in L$ such that $m = i(s)$. There exists a chart $c = (U, \varphi, n)$ centered at m such that $\varphi(U) = V \times W$ for connected open sets $V \in \mathbb{R}^p$, $W \in \mathbb{R}^q$ such that $(\varphi \circ i)^{-1}(\{v\} \times W)$ is an open set in L. Denote by ∂_j, $1 \le j \le n$, the vector fields on U which correspond to the partial derivatives with respect to the j-th coordinate in $\mathbb{R}^n$ under the diffeomorphism φ. Then $T_r(i)T_r(L) \subset T_{i(r)}(M)$ is spanned by vectors $(\partial_j)_r$, $p+1 \le j \le n$, for any $r \in i^{-1}(U)$. Therefore, $T(i)T(L)$ is a vector subbundle of $T(M)$. Moreover, if X and Y are two vector fields on M such that their values are in $T(i)T(L)$, we have $X = \sum_{j=p+1}^n f_j \partial_j$ and $Y = \sum_{j=p+1}^n g_j \partial_j$ on U. Therefore, we have

$$[X,Y] = \sum_{j,k=p+1}^n [f_j\partial_j, g_k\partial_k] = \sum_{j,k=p+1}^n \big(f_j\partial_j(g_k)\partial_k - g_k\partial_k(f_j)\partial_j\big)$$
$$= \sum_{j,k=p+1}^n \big(f_j\partial_j(g_k) - g_j\partial_j(f_k)\big)\partial_k$$

and the value of the vector field $[X,Y]$ is in $L_r(i)T_r(L)$ for any $r \in i^{-1}(U)$. □

In the next section we are going to prove the converse of this result.

3.2. Frobenius theorem. Let E be an involutive vector subbundle of $T(M)$. An *integral manifold* of E is a pair (N,j) where

(i) N is a differentiable manifold;
(ii) $j : N \longrightarrow M$ is an injective immersion;
(iii) $T_s(j)T_s(N) = E_{j(s)}$ for all $s \in N$.

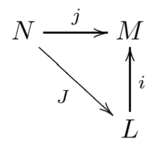
If $m = j(s)$ we say that (N,j) is an integral manifold through $m \in M$.

The observation 3.1.1 has the following converse.

3.2.1. Theorem (Frobenius). *Let M be a differentiable manifold and E an involutive vector subbundle of $T(M)$. Then there exists a foliation (L,i) of M with the following properties:*

(i) *(L,i) is an integral manifold for E;*

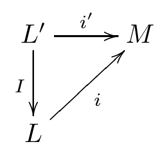
(ii) *for any integral manifold* (N, j) *of* E *there exists a unique differentiable map* $J : N \longrightarrow L$ *such that the diagram*



commutes and $J(N)$ *is an open submanifold of* L.

3.2.2. Remark. The map $J : N \longrightarrow J(N)$ is a diffeomorphism. First, J is an injective immersion. In addition, for any $s \in N$, we have $\dim T_{J(s)}(L) = \dim E_{j(s)} = \dim T_s(N)$ since L and N are integral manifolds. Hence J is also a submersion.

This also implies that the pair (L, i) is unique up to a diffeomorphism. If we have two foliations (L, i) and (L', i') which are integral manifolds for E, then we have a commutative diagram



where the mapping $I : L' \longrightarrow L$ is a diffeomorphism.

The pair (L, i) is the *integral foliation* of M with respect to E.

The proof of Frobenius theorem is based on the following local version of the result.

3.2.3. Lemma. *Let* $m \in M$, $n = \dim_m M$ *and* $q = \dim E_m$. *Then there exists a chart* $c = (U, \varphi, n)$ *centered at* m *and connected open sets* $V \subset \mathbb{R}^p$ *and* $W \subset \mathbb{R}^q$ *such that* $\varphi(U) = V \times W$ *and* $(\{v\} \times W, \varphi^{-1}|_{\{v\}\times W})$ *is an integral manifold of* E *for any* $v \in V$.

Since $\{v\} \times W$ are submanifolds of $\varphi(U)$, $\varphi^{-1}(\{v\} \times W)$ are submanifolds of M.

We postpone the proof of 3.2.3, and show first how it implies the global result.

3.2.4. Lemma. *Let* (N, j) *be a connected integral manifold of* E *such that* $j(N) \subset U$. *Then there exists* $v \in V$ *such that* $j(N) \subset \varphi^{-1}(\{v\} \times W)$ *and* $j(N)$ *is an open submanifold of* $\varphi^{-1}(\{v\} \times W)$.

Proof. Let $p_1 : V \times W \longrightarrow V$ be the projection to the first factor. Then $p_1 \circ \varphi \circ N \longrightarrow V$ is a differentiable map and for any $r \in N$ we have

$$\begin{aligned}(T_{(\varphi\circ j)(r)}(p_1) \circ T_{j(r)}(\varphi) \circ T_r(j))(T_r(N)) &= (T_{(\varphi\circ j)(r)}(p_1) \circ T_{j(r)}(\varphi))(E_{j(r)}) \\ &= T_{(\varphi\circ j)(r)}(p_1)(\{0\} \times \mathbb{R}^q) = \{0\},\end{aligned}$$

i.e., the differential of $p_1 \circ \varphi \circ j$ is equal to 0 and, since N is connected, this map is constant. It follows that there exists $v \in V$ such that $(\varphi \circ j)(N) \subset \{v\} \times W$. □

Let

$$\mathcal{B} = \{j(N) \mid (N, j) \text{ is an integral manifold of } E\}.$$

3.2.5. Lemma. *The family* $\mathcal{B}$ *is a basis of a topology on* M *finer than the natural topology of* M.

PROOF. Let O_1 and O_2 be two elements of $\mathcal{B}$ such that $O_1 \cap O_2 \neq \emptyset$. Let $r \in O_1 \cap O_2$. We have to show that there exists $O_3 \in \mathcal{B}$ such that $r \in O_3 \subset O_1 \cap O_2$.

Let (U, φ, n) be a chart around r satisfying 3.2.3. Let $O_1 = j_1(N_1)$ and $O_2 = j_2(N_2)$ for two integral manifolds (N_i, i_i), $i = 1, 2$, of E. Let C_1 and C_2 be the connected components of $j_1^{-1}(U)$, resp. $j_2^{-1}(U)$, containing $j_1^{-1}(r)$, resp. $j_2^{-1}(r)$. Then C_1, resp. C_2, are open submanifolds of N_1, resp. N_2, and $(C_1, j|_{C_1})$, resp. $(C_2, j|_{C_2})$, are integral manifolds through r. By 3.2.4, there exists $v \in V$ such that $r \in \varphi^{-1}(\{v\} \times W)$ and $j_1(C_1)$ and $j_2(C_2)$ are open submanifolds of $\varphi^{-1}(\{v\} \times W)$ which contain r. Therefore, $O_3 = j_1(C_1) \cap j_2(C_2)$ is an open submanifold of $\varphi^{-1}(\{v\} \times W)$. Hence O_3 is an integral manifold through r and $O_3 \in \mathcal{B}$.

Since we can take U to be arbitrarily small open set, the topology defined by $\mathcal{B}$ is finer than the naturally topology of M. □

Let L be the topological space obtained by endowing the set M with the topology with basis $\mathcal{B}$. Let $i : L \longrightarrow M$ be the natural bijection. By 3.2.5, the map i is continuous. In particular, the topology of L is hausdorff.

Let $l \in L$. By 3.2.3, there exists a chart (U, φ, n) around l, and $v \in V$ such that $(\varphi^{-1}(\{v\} \times W), i)$ is an integral manifold through l. By the definition of the topology on L, $\varphi^{-1}(\{v\} \times W)$ is an open neighborhood of l in L. Any open subset of $\varphi^{-1}(\{v\} \times W)$ in topology of L is an open set of $\varphi^{-1}(\{v\} \times W)$ as a submanifold of M. Therefore, $i : \varphi^{-1}(\{v\} \times W) \longrightarrow M$ is a homeomorphism on its image. Clearly, $\varphi^{-1}(\{v\} \times W)$ has the natural structure of differentiable submanifold of M. We can transfer this structure to $\varphi^{-1}(\{v\} \times W)$ considered as an open subset of L. In this way, L is covered by open subsets with structure of a differentiable manifold. On the intersection of any two of these open sets these differentiable structures agree (since they are induced as differentiable structures of submanifolds of M). Therefore, we can glue them together to a differentiable manifold structure on L. clearly, for that structure, $i : L \longrightarrow M$ is an injective immersion. Moreover, it is clear that (L, i) is an integral manifold for E. This completes the proof of (i).

Let (N, j) be an integral manifold of E. We define $J = i^{-1} \circ j$. Clearly, J is an injection. Let $r \in N$ and $l \in L$ such that $j(r) = i(l)$. Then, by 3.2.3, there exists a chart (U, φ, n) around l, and $v \in V$ such that $(\varphi^{-1}(\{v\} \times W), i)$ is an integral manifold through l. Moreover, there exists a connected neighborhood O of $r \in N$ such that $j(O) \subset U$. By 3.2.4, it follows that $J(O)$ is an open submanifold in $\varphi^{-1}(\{v\} \times W)$. Therefore, $J|_O : O \longrightarrow \varphi^{-1}(\{v\} \times W)$ is differentiable. It follows that $J : N \longrightarrow L$ is differentiable. This completes the proof of (ii).

Now we have to establish 3.2.3. We start with the special case where the fibers of E are one-dimensional. In this case, the involutivity is automatic.

3.2.6. LEMMA. *Let $m \in M$. Let X be a vector field on M such that $X_m \neq 0$. Then there exists a chart (U, φ, n) around m such that X_U corresponds to ∂_1 under the diffeomorphism φ.*

PROOF. Since the assertion is local, we can assume that $U = \varphi(U) \subset \mathbb{R}^n$ and $m = 0 \in \mathbb{R}^n$. Also, since $X_m \neq 0$, we can assume that $X(x_1)(0) \neq 0$. We put

$$F_j(x_1, x_2, \ldots, x_n) = X(x_j)$$

for $1 \leq j \leq n$. Then we can consider the system of first order differential equations

$$\frac{d\varphi_j}{dt} = F_j(\varphi_1, \varphi_2, \ldots, \varphi_n)$$

for $1 \leq j \leq n$, with the initial conditions

$$\begin{aligned} \varphi_1(0, c_2, c_3, \dots, c_n) &= 0 \\ \varphi_2(0, c_2, c_3, \dots, c_n) &= c_2 \\ &\dots \\ \varphi_n(0, c_2, c_3, \dots, c_n) &= c_n \end{aligned}$$

for "small" c_i, $2 \leq i \leq n$. By the existence and uniqueness theorem for systems of first order differential equations, this system has a unique differentiable solutions φ_j, $1 \leq j \leq n$, which depend differentiably on $t, c_1, c_2, \dots, c_n$ for $|t| < \epsilon$ and $|c_j| < \epsilon$ for $2 \leq j \leq n$.

Consider the differentiable map $\Phi : (-\epsilon, \epsilon)^n \longrightarrow \mathbb{R}^n$ given by

$$\Phi(y_1, y_2, \dots, y_n) = (\varphi_1(y_1, y_2, \dots, y_n), \varphi_2(y_1, y_2, \dots, y_n), \dots, \varphi_n(y_1, y_2, \dots, y_n)).$$

Then $\Phi(0) = 0$. Moreover, The Jacobian determinant of this map at 0 is equal to

$$\begin{vmatrix} \partial_1\varphi_1(0) & \partial_2\varphi_1(0) & \dots & \partial_n\varphi_1(0) \\ \partial_1\varphi_2(0) & \partial_2\varphi_2(0) & \dots & \partial_n\varphi_2(0) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1\varphi_n(0) & \partial_2\varphi_n(0) & \dots & \partial_n\varphi_n(0) \end{vmatrix}$$
$$= \begin{vmatrix} F_1(0,0,\dots,0) & 0 & \dots & 0 \\ F_2(0,0,\dots,0) & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ F_n(0,0,\dots,0) & 0 & \dots & 1 \end{vmatrix} = F_1(0,0,\dots,0) = X(x_1)(0) \neq 0.$$

Therefore, Φ is a local diffeomorphism at 0. By reducing ϵ if necessary we can assume that $\Phi : (-\epsilon, \epsilon)^n \longrightarrow \mathbb{R}^n$ is a diffeomorphism onto its image which is contained in U.

Let $y = (y_1, y_2, \dots, y_n) \in (-\epsilon, \epsilon)^n$. Then

$$\begin{aligned} T_y(\Phi)((\partial_1)_y)(x_i) &= \partial_1(x_i \circ \Phi)(\Phi(y)) = \partial_1\varphi_i(\Phi(y)) \\ &= \frac{d\varphi_i}{dt}(\varphi_1(y_1, y_2, \dots, y_n), \varphi_2(y_1, y_2, \dots, y_n), \dots, \varphi_n(y_1, y_2, \dots, y_n)) \\ &= F_i(\varphi_1(y_1, y_2, \dots, y_n), \varphi_2(y_1, y_2, \dots, y_n), \dots, \varphi_n(y_1, y_2, \dots, y_n)) = X_{\Phi(y)}(x_i). \end{aligned}$$

Hence, X and $T(\Phi)\partial_1$ agree on coordinate functions x_i, $1 \leq i \leq n$. Since vector fields are uniquely determined by their action on these functions, $X = T(\Phi)\partial_1$. □

This proves 3.2.3 for vector subbundles such that $\dim E_m = 1$ for all $m \in M$. In this case, the involutivity condition is automatic. To see this, let $m \in M$. Then there exists a vector field X on an open set U around m such that X_s span E_s for any $s \in U$. Therefore, any vector field Y on U such that $Y_s \in E_s$ for all $s \in U$ is of the form $Y = fX$ for some $f \in C^\infty(U)$. Therefore, if Y, Z are two such vector fields, we have $Y = fX$, $Y = gX$ for $f, g \in C^\infty(U)$, and

$$[Y, Z] = [fX, gX] = fX(g)X - gX(f)X = (fX(g) - gX(f))X.$$

It follows that E is involutive.

By 3.2.6, by shrinking U if necessary, we can assume that there exists a chart (U, φ, n) around m such that $\varphi(U) = (-\epsilon, \epsilon) \times V$ where V is an open connected set in $\mathbb{R}^{n-1}$, and X corresponds to ∂_1. In this case $\varphi^{-1}((-\epsilon,\epsilon)\times\{v\})$ are the integral manifolds for E.

It remains to prove the induction the proof of 3.2.3. We assume that the assertion holds for all involutive vector subbundles with fibers of dimension $\le q-1$. Assume that $\dim E_m = q$ for all $m \in M$. Since the statement is local, we can assume, without any loss of generality, that M is an connected open set in $\mathbb{R}^n$ and $X_1, X_2, \dots, X_q$ are vector fields on M such that E_s is spanned by their values $X_{1,s}, X_{2,s}, \dots, X_{q,s}$ in $s \in M$. Since E is involutive,

$$[X_i, X_j] = \sum_{k=1}^{q} c_{ijk} X_k$$

with $c_{ijk} \in C^\infty(M)$. By 3.2.6, after shrinking M if necessary, we can also assume that $X_1 = \partial_1$. If we write

$$X_i = \sum_{j=1}^{n} A_{ij}\partial_j$$

we see that the values of $Y_1 = X_1$ and

$$Y_i = X_i - A_{i1}\partial_1$$

for $i = 2, \dots, q$, also span E_s at any $s \in M$. Therefore, we can assume, after relabeling, that

$$X_1 = \partial_1 \text{ and } X_i = \sum_{j=2}^{n} A_{ij}\partial_j \text{ for } i = 2, \dots, q.$$

Now, for $i, j \ge 2$, we have

$$\begin{aligned}[X_i, X_j] &= \sum_{k,l=2}^{n} [A_{ik}\partial_k, A_{jl}\partial_l] = \sum_{k,l=2}^{n} (A_{ik}\partial_k(A_{jl})\partial_l - A_{jl}\partial_l(A_{ik})\partial_k) \\ &= \sum_{k,l=2}^{n} (A_{ik}\partial_k(A_{jl}) - A_{jk}\partial_k(A_{il}))\,\partial_l = \sum_{k=2}^{n} B_{ijk}\partial_k.\end{aligned}$$

On the other hand, we have

$$[X_i, X_j] = \sum_{k=1}^{n} c_{ijk} X_k = c_{ij1}\partial_1 + \sum_{k,l=2}^{n} c_{ijk} A_{kl}\partial_l,$$

for all $i, j \ge 2$. Hence, we conclude that $c_{ij1} = 0$ for $i, j \ge 2$, i.e.,

$$[X_i, X_j] = \sum_{k=2}^{n} c_{ijk} X_k$$

for $i, j \ge 2$. By shrinking M even more, we can assume that $M = (-\epsilon, \epsilon) \times N$ where N is an open subset in $\mathbb{R}^{n-1}$. Clearly,

$$X_{i,(0,t)} = \sum_{j=2}^{n} A_{ij}(0,t)\partial_j$$

can be considered as a vector field Z_i on N. Moreover, $Z_{2,t}, \dots, Z_{q,t}$ span a $(q-1)$-dimensional subspace F_t of $T(N)$ for any $t \in N$. Therefore, they define a vector subbundle F of $T(N)$. By the above calculation, this subbundle is involutive. Therefore, by the induction assumption, by shrinking N we can assume that there exists a coordinate system $(y_2, \dots, y_n)$ on N such that the submanifolds given by $y_{q+1} = c_{q+1}, \dots, y_n = c_n$ for $|c_i| < \delta$ for $n-q+1 \le i \le n$ are integral submanifolds for F. Relabeling y_i, $2 \le i \le n$, as x_i, $2 \le i \le n$, defines a new coordinate system on M such that

$$X_i = \sum_{j=2}^{n} A_{ij} \partial_j$$

with

$$A_{ij}(0,t) = 0 \text{ for } q+1 \le j \le n,$$

for $2 \le i \le n$. Now, for $2 \le i, j \le n$, we have

$$\frac{\partial}{\partial x_1} A_{ij} = X_1(X_i(x_j)) = [X_1, X_j](x_j) = \sum_{k=1}^{n} c_{1jk} X_k(x_j) = \sum_{k=2}^{n} c_{1jk} A_{kj}.$$

It follows that, for any $q+1 \le j \le n$, the functions $\mathbf{A}_j = (A_{2j}, \dots, A_{nj})$, satisfy the linear system of first order differential equations

$$\frac{\partial}{\partial x_1} A_{ij} = \sum_{k=2}^{n} c_{1jk} A_{kj}$$

on $(-\epsilon, \epsilon) \times (-\delta, \delta)^{n-1}$ with the initial conditions

$$A_{ij}(0,t) = 0,$$

for $2 \le i \le n$. Therefore, by the uniqueness theorem for such systems, it follows that $A_{ij} = 0$ for $2 \le i \le n$ and $q+1 \le j \le n$.

Therefore, we finally conclude that $X_1 = \partial_1$ and $X_i = \sum_{j=2}^{q} A_{ij} \partial_j$ for $2 \le i \le q$. This implies that E_s is spanned by $\partial_{1,s}, \partial_{2,s}, \dots, \partial_{q,s}$ for all $s \in M$. Hence, the submanifolds given by the equations $x_{q+1} = c_{q+1}, \dots, x_n = c_n$, are integral manifolds for E. This completes the proof of 3.2.3.

3.3. Separable leaves. In general, a connected manifold M can have a foliation with only one leaf L such that $\dim(L) < \dim(M)$. In this section, we discuss some topological conditions under which this doesn't happen.

A topological space is called *separable* if it has a countable basis of open sets.

We start with some topological preparation.

3.3.1. Lemma. *Let M be a separable topological space and $\mathcal{U} = \{U_i \mid i \in I\}$ be an open cover of M. Then there exists a countable subcover of $\mathcal{U}$.*

Proof. Let $\mathcal{V} = \{V_n \mid n \in \mathbb{N}\}$ be a countable basis of the topology on M. Every U_i in $\mathcal{U}$ is a union of elements in $\mathcal{V}$. Therefore, there exists a subfamily $\mathcal{A}$ of $\mathcal{V}$ such that $V \in \mathcal{A}$ implies $V \subset U_i$ for some $i \in I$. Since $\mathcal{V}$ is a basis of the topology of M, $\mathcal{A}$ is a cover of M. For each $V \in \mathcal{A}$, we can pick U_i such that $V \subset U_i$. In this way we get a subcover of $\mathcal{U}$ which is countable. □

3.3.2. Lemma. *Let M be a connected topological space. Let $\mathcal{U} = \{U_i \mid i \in I\}$ be an open cover of M with the following properties:*

(i) *U_i are separable for all $i \in I$;*
(ii) *$\{j \in I \mid U_i \cap U_j \neq \emptyset\}$ is countable for each $i \in I$.*

Then M is separable.

PROOF. Let $i_0 \in I$ be such that $U_{i_0} \neq \emptyset$. We say that $i \in I$ is *accessible* in n steps from i_0 if there exists a sequence $(i_1, i_2, \dots, i_n)$, $i = i_n$, such that $U_{i_{k-1}} \cap U_{i_k} \neq \emptyset$ for $k = 1, 2, \dots n$.

Let A_n be the set of all indices accessible in n steps from i_0. We claim that A_n are countable. First, the condition (ii) implies that A_1 is countable. Assume that A_n is countable. If $j \in A_{n+1}$, there exists $i \in A_n$ such that $U_j \cap U_i \neq \emptyset$. Since A_n is countable and (ii) holds we conclude that A_{n+1} must be countable. Therefore $\mathcal{A} = \bigcup_{n=1}^{\infty} A_n$ is countable.

Let $U = \bigcup_{i \in \mathcal{A}} U_i$. Then U is an open subset of M. Since it contains U_{i_0} it must be nonempty. Let $m \in \bar{U}$. Then there exists $i \in I$ such that $m \in U_i$. Hence, we have $U_i \cap U \neq \emptyset$. It follows that $U_i \cap U_j \neq \emptyset$ for some $j \in \mathcal{A}$. If $j \in A_n$, we see that $i \in A_{n+1} \subset \mathcal{A}$. Therefore, we have $m \in U_i \subset U$. Hence, U is also closed. Since M is connected, U must be equal to M.

Therefore, M is a union of a countable family of separable open subsets U_i, $i \in \mathcal{A}$. The union of countable bases of topology on all U_i, $i \in A$, is a countable basis of topology of M. Therefore, M is also separable. □

3.3.3. LEMMA. *Let M be a locally connected, connected topological space. Let $\mathcal{U} = \{U_n \mid n \in \mathbb{N}\}$ be an open cover of M such that each connected component of U_n is separable. Then M is separable.*

PROOF. Since M is locally connected, the connected components of U_n, $n \in \mathbb{N}$, are open in M. Let $U_{n,\alpha}$, $\alpha \in A_n$, be the connected components of U_n. Therefore, $\mathcal{V} = \{U_{n,\alpha} \mid \alpha \in A_n, n \in \mathbb{N}\}$ is an open cover of M.

Let $A_{n,\alpha;m} = \{\beta \in A_m \mid U_{n,\alpha} \cap U_{m,\beta} \neq \emptyset\}$ for $\alpha \in A_n$, $n, m \in \mathbb{N}$. We claim that $A_{n,\alpha;m}$ is countable for any $\alpha \in A_n$, $n, m \in \mathbb{N}$. First we remark that the set $U_m \cap U_{n,\alpha}$ is open in $U_{n,\alpha}$, and since $U_{n,\alpha}$ is separable, $U_m \cap U_{n,\alpha}$ can have only countably many components. We denote them by S_p, $p \in \mathbb{N}$. Since S_p is connected, it must be contained in a unique connected component $U_{m,\beta(p)}$ of U_m. Let $\beta \in A_{n,\alpha;m}$. Then we have $U_{m,\beta} \cap U_{n,\alpha} \neq \emptyset$. If we take $s \in U_{m,\beta} \cap U_{n,\alpha}$, then s is in one of S_p. It follows that $\beta = \beta(p)$. It follows that $A_{n,\alpha;m}$ is countable. Hence, the cover $\mathcal{V}$ satisfies the conditions of 3.3.2, and M is separable. □

The main result which we want to establish is the following theorem.

3.3.4. THEOREM. *Let M be a differentiable manifold such that all of its connected components are separable. Let (L, i) be a foliation of M. Then all leaves of L are separable manifolds.*

PROOF. Let $m \in M$ and L_m be the leaf passing through m. We want to prove that L_m is separable. Since L_m is connected, it lies in a connected component of M. Therefore, we can replace M with this component, i.e., we can assume that M is connected and separable.

By 3.3.1, there exists a countable family of charts $c_n = (U_n, \varphi_n)$, $n \in \mathbb{N}$, such that U_n, $n \in \mathbb{N}$, cover M; $\varphi_n(U_n) = V_n \times W_n$, V_n and W_n are connected and $(\varphi_n \circ i)^{-1}(\{v\} \times W_n)$ are open in L and $(\varphi_n \circ i) : (\varphi_n \circ i)^{-1}(\{v\} \times W_n) \longrightarrow \{v\} \times W_n$ are diffeomorphisms for all $v \in V_n$ and $n \in \mathbb{N}$. Therefore, $\{i^{-1}(U_n); n \in \mathbb{N}\}$ is a countable cover of L. In addition, the connected components of $i^{-1}(U_n)$ are of the form $(\varphi_n \circ i)^{-1}(\{v\} \times W_n)$ for $v \in V_n$, hence they are separable. By 3.3.3, the leaf L_m is separable. □

3.3.5. Remark. A differentiable manifold has separable connected components if and only if it is paracompact. Therefore, 3.3.4 is equivalent to the statement that any foliation of a paracompact differentiable manifold is paracompact.

This result allows us to use the following observation.

3.3.6. Lemma. *Let M be a differentiable manifold and (L, i) a foliation with separable leaves. Let N be a differentiable manifold and $f : N \longrightarrow M$ a differentiable map such that $f(N)$ is contained in countably many leaves. Then there exists a differentiable map $F : N \longrightarrow L$ such that the diagram*

$$\begin{array}{ccc} N & \xrightarrow{f} & M \\ {\scriptstyle F}\downarrow & \nearrow_{i} & \\ L & & \end{array}$$

commutes.

Proof. Let $p \in N$ and (U, φ, n) a chart centered at $f(p)$ such that $\varphi(U) = V \times W$ where V and W are connected and such that $(\varphi \circ i)^{-1}(\{v\} \times W)$ are open in L and $\varphi \circ i : (\varphi \circ i)^{-1}(\{v\} \times W) \longrightarrow \{v\} \times W$ are diffeomorphisms for all $v \in V$. Since the leaves are separable, for a fixed leaf L_m passing through $m \in M$, we have $(\varphi \circ i)^{-1}(\{v\} \times W) \subset L_m$ for countably many $v \in V$. By our assumption, $f(N)$ intersect only countably many leaves, $((\varphi \circ f)^{-1}(\{v\} \times W)$ is nonempty for only countably many $v \in V$.

Let U' be a connected neighborhood of p such that $f(U') \subset U$. Denote by $pr_1 : V \times W \longrightarrow V$ the projection to the first factor. Then $(pr_1 \circ \varphi \circ f)|_{U'}$ maps U' onto a countable subset of V. Therefore, it is a constant map, i.e., $(\varphi \circ f)(U') \subset \{v_0\} \times W$ for some $v_0 \in V$. This implies that F is differentiable at p. □

3.3.7. Corollary. *Let M be a separable, connected differentiable manifold. Let (L, i) be a foliation of M. Then either $L = M$ or L consists of uncountably many leaves.*

Proof. Assume that L consists of countably many leaves. Then the identity map $id : M \longrightarrow M$ factors through L by 3.3.6. Therefore, $i : L \longrightarrow M$ is a diffeomorphism and $L = M$. □

4. Integration on manifolds

4.1. Change of variables formula. Let U and V be two open subsets in $\mathbb{R}^n$ and $\varphi : U \longrightarrow V$ a diffeomorphism of U on V. Then $\varphi(x_1, x_2, \dots, x_n) = (\varphi_1(x_1, x_2, \dots, x_n), \varphi_2(x_1, x_2, \dots, x_n), \dots, \varphi_n(x_1, x_2, \dots, x_n))$ with $\varphi_i : U \longrightarrow R$, $1 \le i \le n$, for all $(x_1, x_2, \dots, x_n) \in U$. Let

$$J(\varphi) = \begin{vmatrix} \frac{\partial \varphi_1}{\partial x_1} & \frac{\partial \varphi_1}{\partial x_2} & \dots & \frac{\partial \varphi_1}{\partial x_n} \\ \frac{\partial \varphi_2}{\partial x_1} & \frac{\partial \varphi_2}{\partial x_2} & \dots & \frac{\partial \varphi_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \varphi_n}{\partial x_1} & \frac{\partial \varphi_n}{\partial x_2} & \dots & \frac{\partial \varphi_n}{\partial x_n} \end{vmatrix}$$

be the Jacobian determinant of the mapping φ. Then, since φ is a diffeomorphism, $J(\varphi)(x_1, x_2, \dots, x_n) \ne 0$ for all $(x_1, x_2, \dots, x_n) \in U$.

Let f be a continuous function with compact support on V. Then we have the change of variables formula

$$\int_V f(y_1, y_2, \dots, y_n)\, dy_1\, dy_2 \dots dy_n \\ = \int_U f(\varphi(x_1, x_2, \dots, x_n))|J(\varphi)(x_1, x_2, \dots, x_n)|\, dx_1\, dx_2 \dots dx_n.$$

Let ω be the differential n-form with compact support in V. Then ω is given by a formula

$$\omega = f(y_1, y_2, \dots, y_n)\, dy_1 \wedge dy_2 \wedge \dots \wedge dy_n$$

for $(y_1, y_2, \dots, y_n) \in V$. On the other hand,

$$\varphi^*(\omega) = f(\varphi(x_1, x_2, \dots, x_n))\, J(\varphi)(x_1, x_2, \dots, x_n) dx_1 \wedge dx_2 \wedge \dots \wedge dx_n$$

for $(x_1, x_2, \dots, x_n) \in U$.

4.2. Positive measure associated to a differential form of top degree. Let M be a manifold of pure dimension n. Let ω be a differential n-form on M with compact support. Let $c = (U, \varphi, n)$ be a chart on M such that $\operatorname{supp} \omega \subset U$. Let $d = (V, \psi, n)$ be another chart such that $\operatorname{supp} \omega \subset V$. Clearly, $\operatorname{supp} \omega \subset U \cap V$. We can consider the differential n-forms $(\varphi^{-1})^*(\omega)$ on $\varphi(U) \subset \mathbb{R}^n$ and $(\psi^{-1})^*(\omega)$ on $\psi(V) \subset \mathbb{R}^n$. These forms are represented by

$$(\varphi^{-1})^*(\omega) = f_U(x_1, x_2, \dots, x_n)\, dx_1 \wedge dx_2 \wedge \dots \wedge dx_n,$$

for all $(x_1, x_2, \dots, x_n) \in \varphi(U)$, and

$$(\psi^{-1})^*(\omega) = f_V(y_1, y_2, \dots, y_n)\, dy_1 \wedge dy_2 \wedge \dots \wedge dy_n,$$

for all $(y_1, y_2, \dots, y_n) \in \psi(V)$, respectively. Moreover, $\alpha = \psi \circ \varphi^{-1} : \varphi(U \cap V) \longrightarrow \psi(U \cap V)$ is a diffeomorphism, and $\alpha^*((\varphi^{-1})^*(\omega)) = (\psi^{-1})^*(\omega)$. By the discussion in 4.1, we have

$$(\varphi^{-1})^*(\omega) = \alpha^*((\psi^{-1})^*(\omega)) \\ = f_V(\alpha(x_1, x_2, \dots, x_n))J(\alpha)(x_1, x_2, \dots, x_n)\, dx_1 \wedge dx_2 \wedge \dots \wedge dx_n$$

for all $(x_1, x_2, \dots, x_n) \in \varphi(U)$. Hence,

$$f_U(x_1, x_2, \dots, x_n) = f_V(\alpha(x_1, x_2, \dots, x_n))J(\alpha)(x_1, x_2, \dots, x_n)$$

for all $(x_1, x_2, \dots, x_n) \in \varphi(U)$.

For any continuous function g on M, by the change of variables formula in 4.1, we also see that

$$\begin{aligned}
&\int_{\varphi(U)} g(\varphi^{-1}(x_1, x_2, \dots, x_n))\, |f_U(x_1, x_2, \dots, x_n))|\, dx_1\, dx_2 \dots dx_n \\
&= \int_{\varphi(U)} g(\varphi^{-1}(x_1, \dots, x_n))\, |f_V(\alpha(x_1, \dots, x_n))|\, |J(\alpha)(x_1, \dots, x_n)|\, dx_1 \dots dx_n \\
&= \int_{\varphi(U\cap V)} g(\varphi^{-1}(x_1, \dots, x_n))\, |f_V(\alpha(x_1, \dots, x_n))|\, |J(\alpha)(x_1, \dots, x_n)|\, dx_1 \dots dx_n \\
&= \int_{\psi(U\cap V)} g(\psi^{-1}(y_1, y_2, \dots, y_n))\, |f_V(y_1, y_2, \dots, y_n)|\, dy_1\, dy_2 \dots dy_n \\
&= \int_{\psi(V)} g(\psi^{-1}(y_1, y_2, \dots, y_n))\, |f_V(y_1, y_2, \dots, y_n)|\, dy_1\, dy_2 \dots dy_n.
\end{aligned}$$

Therefore, the expression

$$\int_{\varphi(U)} g(\varphi^{-1}(x_1, x_2, \dots, x_n))\, |f_U(x_1, x_2, \dots, x_n))|\, dx_1\, dx_2 \dots dx_n$$

is independent of the choice of the chart c such that $\operatorname{supp}\omega \subset U$. Hence we can define

$$\int g|\omega| = \int_{\varphi(U)} g(\varphi^{-1}(x_1, x_2, \dots, x_n))\, |f_U(x_1, x_2, \dots, x_n))|\, dx_1\, dx_2 \dots dx_n,$$

for any chart $c = (U, \varphi, n)$ such that $\operatorname{supp}\omega \subset U$. The linear map $g \longmapsto \int g|\omega|$ defines a positive measure on M with compact support.

Now we want to extend this definition to differential n-forms on M with arbitrary compact support. Let ω be a differentiable n-form with support in a compact set K in M. Let $c_i = (U_i, \varphi_i, n)$, $1 \le i \le p$, be a finite cover of K by charts. Let α_i, $1 \le i \le p$, be a partition of unity such that

(i) α_i, $1 \le i \le p$, are positive smooth functions with compact support on M;
(ii) $\operatorname{supp}\alpha_i \subset U_i$ for all $1 \le i \le p$;
(iii) $\sum_{i=1}^p \alpha_i(m) = 1$ for all $m \in K$.

Then $\omega = \sum_{i=1}^p \alpha_i\omega$. Moreover, the differential n-forms $\alpha_i\omega$ are supported in U_i, hence the measures $|\alpha_i\omega|$ are well-defined.

We claim that the sum $\sum_{i=1}^p |\alpha_i\omega|$ is independent of the choice of the cover U_i and the partition α_i. Let $d_j = (V_j, \psi_j, n)$, $1 \le j \le q$, be another open cover of K by charts on M. Let β_j, $1 \le j \le q$, be the corresponding partition of unity. Then, we have

$$\begin{aligned}
\sum_{i=1}^p |\alpha_i\omega| = \sum_{i=1}^p \left(\sum_{j=1}^q \beta_j|\alpha_i\omega|\right) &= \sum_{j=1}^q \left(\sum_{i=1}^p |\alpha_i\beta_j\omega|\right) \\
&= \sum_{j=1}^q \left(\sum_{i=1}^p \alpha_i|\beta_j\omega|\right) = \sum_{j=1}^q |\beta_j\omega|,
\end{aligned}$$

and this establishes our claim. Therefore, we can define

$$\int g|\omega| = \sum_{i=1}^p \int g|\alpha_i\omega|$$

for any continuous function g on M.

Finally we want to extend the definition to arbitrary differentiable n-forms on M. Let K be a compact set in M and α a positive smooth function with compact support on M such that $\alpha(m) = 1$ for all $m \in K$. Then $\alpha\omega$ is a differentiable n-form with compact support on M. For any continuous function with support in K, the expression $\int g|\alpha\omega|$ doesn't depend on the choice of α. In fact, if β is another positive smooth function on M which is equal to 1 on K, we have

$$\int g|\alpha\omega| = \int g\beta|\alpha\omega| = \int g|\alpha\beta\omega| = \int g\alpha|\beta\omega| = \int g|\beta\omega|.$$

Therefore, we can define

$$\int g|\omega| = \int g|\alpha\omega|$$

for any continuous function g with compact support in M. Therefore, ω defines a positive measure $|\omega|$ on M.

From the construction of the positive measure associated to a differentiable n-form we deduce the following result.

4.2.1. Proposition. *Let M and N be differentiable manifolds and $\varphi : M \longrightarrow N$ a diffeomorphism of M onto N. Let ω be a differentiable n-form on N. Then*

$$\int_M (f \circ \varphi)\, |\varphi^*(\omega)| = \int_N f|\omega|$$

for any compactly supported continuous function f on N.

CHAPTER 2

Lie groups

1. Lie groups

1.1. Lie groups. A set G is a *Lie group* if

(i) G is a differentiable manifold;
(ii) G is a group;
(iii) the map $\alpha_G : (g,h) \longmapsto gh^{-1}$ from the manifold $G \times G$ into G is differentiable.

Let G be a Lie group. Denote by $m : G \times G \longrightarrow G$ the multiplication map $m(g,h) = gh$, $\iota : G \longrightarrow G$ the inversion map $\iota(g) = g^{-1}$ and by $i : G \longmapsto G \times G$ the inclusion $i(g) = (1,g)$. Then we have $\alpha_G \circ i = \iota$, hence the inversion map is differentiable. On the other hand, $m = \alpha_G \circ (1_G \times \iota)$, hence the multiplication map is also differentiable.

For $g \in G$, we define the *left translation* $\gamma(g) : G \longrightarrow G$ by $\gamma(g)(h) = gh$ for $h \in G$, and the *right translation* $\delta(g) : G \longrightarrow G$ by $\delta(g)(h) = hg^{-1}$ for $h \in G$. Clearly, left and right translations are diffeomorphisms. Therefore, the function $g \longmapsto \dim_g G$ is constant on G, i.e., the manifold G is of pure dimension.

Let V be a finite-dimensional linear space over $\mathbb{R}$. Then the group $\mathrm{GL}(V)$ of all linear automorphisms of V has a natural Lie group structure. It is called the *general linear group* of V.

A *morphism* $\phi : G \longrightarrow H$ of a Lie group G into a Lie group H is a group homomorphism which is also a morphism of differentiable manifolds.

Let G be a Lie group. Define the multiplication $(g,h) \longmapsto g \circ h = hg$. The set G with this operation is a group. Moreover, it is a Lie group. We call this Lie group G^{opp} the *opposite Lie group* of G. The map $g \longmapsto g^{-1}$ is an isomorphism of G onto G^{opp}. Evidently, we have $(G^{opp})^{opp} = G$.

Let H be a subgroup of G. If H is a submanifold of G we call it a *Lie subgroup* of G.

Let H be a Lie subgroup of G. Then we have the following commutative diagram:

$$\begin{array}{ccc} H \times H & \longrightarrow & G \times G \\ {\scriptstyle \alpha_H}\Big\downarrow & & \Big\downarrow{\scriptstyle \alpha_G} \\ H & \longrightarrow & G \end{array} .$$

Clearly, the map $\alpha_H : H \times H \longrightarrow G$ is differentiable. This in turn implies that $\alpha_H : H \times H \longrightarrow H$ is differentiable and H is a Lie group.

Clearly, the map $i : H \longrightarrow G$ is a morphism of Lie groups.

By its definition a Lie subgroup is locally closed.

1.1.1. Lemma. *Let G be a topological group and H its locally closed subgroup. Then H is closed in G.*

PROOF. Let x be a point in the closure of H. Let V be a symmetric open neighborhood of 1 in G such that $V \cap H$ is closed in V. Then xV is a neighborhood of x and since x is in the closure of H, $xV \cap H$ is nonempty. Let $y \in xV \cap H$. Then, $x \in yV$. Moreover, $y(V \cap H) = yV \cap H$ is closed in yV. Assume that x is not in H. Then there exists an open neighborhood U of x in yV such that $U \cap H = \emptyset$. But this clearly contradicts our choice of x. Hence, $x \in H$. □

Therefore, we have the following obvious consequence.

1.1.2. COROLLARY. *Any Lie subgroup H of a Lie group G is closed in G.*

A *(left) differentiable action* of G on a manifold M is a differentiable map $\mu : G \times M \longrightarrow M$ satisfying

(i) $\mu(1_G, m) = m$ for all $m \in M$;
(ii) $\mu(g, \mu(h, m)) = \mu(m(g, h), m)$ for all $g, h \in G$ and $m \in M$, i.e., the diagram

$$\begin{array}{ccc} G \times G \times M & \xrightarrow{id_G \times \mu} & G \times M \\ {\scriptstyle m \times id_M}\big\downarrow & & \big\downarrow{\scriptstyle \mu} \\ G \times M & \xrightarrow[\mu]{} & G \end{array}$$

is commutative.

Clearly, $\phi : G \times G \longrightarrow G$ defined by $\phi(g, h) = \gamma(g)h$ and $\psi(g, h) = \delta(g)h$, $g, h \in G$, respectively, define differentiable actions of G on G by left and right translations respectively.

Let $\mu : G \times M \longrightarrow M$ be a differentiable action of G on M. We denote $\mu(g, m) = g \cdot m$ for $g \in G$ and $m \in M$. For any $g \in G$ we define the map $\tau(g) : M \longrightarrow M$ by $\tau(g)(m) = g \cdot m$ for any $m \in M$. It is easy to check that $\tau(gh) = \tau(g)\tau(h)$. Moreover, $\tau(g)$ is differentiable. Hence, for any $g \in G$, $\tau(g)$ is a diffeomorphism of M with inverse $\tau(g^{-1})$.

The set $\Omega = \{g \cdot m \mid g \in G\}$ is called the *G-orbit* of $m \in M$. The differentiable map $\rho(m) : G \longrightarrow M$ given by $\rho(m)(g) = g \cdot m$ is the *orbit map* of m. Its image is the orbit Ω.

The action of G on M is *transitive* if M is a G-orbit.

The set $G_m = \{g \in G \mid gm = m\} = \rho(m)^{-1}(m)$ is a subgroup of G which is called the *stabilizer* of m in G.

1.1.3. LEMMA. *For any $m \in M$, the orbit map $\rho(m) : G \longrightarrow M$ has constant rank. In particular, $\rho(m)$ is a subimmersion.*

PROOF. For any $a, b \in G$ we have

$$(\tau(a) \circ \rho(m))(b) = \tau(a)(b \cdot m) = (ab) \cdot m = \rho(m)(ab) = (\rho(m) \circ \gamma(a))(b),$$

i.e., we have

$$\tau(a) \circ \rho(m) = \rho(m) \circ \gamma(a)$$

for any $a \in G$. If we calculate the differential of this map at the identity in G we get

$$T_m(\tau(a)) \circ T_1(\rho(m)) = T_a(\rho(m)) \circ T_1(\gamma(a))$$

for any $a \in G$. Since $\tau(a)$ and $\gamma(a)$ are diffeomorphisms, their differentials $T_m(\tau(a))$ and $T_1(\gamma(a))$ are isomorphisms of tangent spaces. This implies that $\operatorname{rank} T_1(\rho(m)) =$

rank $T_a(\rho(m))$ for any $a \in G$. Hence the function $a \longmapsto \operatorname{rank}_a \rho(m)$ is constant on G. □

By 1.1.4.4, we have the following consequence.

1.1.4. PROPOSITION. *For any $m \in M$, the stabilizer G_m is a Lie subgroup of G. In addition, $T_1(G_m) = \ker T_1(\rho(m))$.*

Let G and H be Lie groups and $\phi : G \longrightarrow H$ a morphism of Lie groups. Then we can define a differentiable action of G on H by $(g, h) \longmapsto \phi(g)h$ for $g \in G$ and $h \in H$. The stabilizer in G of $1 \in H$ is the Lie subgroup $\ker \phi = \{g \in G \mid \phi(g) = 1\}$. Therefore, we have the following result.

1.1.5. PROPOSITION. *Let $\phi : G \longrightarrow H$ be a morphism of Lie groups. Then:*

(i) *The kernel $\ker \phi$ of a morphism $\phi : G \longrightarrow H$ of Lie groups is a normal Lie subgroup of G.*
(ii) $T_1(\ker \phi) = \ker T_1(\phi)$.
(iii) *The map $\phi : G \longrightarrow H$ is a subimmersion.*

On the contrary the image of a morphism of Lie groups doesn't have to be a Lie subgroup.

1.2. Orbit manifolds. Let G be a Lie group acting on a manifold M. We define an equivalence relation R_G on M by

$$R_G = \{(g \cdot m, m) \in M \times M \mid g \in G,\ m \in M\}.$$

The equivalence classes for this relation are the G-orbits in M. The quotient M/R_G is called the *orbit space* of M and denoted by M/G.

The next result is a variant of 2.1.1 for Lie group actions.

1.2.1. THEOREM. *Let G be a Lie group acting differentiably on a manifold M. Then the following conditions are equivalent:*

(i) *the relation R_G is regular;*
(ii) *R_G is a closed submanifold in $M \times M$.*

PROOF. First, from 2.1.1, it is evident that (i) implies (ii).

To prove that (ii) implies (i), by 2.1.1, we just have to show that $p_2 : R_G \longrightarrow M$ is a submersion.

Define the map $\theta : G \times M \longrightarrow M \times M$ by $\theta(g, m) = (g \cdot m, m)$ for $g \in G$ and $m \in M$. Clearly, θ is differentiable and its image in $M \times M$ is equal to R_G. Therefore, we can view θ as a differentiable map from $G \times M$ onto R_G. Then we have $p_2 \circ \theta = pr_2 : G \times M \longrightarrow M$. Therefore, this composition is a submersion. Since θ is surjective, p_2 must also be a submersion. □

Therefore, if R_G is a closed submanifold, the orbit space M/G has a natural structure of a differentiable manifold and the projection $p : M \longrightarrow M/G$ is a submersion. In this situation, we say that the group action is *regular* and we call M/G the *orbit manifold* of M.

For a regular action, all G-orbits in M are closed submanifolds of M by 2.1.9. Let Ω be an orbit in M in this case. By 1.3.3, the induced map $G \times \Omega \longrightarrow \Omega$ is a differentiable action of G on Ω. Moreover, the action of G on Ω is transitive. For any $g \in G$, the map $\tau(g) : \Omega \longrightarrow \Omega$ is a diffeomorphism. This implies that $\dim_{g \cdot m} \Omega = \dim_m \Omega$, for any $g \in G$, i.e., $m \longmapsto \dim_m \Omega$ is constant on Ω, and Ω

is of pure dimension. Moreover, for $m \in \Omega$, the orbit map $\rho(m) : G \longrightarrow \Omega$ is a surjective subimmersion by 1.1.3.

In addition, the map $\theta : G\times M \longrightarrow R_G$ is a differentiable surjection. Fix $m \in M$ and let Ω denote its orbit. We denote by $j_m : G \longrightarrow G \times M$ the differentiable map $j_m(g) = (g, m)$ for $g \in G$. Clearly, j_m is a diffeomorphism of G onto the closed submanifold $G \times \{m\}$ of $G \times M$. Analogously, we denote by $k_m : \Omega \longrightarrow M \times M$ the differentiable map given by $k_m(n) = (n, m)$ for $n \in \Omega$. Clearly, k_m is a diffeomorphism of Ω onto the closed submanifold $\Omega \times \{m\}$. Since $\Omega \times \{m\} = R_G \cap (M \times \{m\})$, we can view it as a closed submanifold of R_G. It follows that we have the following commutative diagram:

$$\begin{array}{ccc} G & \xrightarrow{\rho(m)} & \Omega \\ {\scriptstyle j_m}\big\downarrow & & \big\downarrow{\scriptstyle k_m} \\ G \times M & \xrightarrow[\theta]{} & M \times M \end{array} .$$

We say that a regular differentiable action of a Lie group G on M is *free* if the map $\theta : G \times M \longrightarrow R_G$ is a diffeomorphism.

The above diagram immediately implies that if the action of G is free, all orbit maps are diffeomorphisms of G onto the orbits. In addition, the stabilizers G_m for $m \in M$ are trivial. In §1.4 we are going to study free actions in more detail.

1.3. Coset spaces and quotient Lie groups. Let G be a Lie group and H be a Lie subgroup of G. Then $\mu_\ell : H\times G \longrightarrow G$ given by $\mu_\ell(h, g) = \gamma(h)(g) = hg$ for $h \in H$ and $g \in G$, defines a differentiable left action of H on G. The corresponding map $\theta_\ell : H \times G \longrightarrow G\times G$ is given by $\theta_\ell(h, g) = (hg, g)$. This map is the restriction to $H\times G$ of the map $\alpha_\ell : G\times G \longrightarrow G\times G$ defined by $\alpha_\ell(h, g) = (hg, g)$ for $g, h \in G$. This map is clearly differentiable, and its inverse is the map $\beta : G \times G \longrightarrow G \times G$ given by $\beta_\ell(h, g) = (hg^{-1}, g)$ for $g, h \in G$. Therefore, α_ℓ is a diffeomorphism. This implies that its restriction θ_ℓ to $H \times G$ is a diffeomorphism on the image R_G. Therefore, R_G is a closed submanifold of $G \times G$, and this action of H on G is regular and free. The quotient manifold is denoted by $H\backslash G$ and called the *right coset manifold* of G with respect to H.

Analogously, $\mu_r : H \times G \longrightarrow G$ given by $\mu_r(h, g) = \delta(h)(g) = gh^{-1}$ for $h \in H$ and $g \in G$, defines a differentiable left action of H on G. The corresponding map $\theta : H \times G \longrightarrow G\times G$ is given by $\theta_r(h, g) = (gh^{-1}, g)$. This map is the restriction to $H\times G$ of the map $\alpha_r : G\times G \longrightarrow G\times G$ defined by $\alpha_r(h, g) = (gh^{-1}, g)$ for $g, h \in G$. This map is clearly differentiable, and its inverse is the map $\beta_r : G \times G \longrightarrow G \times G$ given by $\beta_r(h, g) = (gh, g)$ for $g, h \in G$. Therefore, α_r is a diffeomorphism. This implies that its restriction θ_r to $H \times G$ is a diffeomorphism on the image R_G. Therefore, R_G is a closed submanifold of $G \times G$, and this action of H on G is regular and free. The quotient manifold is denoted by G/H and called the *left coset manifold* of G with respect to H.

Since G acts differentiably on G by right translations, we have a differentiable map $G \times G \xrightarrow{m} G \xrightarrow{p} H\backslash G$. This map is constant on right cosets in the first factor. By the above discussion it induces a differentiable map $\mu_{H,r} : G \times H\backslash G \longrightarrow H\backslash G$. It is easy to check that this map is a differentiable action of G on $H\backslash G$.

Analogously, we see that G acts differentiably on the left coset manifold G/H.

If N is a normal Lie subgroup of G, from the uniqueness of the quotient it follows that $G/N = N\backslash G$ as differentiable manifolds. Moreover, the map $G \times G \longrightarrow G/N$ given by $(g, h) \longrightarrow p(gh^{-1}) = p(g)p(h)^{-1}$ factors through $G/N \times G/N$. This proves that G/N is a Lie group. We call it the *quotient Lie group* G/N of G with respect to the normal Lie subgroup N.

Let G be a Lie group acting differentiably on a manifold M. Let $m \in M$ and G_m the stabilizer of m in G. Then the orbit map $\rho(m) : G \longrightarrow M$ is constant on left G_m-cosets. Therefore, it factors through the left coset manifold G/G_m, i.e., we have a commutative diagram

$$\begin{array}{ccc} G & \xrightarrow{\rho(m)} & M \\ {\scriptstyle p}\downarrow & \nearrow{\scriptstyle o(m)} & \\ G/G_m & & \end{array} .$$

Since $\rho(m)$ is a has constant rank by 1.1.3, we have

$$\begin{aligned} \operatorname{rank}_g \rho(m) &= \operatorname{rank}_1 \rho(m) = \dim \operatorname{im} T_1(\rho(m)) = \dim T_1(G) - \dim \ker T_1(\rho(m)) \\ &= \dim T_1(G) - \dim T_1(G_m) = \dim G - \dim G_m = \dim G/G_m. \end{aligned}$$

On the other hand, since p is a submersion we have $\operatorname{rank}_{p(g)} o(m) = \operatorname{rank}_g \rho(m) = \dim G/G_m$. Since p is surjective, this in turn implies that $o(m)$ is also a subimmersion. On the other hand, $o(m)$ is injective, therefore it has to be an immersion.

1.3.1. LEMMA. *The map $o(m) : G/G_m \longrightarrow M$ is an injective immersion.*

In particular, if $\phi : G \longrightarrow H$ a morphism of Lie groups, we have the commutative diagram

$$\begin{array}{ccc} G & \xrightarrow{\phi} & H \\ {\scriptstyle p}\downarrow & \nearrow{\scriptstyle \Phi} & \\ G/\ker\phi & & \end{array}$$

of Lie groups and their morphisms. The morphism Φ is an immersion. Therefore, any Lie group morphism can be factored into a composition of two Lie group morphisms, one of which is a surjective submersion and the other is an injective immersion.

1.4. Free actions. Let G be a Lie group acting differentiably on a manifold M. Assume that the action is regular. Therefore the quotient manifold M/G exists, and the natural projection $p : M \longrightarrow M/G$ is a submersion. Let U be an open set in M/G. A differentiable map $s : U \longrightarrow M$ is called a *local section* if $p \circ s = id_U$. Since p is a submersion, each point $u \in M/G$ has an open neighborhood U and a local section s on U.

Let $U \subset M/G$ be an open set and $s : U \longrightarrow M$ a local section. We define a differentiable map $\psi = \mu \circ (id_G \times s) : G \times U \longrightarrow M$. Clearly, if we denote by $p_2 : G \times U \longrightarrow U$ the projection to the second coordinate, we have

$$p(\psi(g, u)) = p(\mu(g, s(u))) = p(g \cdot s(u)) = p(s(u)) = u = p_2(g, u)$$

for any $g \in G$ and U, i.e., the diagram

$$\begin{array}{ccc} G \times U & \xrightarrow{\psi} & M \\ {\scriptstyle p_2}\big\downarrow & & \big\downarrow{\scriptstyle p} \\ U & \longrightarrow & M/G \end{array}$$

is commutative.

Clearly, the open subset $p^{-1}(U)$ of M is saturated and $s(U) \subset p^{-1}(U)$. Let $m \in p^{-1}(U)$. Then $p(m)$ corresponds to the orbit Ω of m. Moreover, $p(s(p(m))) = p(m)$ and $s(p(m))$ is also in Ω. This implies that $m = g \cdot s(p(m)) = \psi(g, p(m))$ for some $g \in G$, and the map ψ is a differentiable surjection of $G \times U$ onto $p^{-1}(U)$.

Let $m = s(u)$ for $u \in U$ and $g \in G$. Denote by Ω the G-orbit through m. Then $T_m(p) \circ T_u(s) = 1_{T_u(M/G)}$. Therefore, $T_u(s) : T_u(M/G) \longrightarrow T_m(M)$ is a linear injection, $T_m(p)$ is a linear surjection, $\ker T_m(p) \cap \operatorname{im} T_u(s) = \{0\}$ and $T_m(M) = \ker T_m(p) \oplus \operatorname{im} T_u(s)$. By 1.1.4.4, we have $\ker T_m(p) = T_m(\Omega)$. Hence, we have $T_m(M) = T_m(\Omega) \oplus \operatorname{im} T_u(s)$.

Now we want to calculate the differential $T_{(g,u)}(\psi) : T_{(g,u)}(G \times U) \longrightarrow T_{g \cdot m}(M)$. Let $i_u : G \longrightarrow G \times \{u\}$ and $i_g : U \longrightarrow \{g\} \times U$. First, we have

$$(\psi \circ i_u)(h) = h \cdot m = \tau(g)(g^{-1} \cdot h \cdot m) = (\tau(g) \circ \rho(m))(g^{-1}h) = (\tau(g) \circ \rho(m) \circ \gamma(g^{-1}))(h),$$

for any $h \in G$. So, by taking the differentials

$$T_g(\psi \circ i_u) = T_m(\tau(g)) \circ T_1(\rho(m)) \circ T_g(\gamma(g^{-1})).$$

Second, we have

$$(\psi \circ i_g)(v) = g \cdot s(v) = (\tau(g) \circ s)(v)$$

so, by taking differentials we have

$$T_u(\psi \circ i_g) = T_m(\tau(g)) \circ T_u(s).$$

Since $T_{(g,u)}(G \times U) = T_g(G) \oplus T_u(M/G)$, we have the formula

$$\begin{aligned} T_{(g,u)}(\psi)(X, Y) &= T_m(\tau(g))(T_1(\rho(m))(T_g(\gamma(g^{-1}))(X))) + T_m(\tau(g))(T_u(s)(Y)) \\ &= T_m(\tau(g))\Big(T_1(\rho(m))(T_g(\gamma(g^{-1}))(X)) + T_u(s)(Y)\Big) \end{aligned}$$

for $X \in T_g(G)$ and $Y \in T_u(M/G)$. Since $\tau(g)$ is a diffeomorphism, $T_m(\tau(g)) : T_m(M) \longrightarrow T_{g \cdot m}(M)$ is a linear isomorphism. Moreover, since $\gamma(g)$ is a diffeomorphism, $T_g(\gamma(g^{-1})) : T_g(G) \longrightarrow T_1(G)$ is a linear isomorphism. Hence, $T_{(g,u)}(\psi)$ is surjective if and only if

$$\operatorname{im} T_1(\rho(m)) + \operatorname{im} T_u(s) = T_m(M).$$

Clearly, $\operatorname{im} T_1(\rho(m)) \subset T_m(\Omega)$ and as we already remarked $T_m(\Omega) \oplus \operatorname{im} T_u(s) = T_m(M)$. Hence, $T_{(g,u)}(\psi)$ is surjective if and only if $T_1(\rho(m)) : T_1(G) \longrightarrow T_m(\Omega)$ is surjective.

Therefore, ψ is a surjective submersion of $G \times U$ onto $p^{-1}(U)$ if and only if all orbit maps $\rho(m)$ are submersions of G onto the orbits of $m \in s(U)$. Since

$$\rho(h \cdot m) = \rho(m) \circ \delta(h^{-1})$$

and $\delta(h^{-1})$ is a diffeomorphism, we see that $\rho(h \cdot m)$, $h \in G$, are subimmersions of the same rank. Therefore, the above condition is equivalent to all maps $\rho(m)$

being submersions of G onto orbits of $m \in p^{-1}(U)$. By 1.3.1, this is equivalent to all maps $o(m)$ being diffeomorphisms of G/G_m onto orbits of $m \in p^{-1}(U)$.

Let M be a manifold. Consider the action of G on $G \times M$ given by $\mu_M(g,(h,m)) = (gh,m)$ for any $g,h \in G$ and $m \in M$. This is clearly a differentiable action and $R_G = \{(g,m,h,m) \in G \times M \times G \times M\}$. Therefore, R_G is a closed submanifold of $G \times M \times G \times M$ and this action is regular. Moreover, the corresponding map $\theta_M : G \times G \times M \longrightarrow G \times M \times G \times M$ is given by the formula $\theta_M(g,h,m) = (gh,m,h,m)$ for $g,h \in G$ and $m \in M$, hence it is a diffeomorphism of $G \times G \times M$ onto R_G and the action of G on $G \times M$ is free.

No we we want to give a natural characterization of free actions and show that they locally look like the free action from the above example.

1.4.1. Theorem. *Let G be a Lie group acting differentiably on a manifold M. Assume that the action is regular. Then the following conditions are equivalent:*

(i) *the action of G is free;*
(ii) *all orbit maps $\rho(m) : G \longrightarrow \Omega$, $m \in M$, are diffeomorphisms;*
(iii) *for any point $u \in M/G$ there exists an open neighborhood U of u in M/G and a local section $s : U \longrightarrow M$ such that the map $\psi : G \times U \longrightarrow M$ is a diffeomorphism of $G \times U$ onto the open submanifold $p^{-1}(U)$ of M.*

Proof. We already established that if the action of G is free, all orbit maps are diffeomorphisms. Hence, (i) implies (ii). If (ii) holds, by the above discussion, we see that ψ is a surjective submersion. On the other hand,

$$\dim_{(g,u)}(G \times U) = \dim G + \dim_u(M/G) = \dim \Omega + \dim_u(M/G) = \dim_{g \cdot m} M,$$

so $T_{(g,u)}(\psi)$ is also injective. Therefore, ψ is a local diffeomorphism. On the other hand, if $\psi(g,u) = \psi(h,v)$, we have $u = p(\psi(g,u)) = p(\psi(h,v)) = v$. Moreover, $g \cdot u = h \cdot u$ implies that $g = h$, since the orbit maps are diffeomorphisms. It follows that ψ is a bijection. Since it is a local diffeomorphism, it must be a diffeomorphism. Therefore, (iii) holds.

It remains to show that (iii) implies (i). First assume that we have an open set U in M/G and a local section s on U such that $\psi : G \times U \longrightarrow p^{-1}(U)$ is a diffeomorphism. Then, $p^{-1}(U)$ is G-invariant and we can consider the G-action induced on $p^{-1}(U)$. Clearly, this action of G is differentiable. If we consider the action of G onto $G \times U$ from the previous example, the diagram

$$\begin{array}{ccc} G \times G \times U & \xrightarrow{\mu_U} & G \times U \\ {\scriptstyle id_G \times \psi}\downarrow & & \downarrow{\scriptstyle \psi} \\ G \times p^{-1}(U) & \xrightarrow{\mu} & p^{-1}(U) \end{array}$$

is commutative, since

$$\psi(\mu_U(g,(h,u))) = \psi(gh,u) = gh \cdot s(u) = \mu(g, \psi(h,s(u)))$$

for all $g,h \in G$ and $u \in U$. This implies that the diagram

$$\begin{array}{ccc} G \times G \times U & \xrightarrow{\theta_U} & G \times U \times G \times U \\ {\scriptstyle id_G \times \psi}\downarrow & & \downarrow{\scriptstyle \psi \times \psi} \\ G \times p^{-1}(U) & \xrightarrow{\theta} & p^{-1}(U) \times p^{-1}(U) \end{array}$$

is commutative and the vertical arrows are diffeomorphisms. The diffeomorphism $\psi \times \psi$ maps the graph of the equivalence relation on $G \times U$ onto the graph of the equivalence relation on $p^{-1}(U)$. Since the action on $G \times U$ is free, the action on $p^{-1}(U)$ is also free. Therefore, the restriction of θ to $G \times p^{-1}(U)$ is a diffeomorphism onto $R_G \cap (p^{-1}(U) \times p^{-1}(U))$.

Therefore (iii) implies that θ is a local diffeomorphism of $G \times M$ onto R_G. In addition, the orbit maps are diffeomorphisms.

It remains to show that $\theta : G \times M \longrightarrow M \times M$ is an injection. Assume that $\theta(g,m) = \theta(h,n)$ for $g, h \in G$ and $m, n \in M$. Then we have $(g \cdot m, m) = (h \cdot n, n)$, i.e., $m = n$ and $g \cdot m = h \cdot m$. Since the orbit maps are bijections, this implies that $g = h$. □

1.5. Lie groups with countably many components. Let G be a Lie group. The connected component G_0 of G containing the identity is called the *identity component* of G. Clearly, G_0 is an open and closed subset of G. For any $g \in G_0$ the right translation $\delta(g)$ permutes connected components of G. Moreover, it maps the g into 1, hence it maps G_0 onto itself. It follows that G_0 is a Lie subgroup of G.

Moreover, the map $\operatorname{Int}(g) : G \longrightarrow G$ is a Lie group automorphism of G. Therefore, it also permutes the connected components of G. In particular it maps G_0 onto itself. This implies that G_0 is a normal Lie subgroup of G. The quotient Lie group G/G_0 is discrete and its cardinality is equal to the number of connected components of G.

1.5.1. Lemma. *Let G be a connected Lie group. For any neighborhood U of the identity 1 in G, we have*

$$G = \bigcup_{n=1}^{\infty} U^n.$$

Proof. Let V be a symmetric neighborhood of identity contained in U. Let $H = \bigcup_{n=1}^{\infty} V^n$. If $g \in V^n$ and $h \in V^m$, it follows that $gh \in V^{n+m} \subset H$. Therefore, H is closed under multiplication. In addition, if $g \in V^n$, we see that $g^{-1} \in V^n$ since V is symmetric, i.e., H is a subgroup of G. Since $V \subset H$, H is a neighborhood of the identity in G. Since H is a subgroup, it follows that H is a neighborhood of any of its points, i.e., H is open in G. This implies that the complement of H in G is a union of H-cosets, which are also open in G. Therefore, H is also closed in G. Since G is connected, $H = G$. □

This result has the following consequence.

1.5.2. Corollary. *Let G be a connected Lie group. Then G is separable.*

Proof. Let U be a neighborhood of 1 which is domain of a chart. Then, U contains a countable dense set C. By continuity of multiplication, it follows that C^n is dense in U^n for any $n \in \mathbb{Z}_+$. Therefore, by 1.5.1, $D = \bigcup_{n=1}^{\infty} C^n$ is dense in G. In addition, D is a countable set. Therefore, there exists a countable dense set D in G.

Let $(U_n; n \in \mathbb{Z}_+)$, be a fundamental system of neighborhoods of 1 in G. Without any loss of generality we can assume that U_n are symmetric. We claim that $\mathcal{U} = \{U_n d \mid m \in \mathbb{Z}_+, d \in D\}$ is a basis of the topology on G. Let V be an open set in G and $g \in V$. Then there exists $n \in \mathbb{Z}_+$ such that $U_n^2 g \subset V$. Since D is dense in

G, there exists $d \in D$ such that $d \in U_n g$. Since U_n is symmetric, this implies that $g \in U_n d$. Moreover, we have

$$U_n d \subset U_n^2 g \subset V.$$

Therefore, V is a union of open sets from $\mathcal{U}$. □

A locally compact space is *countable at infinity* if it is a union of countably many compact subsets.

1.5.3. LEMMA. *Let G be a Lie group. Then the following conditions are equivalent:*

(i) *G is countable at infinity;*
(ii) *G has countably many connected components.*

PROOF. $(i) \Rightarrow (ii)$ Let K be a compact set in G. Since it is covered by the disjoint union of connected components of G, it can intersect only finitely many connected components of G. Therefore, if G is countable at infinity, it can must have countably many components.

$(ii) \Rightarrow (i)$ Let g_i, $i \in I$, be a set of representatives of connected components in G. Then $G = \bigcup_{i\in I} g_i G_0$. Let K be a connected compact neighborhood of the identity in G. Then $K \subset G_0$ and by 1.5.1, we have $G_0 = \bigcup_{n=1}^\infty K^n$. Moreover, K^n, $n \in \mathbb{N}$, are all compact. It follows that

$$G = \bigcup_{i\in I} \bigcup_{n=1}^{\infty} g_i K^n.$$

Therefore, if I is countable, G is countable at infinity. □

A topological space X is a *Baire space* if the intersection of any countable family of open, dense subsets of X is dense in X.

1.5.4. LEMMA (Category theorem). *Any locally compact space X is a Baire space.*

PROOF. Let U_n, $n \in \mathbb{N}$, be a countable family of open, dense subsets of X. Let $V = V_1$ be a nonempty open set in X with compact closure. Then $V_1 \cap U_1$ is a nonempty open set in X. Therefore, we can pick a nonempty open set with compact closure $V_2 \subset \bar{V}_2 \subset V_1 \cap U_1$. Then $V_2 \cap U_2$ is a nonempty open subset of X. Continuing this procedure, we can construct a sequence V_n of nonempty open subsets of X with compact closure such that $V_{n+1} \subset \bar{V}_{n+1} \subset V_n \cap U_n$. Therefore, $\bar{V}_{n+1} \subset \bar{V}_n$ for $n \in \mathbb{N}$, i.e., $\bar{V}_n$, $n \in \mathbb{N}$, is a decreasing family of compact sets. Therefore, $W = \bigcap_{n=1}^\infty \bar{V}_n \neq \emptyset$. On the other hand, $W \subset \bar{V}_{n+1} \subset U_n$ for all $n \in \mathbb{N}$. Hence the intersection of all U_n, $n \in \mathbb{N}$, with V is not empty. □

1.5.5. PROPOSITION. *Let G be a locally compact group countable at infinity acting continuously on a hausdorff Baire space M. Assume that the action of G on M is transitive. Then the orbit map $\rho(m) : G \longrightarrow M$ is open for any $m \in M$.*

PROOF. Let U be a neighborhood of 1 in G. We claim that $\rho(m)(U)$ is a neighborhood of m in M.

Let V be a symmetric compact neighborhood of 1 in G such that $V^2 \subset U$. Clearly, $(gV; g \in G)$, is a cover of G. Since G is countable at infinity, this cover has a countable subcover $(g_n V; n \in \mathbb{N})$, i.e., $G = \bigcup_{n=1}^\infty g_n V$. Therefore, M is equal to the union of compact sets $(g_n V) \cdot m$, $n \in \mathbb{N}$. Let $U_n = M - (g_n V) \cdot m$ for $n \in \mathbb{N}$.

Since V is compact, all $g_n V$, $n \in \mathbb{N}$, are compact. Hence their images $(g_n V) \cdot m$ are compact and therefore closed in M. It follows that U_n, $n \in \mathbb{N}$, are open in M. Moreover, we have

$$\bigcap_{n=1}^{\infty} U_n = \bigcap_{n=1}^{\infty} (M - (g_n V) \cdot m) = M - \bigcup_{n=1}^{\infty} (g_n V) \cdot m = \emptyset.$$

Therefore, by 1.5.4, at least one U_n cannot be dense in M. Hence $M - V \cdot m = \tau(g_n^{-1})(M - (g_n V) \cdot m)$ is not dense in M. It follows that $V \cdot m$ has a nonempty interior. Assume that $g \cdot m$, $g \in V$, is an interior point of $V \cdot m$. Then $(g^{-1}V) \cdot m$ is a neighborhood of m. Therefore,

$$(g^{-1}V) \cdot m \subset V^2 \cdot m \subset U \cdot m = \rho(m)(U)$$

is a neighborhood of $m \in M$. This establishes our claim.

Assume now that U is an arbitrary open set in G. Let $g \in U$. Then $g^{-1}U$ is a neighborhood of $1 \in G$. Hence, by the claim, $g^{-1} \cdot \rho(m)(U) = \rho(m)(g^{-1}U)$ is a neighborhood of $m \in M$. This implies that $\rho(m)(U)$ is a neighborhood of $g \cdot m$. Therefore, $\rho(m)(U)$ is a neighborhood of any of its points, i.e., it is an open set. □

Let G be a Lie group acting differentiably on a manifold M. If the action is transitive, the orbit map $\rho(m) : G \longrightarrow M$ is a surjective subimmersion. If G has countably many connected components, it is countable at infinity by 1.5.3. Therefore, by 1.5.5, $\rho(m)$ is an open map. By 1.1.3.2, it has to be a submersion. As we remarked before, it factors through a differentiable map $o(m) : G/G_m \longrightarrow M$. Clearly, in our situation, the map $o(m)$ is an bijective submersion. By 1.3.1, it is also an immersion. Therefore, we have the following result.

1.5.6. Theorem. *Let G be a Lie group with countably many connected components acting differentiably on a manifold M. Assume that the action of G on M is transitive. Then the orbit map induces a diffeomorphism $o(m) : G/G_m \longrightarrow M$.*

This has the following direct consequences.

1.5.7. Corollary. *Let $\phi : G \longrightarrow H$ be a surjective Lie group morphism. If G has countably many connected components the induced homomorphism $\Phi : G/\ker\phi \longrightarrow H$ is an isomorphism.*

1.5.8. Theorem. *Let G be a Lie group with countably many connected components acting differentiably on a manifold M. Assume that the action is regular. Then the following conditions are equivalent:*

(i) *all stabilizers G_m, $m \in M$, are trivial;*
(ii) *the action of G on M is free.*

Another consequence of the argument in the proof of 1.5.5 is the following observation.

1.5.9. Lemma. *Let G be a locally compact group countable at infinity acting continuously on a hausdorff Baire space M. Assume that G has countably many orbits in M. Then there exists an open orbit in M.*

Proof. Let m_i, $i \in I$, be a family of representatives of all G-orbits in M. Let V be a compact neighborhood of $1 \in G$. Then, as in the proof of 1.5.5, there exists a sequence $(g_n; n \in \mathbb{N})$ such that $G = \bigcup_{n=1}^{\infty} g_n V$. Therefore, we have

$$M = \bigcup_{i \in I} \bigcup_{n=1}^{\infty} (g_n V) \cdot m_i.$$

If we define $U_{i,n} = M - (g_n V) \cdot m_i$, $i \in I, n \in \mathbb{N}$, the sets $U_{i,n}$ are open sets in M. In addition,

$$\bigcap_{i \in I} \bigcap_{n=1}^{\infty} U_{i,n} = M - \bigcup_{i \in I} \bigcup_{n=1}^{\infty} (g_n V) \cdot m_i = \emptyset.$$

Since I is countable, by 1.5.4, at least one $U_{i,n}$ cannot be dense in M. Therefore, that $g_n V \cdot m_i$ has a nonempty interior. This implies that the orbit $G \cdot m_i$ has nonempty interior. Let m be an interior point of $G \cdot m_i$. Then, for any $g \in G$, $g \cdot m$ is another interior point of $G \cdot m_i$. Therefore, all points in $G \cdot m_i$ are interior, i.e., the orbit $G \cdot m_i$ is open in M. □

This has the following consequence.

1.5.10. Proposition. *Let G be a Lie group with countably many components acting differentiably on a manifold M. If G acts on M with countably many orbits, all orbits are submanifolds in M.*

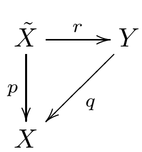
Proof. Let Ω be an orbit in M. Since Ω is G-invariant, its closure $\bar{\Omega}$ is G-invariant. Therefore $\bar{\Omega}$ is a union of countably many orbits. Moreover, it is a locally compact space. Hence, by 1.5.4, it is a Baire space. If we apply 1.5.9 to the action of G on $\bar{\Omega}$, we conclude that $\bar{\Omega}$ contains an orbit Ω' which is open in $\bar{\Omega}$. Since Ω is dense in $\bar{\Omega}$, we must have $\Omega' = \Omega$. Therefore, Ω is open in $\bar{\Omega}$. Therefore, there exists an open set U in M such that $\bar{\Omega} \cap U = \Omega$, i.e., Ω is closed in U. Therefore, the orbit Ω is locally closed in M. In particular, Ω is a locally compact space with the induced topology. Let $m \in \Omega$. Using again 1.5.4 and 1.5.5 we see that the map $o(m) : G/G_m \longrightarrow \Omega$ is a homeomorphism. By 1.3.1, Ω is the image of an immersion $o(m) : G/G_m \longrightarrow M$. Therefore, by 1.1.4.2, Ω is a submanifold of M. □

1.6. Universal covering Lie group. Let X be a connected manifold with base point x_0. A *covering* of (X, x_0) is a triple consisting of a connected manifold Y with a base point y_0 and a projection $q : Y \longrightarrow X$ such that

(i) q is a surjective local diffeomorphism;
(ii) $q(y_0) = x_0$;
(iii) for any $x \in X$ there exists a connected neighborhood U of X such that q induces a diffeomorphism of every connected component of $q^{-1}(U)$ onto U.

The map q is called the *covering projection* of Y onto X.

A cover $(\tilde{X}, p, \tilde{x}_0)$ of (X, x_0) is called a *universal covering* if for any other covering (Y, q, y_0) of (X, x_0) there exists a unique differentiable map $r : \tilde{X} \longrightarrow Y$ such that $(\tilde{X}, r, \tilde{x}_0)$ is a covering of (Y, y_0) and the diagram



is commutative.

Clearly, the universal covering is unique up to an isomorphism.

Any connected manifold X with base point x_0 has a universal cover $\tilde{X}$ and $\pi_1(\tilde{X}, \tilde{x}_0)$ is trivial, i.e., $\tilde{X}$ is simply connected.

1.6.1. Lemma. *Let (X, x_0) be a connected manifold and (Y, p, y_0) its covering. Let (Z, z_o) be a connected and simply connected manifold, and $F : Z \longrightarrow X$ a differentiable map such that $F(z_0) = x_0$. Then there exists a unique differentiable map $F' : Z \longrightarrow Y$ such that*

(i) *$F'(z_0) = y_0$;*
(ii) *the diagram*

$$\begin{array}{ccc} Z & \xrightarrow{F'} & Y \\ & \searrow^{F} & \downarrow p \\ & & X \end{array}$$

is commutative.

Let (Y, q, y_0) be a covering space of (X, x_0). A diffeomorphism $\phi : Y \longrightarrow Y$ is called a *deck transformation* if $q \circ \phi = q$.

Let $(\tilde{X}, \tilde{x}_0)$ be the universal covering space of (X, x_0). Then any loop $\gamma : [0, 1] \longrightarrow X$ such that $\gamma(0) = \gamma(1) = x_0$ can be lifted to the unique curve $\tilde{\gamma} : [0, 1] \longrightarrow \tilde{X}$ such that

(i) $\tilde{\gamma}(0) = \tilde{x}_0$;
(ii) $p \circ \tilde{\gamma} = \gamma$.

The end point $\tilde{\gamma}(1)$ of $\tilde{\gamma}$ is in $p^{-1}(x_0)$. This map induces a bijection of $\pi_1(X, x_0)$ onto $p^{-1}(x_0)$. On the other hand, for any $x \in p^{-1}(x_0)$ there exists a unique deck transformation of $\tilde{X}$ which maps $\tilde{x}_0$ into x. In this way, we construct a map from the fundamental group $\pi_1(X, x_0)$ onto the group of deck transformations of $\tilde{X}$. This map is a group isomorphism. Therefore, $\pi_1(X, x_0)$ acts on $\tilde{X}$ and X is the quotient of $\tilde{X}$ with respect to this action.

Let G be a connected Lie group. Denote by $(\tilde{G}, p, \tilde{1})$ the universal covering space of $(G, 1)$. Then $\tilde{G} \times \tilde{G}$ is connected and simply connected. Therefore, the mapping $m \circ (p \times p) : \tilde{G} \times \tilde{G} \longrightarrow G$ has a lifting $\tilde{m} : \tilde{G} \times \tilde{G} \longrightarrow \tilde{G}$ such that $\tilde{m}(\tilde{1}, \tilde{1}) = \tilde{1}$.

We claim that $\tilde{G}$ with the multiplication defined by $\tilde{m}$ is a group. First, we have

$$\begin{aligned} p \circ (\tilde{m} \circ (id_{\tilde{G}} \times \tilde{m})) &= m \circ (p \times p) \circ (id_{\tilde{G}} \times \tilde{m}) \\ &= m \circ (p \times p \circ \tilde{m}) = m \circ (p \times m \circ (p \times p)) = m \circ (id_G \times m) \circ (p \times p \times p) \end{aligned}$$

and

$$\begin{aligned} p \circ (\tilde{m} \circ (\tilde{m} \times id_{\tilde{G}})) &= m \circ (p \times p) \circ (\tilde{m} \times id_{\tilde{G}}) \\ &= m \circ (p \circ \tilde{m} \times p) = m \circ ((m \circ (p \times p)) \times p) = m \circ (m \times id_G) \circ (p \times p \times p). \end{aligned}$$

Since the multiplication on G is associative, it follows that $\tilde{m} \circ (id_{\tilde{G}} \times \tilde{m})$ and $\tilde{m} \circ (\tilde{m} \times id_{\tilde{G}})$ are the lifts of the same map from $\tilde{G} \times \tilde{G} \times \tilde{G}$ into G. Since both maps map $(\tilde{1}, \tilde{1}, \tilde{1})$ into $\tilde{1}$, it follows that they are identical, i.e., the operation $\tilde{m}$ is associative.

Also, we have

$$p(\tilde{m}(\tilde{g}, \tilde{1})) = m(p(\tilde{g}), 1) = p(\tilde{g})$$

for any $g \in G$, hence $\tilde{g} \longmapsto \tilde{m}(\tilde{g}, \tilde{1})$ is the lifting of $p : \tilde{G} \longrightarrow G$. Since $\tilde{m}(\tilde{1}, \tilde{1}) = \tilde{1}$, this map is the identity on $\tilde{G}$, i.e., $\tilde{m}(\tilde{g}, \tilde{1}) = \tilde{g}$ for all $\tilde{g} \in \tilde{G}$.

Analogously, we have

$$p(\tilde{m}(\tilde{1}, \tilde{g})) = m(1, p(\tilde{g})) = p(\tilde{g})$$

for any $g \in G$, hence $\tilde{g} \longmapsto \tilde{m}(\tilde{1}, \tilde{g})$ is the lifting of $p : \tilde{G} \longrightarrow G$. Since $\tilde{m}(\tilde{1}, \tilde{1}) = \tilde{1}$, this map is the identity on $\tilde{G}$, i.e., $\tilde{m}(\tilde{1}, \tilde{g}) = \tilde{g}$ for all $\tilde{g} \in \tilde{G}$.

It follows that $\tilde{1}$ is the identity in $\tilde{G}$.

Let $\tilde{\iota} : \tilde{G} \longrightarrow \tilde{G}$ be the lifting of the map $\iota \circ p : \tilde{G} \longrightarrow G$ such that $\tilde{\iota}(\tilde{1}) = \tilde{1}$.

Then we have

$$p(\tilde{m}(\tilde{g}, \tilde{\iota}(\tilde{g}))) = m(p(\tilde{g}), p(\tilde{\iota}(\tilde{g}))) = m(p(\tilde{g}), p(\tilde{g})^{-1}) = 1.$$

Therefore, $\tilde{g} \longmapsto \tilde{m}(\tilde{g}, \tilde{\iota}(\tilde{g}))$ is the lifting of the constant map of $\tilde{G}$ into 1. Since $(\tilde{m}(\tilde{1}, \tilde{\iota}(\tilde{1})) = \tilde{1}$, we conclude that this map is constant and its value is equal to $\tilde{1}$. Therefore, we have

$$\tilde{m}(\tilde{g}, \tilde{\iota}(\tilde{g})) = \tilde{1}$$

for all $\tilde{g} \in \tilde{G}$.

Analogously, we have

$$p(\tilde{m}(\tilde{\iota}(\tilde{g}), \tilde{g})) = m(p(\tilde{\iota}(\tilde{g})), p(\tilde{g})) = m(p(\tilde{g})^{-1}, p(\tilde{g})) = 1.$$

Therefore, $\tilde{g} \longmapsto \tilde{m}(\tilde{\iota}(\tilde{g}), \tilde{g})$ is the lifting of the constant map of $\tilde{G}$ into $1 \in G$. Since $(\tilde{m}(\tilde{\iota}(\tilde{1}), \tilde{1}) = \tilde{1}$, we conclude that this map is constant and its value is equal to $\tilde{1}$. Therefore, we have

$$\tilde{m}(\tilde{\iota}(\tilde{g}), \tilde{g}) = \tilde{1}$$

for all $\tilde{g} \in \tilde{G}$.

This implies that any element $\tilde{g} \in \tilde{G}$ has an inverse $\tilde{g}^{-1} = \tilde{\iota}(\tilde{g})$. Therefore, $\tilde{G}$ is a group. Moreover, since $\tilde{m}$ and $\tilde{\iota}$ are differentiable maps, $\tilde{G}$ is a Lie group. It is called the *universal covering Lie group* of G.

By the construction we have $m \circ (p \times p) = p \circ \tilde{m}$, i.e., $p : \tilde{G} \longrightarrow G$ is a Lie group homomorphism. Let $D = \ker p$. Then D is a normal Lie subgroup of $\tilde{G}$. Since p is a covering projection, D is also discrete.

For any $d \in D$, $\gamma(d) : \tilde{G} \longrightarrow \tilde{G}$ is a deck transformation which moves $\tilde{1}$ into d. Therefore $d \longmapsto \gamma(d)$ defines an isomorphism of D with the group of all deck transformations of $\tilde{G}$. Composing this with the isomorphism of the fundamental group $\pi_1(G, 1)$ with the group of all deck transformations we see that

$$\pi_1(G, 1) \cong \ker p.$$

1.6.2. Lemma. *Let D be a discrete subgroup of a Lie group G. Then D is a closed subgroup.*

Proof. Clearly, D is locally closed. Hence, by 1.1.1, D is closed in G. □

1.6.3. Lemma. *Let G be a connected Lie group and D its discrete normal subgroup. Then D is a central subgroup.*

Proof. Let $d \in D$. Then $\alpha : g \longmapsto gdg^{-1}$ is a continuous map from G into G and the image of α is contained in D. Therefore, the map $\alpha : G \longrightarrow D$ is continuous. Since G is connected, and D discrete it must be a constant map. Therefore, $gdg^{-1} = \alpha(g) = \alpha(1) = g$ for any $g \in G$. It follows that $gd = dg$ for any $g \in G$, and d is in the center of G. □

In particular, the kernel $\ker p$ of the covering projection $p : \tilde{G} \longrightarrow G$ is a discrete central subgroup of $\tilde{G}$. From the above discussion, we conclude that the following result holds.

1.6.4. Proposition. *The fundamental group $\pi_1(G, 1)$ is abelian.*

Let (Y, q, y_0) be another covering of $(G, 1)$. Then there exists a covering map $r : \tilde{G} \longrightarrow Y$ such that $p = q \circ r$ and $r(\tilde{1}) = y_0$. All deck transformations of $\tilde{G}$ corresponding to the covering $r : \tilde{G} \longrightarrow Y$ are also deck transformations for $p : \tilde{G} \longrightarrow G$. Therefore they correspond to a subgroup C of D. Since D is a central subgroup of $\tilde{G}$, C is also a central subgroup of $\tilde{G}$. It follows that r is constant on C-cosets in $\tilde{G}$ and induces a quotient map $\tilde{G}/C \longrightarrow Y$. This map is a diffeomorphism, hence Y has a Lie group structure for which y_0 is the identity. This proves the following statement which describes all covering spaces of a connected Lie group.

1.6.5. Theorem. *Any covering of $(G, 1)$ has a unique Lie group structure such that the base point is the identity element and the covering projection is a morphism of Lie groups.*

On the other hand, we have the following characterization of covering projections.

1.6.6. Proposition. *Let $\varphi : G \longrightarrow H$ be a Lie group homomorphism of connected Lie groups. Then φ is a covering projection if and only if $T_1(\varphi) : T_1(G) \longrightarrow T_1(H)$ is a linear isomorphism.*

Proof. If φ is a covering projection, it is a local diffeomorphism and the assertion is obvious.

If $T_1(\varphi) : T_1(G) \longrightarrow T_1(H)$ is a linear isomorphism, φ is a local diffeomorphism at 1. By 1.1.5, φ has constant rank, i.e., it is a local diffeomorphism. In particular, it is open and the image contains a neighborhood of identity in H. Since the image is a subgroup, by 1.5.1 it is equal to H. Therefore, φ is surjective. Moreover, $T_1(\ker\varphi) = \{0\}$ by 1.1.5, i.e., $D = \ker\varphi$ is discrete. By 1.6.3, D is a discrete central subgroup. It follows that φ induces an isomorphism of G/D onto H. Therefore, H is evenly covered by G because of 1.4.1. □

Let G and H be connected Lie groups and $\varphi : G \longrightarrow H$ be a Lie group homomorphism. Assume that G is simply connected. Then there exists a unique lifting $\tilde{\varphi} : G \longrightarrow \tilde{H}$ such that $\tilde{\varphi}(1) = \tilde{1}$. Since, we have

$$p\circ\tilde{m}\circ(\tilde{\varphi}\times\tilde{\varphi}) = m\circ(p\times p)\circ(\tilde{\varphi}\times\tilde{\varphi}) = m\circ((p\circ\tilde{\varphi})\times(p\circ\tilde{\varphi})) = m\circ(\varphi\times\varphi) = \varphi\circ m = p\circ\tilde{\varphi}\circ m$$

the maps $\tilde{m} \circ (\tilde{\varphi} \times \tilde{\varphi})$ and $\tilde{\varphi} \circ m$ are the lifts of the same map. They agree on $(1, 1)$ in $G \times G$, hence they are identical. This implies that $\tilde{\varphi} : G \longrightarrow \tilde{H}$ is a Lie group homomorphism.

Therefore, we have the following result.

1.6.7. LEMMA. *Let $\varphi : G \longrightarrow H$ be a Lie group homomorphism of a simply connected, connected Lie group G into a connected Lie group H. Let $\tilde{H}$ be the universal covering Lie group of H and $p : \tilde{H} \longrightarrow H$ the covering projection. Then there exists a unique Lie group homomorphism $\tilde{\varphi} : G \longrightarrow \tilde{H}$ such that $p \circ \tilde{\varphi} = \varphi$.*

In addition, if $\varphi : G \longrightarrow H$ is a Lie group morphism of connected Lie groups, there exists a unique Lie group homomorphism $\tilde{\varphi} : \tilde{G} \longrightarrow \tilde{H}$ such that the diagram

$$\begin{array}{ccc} \tilde{G} & \xrightarrow{\tilde{\varphi}} & \tilde{H} \\ {\scriptstyle p_G}\downarrow & & \downarrow{\scriptstyle p_H} \\ G & \xrightarrow{\varphi} & H \end{array},$$

where the vertical arrows are covering projections, is commutative.

1.7. A categorical interpretation. Let $\mathcal{L}ie$ be the category of Lie groups. Denote by $\mathcal{C}onn\mathcal{L}ie$ its full subcategory of connected Lie groups. If G is a Lie group, its identity component G_0 is a connected Lie group. Moreover, if $\varphi : G \longrightarrow H$ is a Lie group morphism, $\varphi(G_0) \subset H_0$. Therefore, the restriction φ_0 of φ to G_0 is a morphism $\varphi_0 : G_0 \longrightarrow H_0$. It is easy to check that this defines a functor from the category $\mathcal{L}ie$ into the category $\mathcal{C}onn\mathcal{L}ie$. In addition, we have

$$\operatorname{Hom}(G, H) = \operatorname{Hom}(G, H_0)$$

for any connected Lie group G and arbitrary Lie group H. Therefore, taking the identity component is the right adjoint to the forgetful functor $\mathrm{For} : \mathcal{C}onn\mathcal{L}ie \longrightarrow \mathcal{L}ie$.

Let $\mathcal{S}imply\mathcal{C}onn\mathcal{L}ie$ be the full subcategory of $\mathcal{L}ie$ consisting of simply connected connected Lie groups. It follows from the above discussion that $\tilde{}$ is a functor from $\mathcal{C}onn\mathcal{L}ie$ into $\mathcal{S}imply\mathcal{C}onn\mathcal{L}ie$. By 1.6.7, the universal covering functor $\tilde{}$ is the right adjoint to the forgetful functor $\mathrm{For} : \mathcal{S}imply\mathcal{C}onn\mathcal{L}ie \longrightarrow \mathcal{C}onn\mathcal{L}ie$.

It follows that the composition of the identity component functor and the universal covering functor is the right adjoint to the forgetful functor from the category $\mathcal{S}imply\mathcal{C}onn\mathcal{L}ie$ into $\mathcal{L}ie$.

In the next section we are going to show that $\mathcal{S}imply\mathcal{C}onn\mathcal{L}ie$ is equivalent to a category with purely algebraic objects.

1.8. Some examples. Let M be a manifold with an differentiable map $m : M \times M \longrightarrow M$ which defines an associative multiplication operation on M. Assume that this operation has the identity 1.

Let $i : M \longrightarrow M \times M$ and $j : M \longrightarrow M \times M$ be the maps given by $i(m) = (m, 1)$ and $j(m) = (1, m)$ for all $m \in M$. Then $m \circ i = id_M$ and $m \circ j = id_M$. Hence, their differentials satisfy

$$T_{(1,1)}(m) \circ T_1(i) = I_{T_1(M)} \text{ and } T_{(1,1)}(m) \circ T_1(j) = I_{T_1(M)}.$$

Moreover, $T_1(i)(X) = (X, 0)$ and $T_1(j)(Y) = (0, Y)$ for any $X, Y \in T_1(M)$. Therefore,

$$\begin{aligned} T_{(1,1)}(m)(X, Y) = T_{(1,1)}(m)((X, 0) + (0, Y)) &= T_{(1,1)}(m)(X, 0) + T_{(1,1)}(m)(0, Y) \\ &= T_{(1,1)}(m)(T_1(i)(X)) + T_{(1,1)}(m)(T_1(j)(Y)) = X + Y \end{aligned}$$

for any $X, Y \in T_1(M)$. Hence, we proved the following result.

1.8.1. Lemma. *The differential of the multiplication map at* $(1,1)$ *is given by*

$$T_{(1,1)}(m)(X,Y) = X + Y$$

for any $X, Y \in T_1(M)$.

Let G be the set of all invertible elements in M. Then, G is a group.

1.8.2. Lemma. *The group* G *is an open submanifold of* M. *With this manifold structure,* G *is a Lie group.*

Proof. Consider the map $\phi : M \times M \longrightarrow M \times M$ defined by $\phi(a,b) = (a, m(a,b))$ for $m, n \in M$. Then, by 1.8.1 we see that $T_{1,1}(\phi)(X,Y) = (X, X+Y)$ for any $X, Y \in T_1(M)$. Therefore, ϕ is a local diffeomorphism at 1. Therefore, there exists neighborhoods U and V of $(1,1) \in M \times M$ such that $\phi : U \longrightarrow V$ is a diffeomorphism. Let $\psi : V \longrightarrow U$ be the inverse map. Then $\psi(a, m(a,b)) = (a,b)$ for all $(a,b) \in U$. Hence, if we shrink V to be of the form $W \times W$ for some open neighborhood W of 1 in M, we have $\psi(a,b) = (a, \alpha(a,b))$ for some differentiable function $\alpha : W \times W \longrightarrow M$ and $a, b \in W$. In particular, if we put $\iota(a) = \alpha(a,1)$, we have

$$(a, m(a, \iota(a))) = \phi(a, \iota(a)) = \phi(a, \alpha(a,1)) = (a,1)$$

for $a \in W$. Therefore, all elements in W have a right inverse.

Analogously, by considering the opposite multiplication $m^\circ(a,b) = m(b,a)$ for $a, b \in M$, we conclude that there exists an open neighborhoods W' where all elements have a left inverse. Therefore, elements of $O = W \cap W'$ have left and right inverses. Let $a \in O$, b a left inverse and c a right inverse. Then

$$b = b(ac) = (ba)c = c,$$

i.e., any left inverse is equal to the right inverse c. This implies that the left and right inverses are equal and unique. In particular the elements of W are invertible.

It follows that $W \subset G$. Let $g \in G$. Then the left multiplication $\gamma(g) : M \longrightarrow M$ by g is a diffeomorphism. Therefore, $g \cdot W \subset G$ is an open neighborhood of g. It follows that G is an open submanifold of M.

Since the map $g \longrightarrow g^{-1}$ is given by ι on W, it is differentiable on W. If $h \in g \cdot W$, we have $h^{-1} = (g(g^{-1}h))^{-1} = (g^{-1}h)^{-1}g^{-1} = \iota(g^{-1}h)g^{-1}$, and this implies that the inversion is differentiable at g. It follows that G is a Lie group. □

In particular, this implies that checking the differentiability of the inversion map in a Lie group is redundant. If G is a manifold and a group and the multiplication map $m : G \times G \longrightarrow G$ is differentiable, then G is automatically a Lie group.

Let A be a finite dimensional associative algebra over $\mathbb{R}$ with identity. Then the group G of invertible elements in A is an open submanifold of A and with induced structure it is a Lie group. The tangent space $T_1(G)$ can be identified with A.

In particular, if A is the algebra $\mathcal{L}(V)$ of all linear endomorphisms of a linear space V, this group is the group $\mathrm{GL}(V)$. If $V = \mathbb{R}^n$, the algebra $\mathcal{L}(V)$ is the algebra $M_n(\mathbb{R})$ of $n \times n$ real matrices and the corresponding group is the *real general linear group* $\mathrm{GL}(n, \mathbb{R})$. Its dimension is equal to n^2.

Let $\det : \mathrm{GL}(n,\mathbb{R}) \longrightarrow \mathbb{R}^*$ be the determinant map. Then it defines a Lie group homomorphism of $\mathrm{GL}(n,\mathbb{R})$ into $\mathbb{R}^*$. Its kernel is the *real special linear group* $\mathrm{SL}(n,\mathbb{R})$.

The tangent space at $I \in M_n(\mathbb{R})$ can be identified with $M_n(\mathbb{R})$. To calculate the differential of det at I, consider the function

$$t \longmapsto \det(I + tT) = 1 + t\,\mathrm{tr}(T) + t^2(\dots)$$

for arbitrary $T \in M_n(\mathbb{R})$. Since T is the tangent vector to the curve $t \longmapsto I + tT$ at $t = 0$, we see that the differential of det is the linear form $\mathrm{tr} : M_n(\mathbb{R}) \longrightarrow \mathbb{R}$.

It follows that the tangent space to $\mathrm{SL}(n, \mathbb{R})$ at I is equal to the subspace of all traceless matrices in $M_n(\mathbb{R})$.

Therefore, the dimension of $\mathrm{SL}(n, \mathbb{R})$ is equal to $n^2 - 1$.

Let A be a finite dimensional associative algebra over $\mathbb{R}$ with identity. An *involution* τ on A is a linear map $a \longmapsto a^\tau$ such that

(i) $(a^\tau)^\tau = a$ for any $a \in A$;
(ii) $(ab)^\tau = b^\tau a^\tau$ for all $a, b \in A$.

Clearly, τ is a linear isomorphism of A and

$$1^\tau = 1^\tau (1^\tau)^\tau = (1^\tau 1)^\tau = (1^\tau)^\tau = 1.$$

Let G be the Lie group of all regular elements in A. Let

$$H = \{g \in G \mid gg^\tau = g^\tau g = 1\}.$$

Then H is a subgroup of G. Moreover, it is a closed subset of A.

1.8.3. LEMMA. *The group H is a Lie subgroup of G.*

The tangent space $T_1(H)$ can be identified with the linear subspace $\{a \in A \mid a = -a^\tau\}$.

PROOF. The tangent space to G at 1 can be identified with A, by attaching to $a \in A$ the tangent vector at 1 to the line $\mathbb{R} \ni t \longmapsto 1 + ta$. Let $\Psi : A \longrightarrow A$ be the map $\Psi(a) = aa^\tau$. Then we can view it as the composition of $a \longmapsto (a, a^\tau)$ with the multiplication map. Since $a \longmapsto (a, a^\tau)$ is a linear map, its differential is also $a \longmapsto (a, a^\tau)$. Hence, by 1.8.1, we see that the differential of Ψ at 1 is given by $T_1(\Psi)(a) = a + a^\tau$.

Let $S = \{a \in A \mid a = a^\tau\}$. then S is a linear subspace of A and therefore a submanifold. The image of Ψ is in S. Therefore, $\Psi : A \longrightarrow S$ is differentiable. Moreover, by the above calculation, Ψ is a submersion at 1. Hence, there exists an open neighborhood U of 1 in G such that the restriction $\Psi : U \longrightarrow S$ is a submersion. By 1.1.4.4, $H \cap U = U \cap \Psi^{-1}(1)$ is a submanifold of G. This implies that $\gamma(h)(H \cap U) = H \cap h \cdot U$ is a submanifold of G for any $h \in H$. Therefore, H is a submanifold of G and a Lie subgroup of G. In addition, $T_1(H) = \ker T_1(\Psi) = \{a \in A \mid a = -a^\tau\}$. □

Let V be a finite dimensional real linear space and $\varphi : V \times V \longrightarrow \mathbb{R}$ a symmetric (resp. skewsymmetric) nondegenerate bilinear form. Then for any $T \in \mathcal{L}(V)$ there exists a unique $T^* \in \mathcal{L}(V)$ such that

$$\varphi(Tv, w) = \varphi(v, T^* w) \text{ for all } v, w \in V.$$

The mapping $T \longmapsto T^*$ is an involution on $\mathcal{L}(V)$. The Lie group

$$G = \{T \in \mathrm{GL}(V) \mid TT^* = T^*T = 1\}$$

is called the *orthogonal* (resp. *symplectic*) group of φ.

For example, if $V = \mathbb{R}^{p+q}$ and

$$\varphi(v, w) = \sum_{i=1}^{p} v_i w_i - \sum_{i=p+1}^{p+q} v_i w_i,$$

then the corresponding orthogonal group is denoted by $\mathrm{O}(p, q)$. It is a Lie subgroup of $\mathrm{GL}(p + q, \mathbb{R})$.

Then $\det : \mathrm{O}(p, q) \longrightarrow \mathbb{R}^*$ is a Lie group homomorphism. Its kernel is the *special orthogonal group* $\mathrm{SO}(p, q)$ which is also a Lie subgroup of the special linear group $\mathrm{SL}(p + q, \mathbb{R})$.

If $V = \mathbb{R}^{2n}$ and

$$\varphi(v, w) = \sum_{i=1}^{n} (v_i w_{n+i} - v_{n+i} w_i),$$

then the corresponding symplectic group is denoted by $\mathrm{Sp}(n, \mathbb{R})$. It is a Lie subgroup of $\mathrm{GL}(2n, \mathbb{R})$.

Consider now the Lie subgroup $\mathrm{O}(n) = \mathrm{O}(n, 0)$ of $\mathrm{GL}(n, \mathbb{R})$. For any $T \in \mathrm{O}(n)$, its matrix entries are in $[-1, 1]$. Therefore, $\mathrm{O}(n)$ is a bounded closed submanifold of $M_n(\mathbb{R})$. It follows that $\mathrm{O}(n)$ is a compact Lie group. Hence, it has finitely many connected components.

Clearly, for a matrix $T \in \mathrm{O}(n)$, T^* is its transpose. Therefore, $\det(T) = \det(T^*)$ and

$$1 = \det(I) = \det(TT^*) = \det(T)\det(T^*) = (\det(T))^2$$

i.e., $\det(T) = \pm 1$. It follows that the homomorphism det maps $\mathrm{O}(n)$ onto the subgroup $\{\pm 1\}$ of $\mathbb{R}^*$. Therefore, $\mathrm{SO}(n)$ is a normal Lie subgroup of $\mathrm{O}(n)$ of index 2. In particular, $\mathrm{SO}(n)$ is open in $\mathrm{O}(n)$. Moreover, it has finitely many connected components.

The natural action of the group $\mathrm{O}(n)$ on $\mathbb{R}^n$ preserves the euclidean distance. Therefore, it acts differentiably on the unit sphere $S^{n-1} = \{x \in \mathbb{R}^n \mid x_1^2 + x_2^2 + \cdots + x_n^2 = 1\}$. Let $e = (1, 0, \ldots, 0) \in S^{n-1}$. Consider the orbit map $\rho(e) : \mathrm{SO}(n) \longrightarrow \mathbb{R}^n$ given by $T \longrightarrow T \cdot e$ (viewing e as a column vector). Then

$$\rho(e)(T) = T \cdot e = \begin{pmatrix} T_{11} & \cdots & T_{1n} \\ \vdots & \ddots & \vdots \\ T_{n1} & \cdots & T_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} T_{11} \\ \vdots \\ T_{n1} \end{pmatrix} = t,$$

where t is the first column of the matrix T.

Let s be a point on the sphere S^{n-1}. Then we can view $s_1 = s$ as a vector in $\mathbb{R}^n$ of length 1. We can add vectors $s_2, \ldots, s_n$ so that $s_1, s_2, \ldots, s_n$ is an orthonormal basis of $\mathbb{R}^n$. Let S be an orthogonal matrix with columns $s_1, s_2, \ldots, s_n$. As we remarked before, its determinant is ± 1. By changing the signs in the last column s_n if necessary, we can assume that it is 1, i.e., S is in $\mathrm{SO}(n)$. Then, by above calculation, we have $s = \rho(e)(S)$. Hence, $\mathrm{SO}(n)$ acts transitively on S^{n-1}.

Moreover, $S \in \mathrm{SO}(n)$ is in the stabilizer of e if and only if the first column of S is equal to e. Since other columns are orthogonal to e, their first coordinates are 0. Hence, the stabilizer of e in $\mathrm{SO}(n)$ is the group

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & T \end{pmatrix} \;\middle|\; T \in \mathrm{SO}(n-1) \right\}.$$

which is isomorphic to $\mathrm{SO}(n-1)$. Therefore, by 1.5.6, the orbit map induces a diffeomorphism of $\mathrm{SO}(n)/\mathrm{SO}(n-1)$ with S^{n-1}.

The dimension of $\mathrm{SO}(n)$ is equal to the dimension of its tangent space at I. Therefore, by 1.8.3, it is equal to the dimension of the space of all real skewsymmetric $n \times n$ matrices, i.e., we have

$$\dim \mathrm{SO}(n) = \frac{n(n-1)}{2}.$$

1.8.4. Lemma. *For any $n \in \mathbb{N}$, the group $\mathrm{SO}(n)$ is a connected compact Lie group of dimension $\frac{n(n-1)}{2}$.*

For any $n \in \mathbb{N}$, the group $\mathrm{O}(n)$ has two connected components. The identity component is $\mathrm{SO}(n)$.

We need to prove the first statement only. It is a consequence of the following discussion.

Let G be a Lie group and H a Lie subgroup of G. Denote by G_0 and H_0 the identity components of G and H respectively. Then we have a natural Lie group homomorphism $H \longrightarrow G \longrightarrow G/G_0$ from H into the discrete group G/G_0. Therefore, this homomorphism factors through H/H_0, i.e., we have a natural homomorphism $H/H_0 \longrightarrow G/G_0$ of discrete groups.

1.8.5. Lemma. *Let G be a Lie group and H its Lie subgroup. Assume that G/H is connected. Then the homomorphism $H/H_0 \longrightarrow G/G_0$ is surjective.*

Proof. Let e be the identity coset in G/H. Then the orbit map $\rho(e) : G \longrightarrow G/H$ is a submersion. Let G_0 be the identity component of G. Then, the restriction of $\rho(e)$ to G_0 is also a submersion. It follows that the orbit of e under G_0 is open. Since G_0 is a normal subgroup, all orbits of G_0 in G/H are open. Moreover, since G/H is connected, it follows that G_0 acts transitively on G/H.

Let $g \in G$. Then there exists $h \in G_0$ such that $ge = he$. It follows that $h^{-1}ge = e$ and $h^{-1}g$ is in the stabilizer of e, i.e., in H. It follows that g is in a G_0-coset intersecting H, i.e., its coset is in the image of the homomorphism $H/H_0 \longrightarrow G/G_0$. □

This result has the following immediate consequence.

1.8.6. Lemma. *Let G be a Lie group and H its Lie subgroup. Assume that H and G/H are connected. Then G is a connected Lie group.*

Proof. By our assumption, $H/H_0 = \{1\}$. By 1.8.5, it follows that $G/G_0 = \{1\}$. □

Now we prove 1.8.4 by induction in $n \in \mathbb{N}$. If $n = 1$, $\mathrm{SO}(1) = \{1\}$ and the statement is obvious. Moreover, S^{n-1} is connected for $n \geq 2$. By inducition, we can assume that $\mathrm{SO}(n-1)$ is connected. As we remarked above, $\mathrm{SO}(n)/\mathrm{SO}(n-1)$ is diffeomorphic to S^{n-1}. Hence, the assertion follows from 1.8.6.

Consider now $V = \mathbb{C}^n$. The algebra of complex linear transformations on V can be identified with the algebra $M_n(\mathbb{C})$ of $n \times n$ complex matrices. It can be viewed as a real algebra with identity. The corresponding group of regular elements is the group $\mathrm{GL}(n, \mathbb{C})$ of regular matrices in $M_n(\mathbb{C})$. It is called the *complex general linear group*. Clearly, it is an open submanifold of $M_n(\mathbb{C})$ and also a Lie group. Its tangent space at I can be identified with $M_n(\mathbb{C})$. Therefore the dimension of $\mathrm{GL}(n, \mathbb{C})$ is equal to $2n^2$.

The determinant $\det : \mathrm{GL}(n,\mathbb{C}) \longrightarrow \mathbb{C}^*$ is again a Lie group homomorphism. Its kernel is the *complex special linear group* $\mathrm{SL}(n,\mathbb{C})$. As before, we can calculate its differential which is the complex linear form $\mathrm{tr} : M_n(\mathbb{C}) \longrightarrow \mathbb{C}$. Therefore, the tangent space to $\mathrm{SL}(n,\mathbb{C})$ at I can be identified with the space of traceless matrices in $M_n(\mathbb{C})$. It follows that the dimension of $\mathrm{SL}(n,\mathbb{C})$ is equal to $2n^2 - 2$.

Let $V = \mathbb{C}^{p+q}$ and

$$\varphi(v,w) = \sum_{i=1}^{p} v_i \bar{w}_i - \sum_{i=p+1}^{p+q} v_i \bar{w}_i$$

for $v, w \in V$. This form is linear in the first variable and antilinear in the second, but if we forget the complex structure, it is bilinear. Therefore, the above discussion applies again. If $T \longmapsto T^*$ is the corresponding involution on $M_n(\mathbb{C})$, the group $H = \{T \in \mathrm{GL}(n,\mathbb{C}) \mid TT^* = T^*T = 1\}$, is called the *unitary group* with respect to φ and denoted by $\mathrm{U}(p,q)$.

If $V = \mathbb{C}^n$, we put $\mathrm{U}(n) = \mathrm{U}(n,0)$. In this case T^* is the hermitian adjoint of the matrix T. The absolute values of all matrix entries of $T \in \mathrm{U}(n)$ are ≤ 1. Therefore, $\mathrm{U}(n)$ is a bounded closed submanifold of $M_n(\mathbb{C})$. It follows that $\mathrm{U}(n)$ is a compact Lie group. In particular, it has finitely many connected components. In addition, we have

$$1 = \det(TT^*) = \det(T)\det(T)^* = |\det(T)|^2$$

for $T \in \mathrm{U}(n)$, i.e. det is a Lie group homomorphism of $\mathrm{U}(n)$ into the multiplicative group of complex numbers of absolute value 1. The kernel of this homomorphism is the *special unitary group* $\mathrm{SU}(n)$. It is also compact and has finitely many connected components.

By 1.8.3, the tangent space to $\mathrm{U}(n)$ at I is equal to the space of all skewadjoint matrices in $M_n(\mathbb{C})$. Therefore, we have

$$\dim \mathrm{U}(n) = n^2.$$

The tangent space to $\mathrm{SU}(n)$ at I is the kernel of the linear map induced by tr, i.e., the space of all traceless skewadjoint matrices in $M_n(\mathbb{C})$. Therefore, we have

$$\dim \mathrm{SU}(n) = n^2 - 1.$$

The natural action of the group $\mathrm{U}(n)$ on $\mathbb{C}^n$ preserves the euclidean distance. Therefore, it acts differentiably on the unit sphere $S^{2n-1} = \{z \in \mathbb{C}^n \mid |z_1|^2 + |z_2|^2 + \cdots + |z_n|^2 = 1\}$. Let $e = (1, 0, \ldots, 0) \in S^{2n-1}$. Consider the orbit map $\rho(e) : \mathrm{SU}(n) \longrightarrow \mathbb{C}^n$ given by $T \longrightarrow T \cdot e$ (viewing e as a column vector). Then

$$\rho(e)(T) = T \cdot e = \begin{pmatrix} T_{11} & \cdots & T_{1n} \\ \vdots & \ddots & \vdots \\ T_{n1} & \cdots & T_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} T_{11} \\ \vdots \\ T_{n1} \end{pmatrix} = t,$$

where t is the first column of the matrix T.

Let s be a point on the sphere S^{2n-1}. Then we can view $s_1 = s$ as a vector in $\mathbb{C}^n$ of length 1. We can add vectors $s_2, \ldots, s_n$ so that $s_1, s_2, \ldots, s_n$ is an orthonormal basis of $\mathbb{C}^n$. Let S be a unitary matrix with columns $s_1, s_2, \ldots, s_n$. As we remarked before, its determinant is a complex number of modulus 1. By multiplying the last column s_n by the complex conjugate of this determinant if necessary, we can assume that the determinant is equal to 1, i.e., S is in $\mathrm{SU}(n)$. Then, by above calculation, we have $s = \rho(e)(S)$. Hence, $\mathrm{SU}(n)$ acts transitively on S^{2n-1}.

Moreover, $S \in \mathrm{SU}(n)$ is in the stabilizer of e if and only if the first column of S is equal to e. Since other columns are orthogonal to e, their first coordinates are 0. Hence, the stabilizer of e in $\mathrm{SU}(n)$ is the group

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & T \end{pmatrix} \,\middle|\, T \in \mathrm{SU}(n-1) \right\}.$$

which is isomorphic to $\mathrm{SU}(n-1)$. Therefore, by 1.5.6, the orbit map induces a diffeomorphism of $\mathrm{SU}(n)/\mathrm{SU}(n-1)$ with S^{2n-1}.

Clearly, $\mathrm{SU}(1) = \{1\}$. By induction in n, this leads to the following result.

1.8.7. Lemma. *For any $n \in \mathbb{N}$, the group $\mathrm{SU}(n)$ is a connected compact Lie group of dimension $n^2 - 1$.*

Proof. This follows immediately from the above discussion and 1.8.6 as in the proof of 1.8.4. □

On the other hand, $\mathrm{U}(n)/\mathrm{SU}(n)$ is isomorphic to the multiplicative group of complex numbers of absolute value 1. Hence, applying 1.8.6 again, we conclude that the following result holds.

1.8.8. Corollary. *The group $\mathrm{U}(n)$ is a connected compact Lie group.*

Let $\mathbb{H}$ denote the quaternions (for all unexplained notation about quaternions see 1.9). Let $M_n(\mathbb{H})$ be the associative algebra of n-by-n matrices with quaternionic coefficients. Denote by I the identity matrix in $N_n(\mathbb{H})$. Clearly, $M_n(\mathbb{H})$ is isomorphic to $\mathbb{R}^{4n}$ as a real vector space. Hence, it has a natural structure of a $4n$-dimensional manifold. The matrix multiplication on $M_n(\mathbb{H})$ is a differentiable map from $M_n(\mathbb{H}) \times M_n(\mathbb{H})$ into $M_n(\mathbb{H})$. Therefore, all ivertible matrices in $M_n(\mathbb{H})$ form an open submanifold of $M_n(\mathbb{H})$ which is the Lie group $\mathrm{GL}(n, \mathbb{H})$ of dimension $4n^2$ by 1.8.2. Let $T \longmapsto T^*$ be the involution on $M_n(\mathbb{H})$ which attaches to a matrix T its adjoint T^*. We put

$$\mathrm{Sp}(n) = \{T \in M_n(\mathbb{H}) \mid TT^* = T^*T = I\}.$$

By 1.8.3, $\mathrm{Sp}(n)$ is a Lie subgroup of the Lie group $\mathrm{GL}(n, \mathbb{H})$. We call it the *symplectic group* $\mathrm{Sp}(n)$.

Clearly, $\mathrm{Sp}(n)$ is a closed submanifold of $M_n(\mathbb{H})$. Also, all coefficients of a matrix T in $\mathrm{Sp}(n)$ have norm ≤ 1 by 1.9.2. Hence, $\mathrm{Sp}(n)$ is a bounded subset in $\mathbb{R}^{4n}$. Therefore, $\mathrm{Sp}(n)$ is a compact Lie group. In particular, it has finitely many connected components.

The tangent space to $\mathrm{Sp}(n)$ at I is the space of all skewadjoint matrices in $M_n(\mathbb{H})$. Therefore, we have

$$\dim \mathrm{Sp}(n) = 4\frac{n^2 - n}{2} + 3n = 2n^2 + n.$$

The natural action of the group $\mathrm{Sp}(n)$ on $\mathbb{H}^n$ preserves the norm by 1.9.3. Therefore, it acts differentiably on the unit sphere $S^{4n-1} = \{q \in \mathbb{H}^n \mid |q_1|^2 + |q_2|^2 + \cdots + |q_n|^2 = 1\}$. Let $e = (1, 0, \ldots, 0) \in S^{4n-1}$. Consider the orbit map

$\rho(e) : \mathrm{Sp}(n) \longrightarrow \mathbb{H}^n$ given by $T \longrightarrow T \cdot e$ (viewing e as a column vector). Then

$$\rho(e)(T) = T \cdot e = \begin{pmatrix} T_{11} & \cdots & T_{1n} \\ \vdots & \ddots & \vdots \\ T_{n1} & \cdots & T_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} T_{11} \\ \vdots \\ T_{n1} \end{pmatrix} = t,$$

where t is the first column of the matrix T. If $S \in \mathrm{Sp}(n)$ is in the stabilizer of e if and only if the first column of S is equal to e. On the other hand, $S^{-1} = S^*$ is also in the stabilizer. That implies that the first column of S^* is also in the stabilizer. Therefore, the first row of S is $(1, 0, \ldots, 0)$. Hence, the stabilizer of e in $\mathrm{Sp}(n)$ is the group

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & T \end{pmatrix} \,\middle|\, T \in \mathrm{Sp}(n-1) \right\}.$$

which is isomorphic to $\mathrm{Sp}(n-1)$. Clearly, we have

$$\dim \mathrm{Sp}(n)/\mathrm{Sp}(n-1) = \dim \mathrm{Sp}(n) - \dim Sp(n-1) = 2n^2 + n - 2(n-1)^2 - n + 1 = 4n - 1.$$

By 1.3.1, the the orbit map induces a diffeomorphism of $\mathrm{Sp}(n)/\mathrm{Sp}(n-1)$ onto an open set in S^{4n-1}. In particular, the orbit of e is open in the sphere S^{4n-1}. Since $\mathrm{Sp}(n)$ is compact, this orbit is also closed in S^{4n-1}. Since S^{4n-1} is connected, we see that the orbit is equal to S^{4n-1}.

Irt follows that the orbit map induces a diffeomorphism of $\mathrm{Sp}(n)/\mathrm{Sp}(n-1)$ onto S^{4n-1}.

The group $\mathrm{Sp}(1)$ consists of quaternions q such that $q\bar{q} = \bar{q}q = 1$. By 1.9.1 this is equivalent to q being of norm 1. Hence, we have

$$\mathrm{Sp}(1) = \{q \in \mathbb{H} \mid |q| = 1\} = \{q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \mid a^2 + b^2 + c^2 + d^2 = 1\}.$$

Hence, $\mathrm{Sp}(1)$ is diffeomorphic to the sphere S^3. In particular, it is connected and simply connected.

This leads to the following result.

1.8.9. Lemma. *For any $n \in \mathbb{N}$, the group $\mathrm{Sp}(n)$ is a connected compact Lie group.*

Proof. This follows immediately from the above discussion and 1.8.6 as in the proof of 1.8.4. □

Now we want to study the fundamental groups of the above examples.

Assume that G is a Lie group and H a Lie subgroup. Let H_0 be the identity component of H. Then H_0 is a Lie subgroup of G. Moreover, the natural quotient morphism $p : G \longrightarrow G/H$ is constant on left H_0-cosets. Therefore it induces a differentiable map $\pi : G/H_0 \longrightarrow G/H$. Let $p_0 : G \longrightarrow G/H_0$ be the natural quotient morphism. Then we have $\pi \circ p_0 = p$. Since p is a submersion, π has to be a submersion. Moreover, we have

$$\dim(G/H) = \dim G - \dim H = \dim G - \dim H_0 = \dim(G/H_0),$$

and π is a local diffeomorphism.

Consider the differentiable map $H \times G \longrightarrow G$ given by $\mu_r(h, g) = gh^{-1}$ for $h \in H$ and $g \in G$. It defines an action of H on G. The composition of this map with p_0 defines a differentiable map $\lambda : H \times G \longrightarrow G/H_0$. We can view $H_0 \times H_0$ as a Lie subgroup of $H \times G$. Clearly, since H_0 is a normal subgroup in H, we have

$$\lambda(hh_0, gh_1) = gh_1h_0^{-1}h^{-1}H_0 = gh^{-1}(hh_1h_0^{-1}h^{-1})H_0 = gh^{-1}H_0$$

for any $g \in G$, $h \in H$ and $h_0, h_1 \in H_0$. Hence, the differentiable map λ factors through a differentiable map $\kappa : H/H_0 \times G/H_0 \longrightarrow G/H_0$ which satisfies

$$\kappa(hH_0, gH_0) = gh^{-1}H_0$$

for any $g \in G$ and $h \in H$. Clearly, this is a differentiable action of the discrete group H/H_0 on G/H_0. The orbits on this action in G/H_0 are the projections of left H-cosets in G in G/H_0. Therefore, G/H is the quotient manifold of G/H_0 for that action. Hence, the action of the discrete group H/H_0 on G/H_0 is regular. By 1.4.1, this action is free. By (iii) in 1.4.1, we also see that π is a covering projection.

Assume that G is a connected Lie group. Then G/H and G/H_0 are connected manifolds, and $\pi : G/H_0 \longrightarrow G/H$ is a covering projection. Moreover, the order of the covering is equal to $\mathrm{Card}(H/H_0)$, i.e., to the number of the components of H.

1.8.10. LEMMA. *Let G be a connected Lie group and H a Lie subgroup of G. If G/H is simply connected, the group H is connected.*

In addition, the natural morphism of fundamental groups $\pi_1(H, 1) \longrightarrow \pi_1(G, 1)$ is surjective.

PROOF. Let H_0 be the identity component of H. Then $\pi : G/H_0 \longrightarrow G/H$ is a covering projection. Since G/H is simply connected, π has to be a diffeomorphism, i.e., the order of the covering is 1. This implies that $H = H_0$.

Let $\tilde{G}$ be the universal cover of G and $r : \tilde{G} \longrightarrow G$ the natural projection. Since r is a local diffeomorphism, $K = r^{-1}(H)$ is a Lie subgroup of $\tilde{G}$. Moreover, the differentiable map $p \circ r : \tilde{G} \longrightarrow G/H$ induces a diffeomorphism $\tilde{G}/K \longrightarrow G/H$. Hence, $\tilde{G}/K$ is simply connected. By the first part, K is connected. Let $C = \ker r$. As we established before, C is isomorphic to the fundamental group of G. Then $C \subset K$ is a discrete central subgroup of K and $K/C = H$. It follows that the universal covering group of K is also the universal covering group $\tilde{H}$ of H. Moreover, C is a quotient of the kernel of the projection $\tilde{H} \longrightarrow H$, which is isomorphic to the fundamental group of H. □

This has the following consequence.

1.8.11. COROLLARY. *Let G be a connected Lie group and H a connected Lie subgroup of G. If H and G/H are simply connected, G is also simply connected.*

We know that $\mathrm{SU}(n)$, $n \geq 1$, are connected by 1.8.7, all spheres of dimension ≥ 2 are simply connected and $\mathrm{SU}(1) = \{1\}$. Hence, by induction in n, from the isomorphism of $\mathrm{SU}(n)/\mathrm{SU}(n-1)$ with S^{2n-1} for $n \geq 2$, the next result follows.

1.8.12. LEMMA. *All groups $\mathrm{SU}(n)$, $n \geq 1$, are simply connected.*

We know that $\mathrm{Sp}(n)$, $n \geq 1$, are connected by 1.8.9, all spheres of dimension ≥ 2 are simply connected and $\mathrm{Sp}(1)$ is simply connected. Hence, by induction in n, from the isomorphism of $\mathrm{Sp}(n)/\mathrm{Sp}(n-1)$ with S^{4n-1} for $n \geq 2$, the next result follows.

1.8.13. LEMMA. *All groups $\mathrm{Sp}(n)$, $n \geq 1$, are simply connected.*

Now we want to discuss some low dimensional examples. Let $T \in \mathrm{SL}(2,\mathbb{C})$ be given by the matrix

$$T = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} ,$$

with $\alpha, \beta, \gamma, \delta \in \mathbb{C}$ satisfying $\alpha\delta - \beta\gamma = 1$, then its inverse is

$$T^{-1} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix} .$$

If $T \in \mathrm{SU}(2)$, then we must also have

$$T^{-1} = T^* = \begin{pmatrix} \bar{\alpha} & \bar{\gamma} \\ \bar{\beta} & \bar{\delta} \end{pmatrix} .$$

Therefore, $\delta = \bar{\alpha}$ and $\gamma = -\bar{\beta}$. It follows that

$$T = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} ,$$

with $|\alpha|^2 + |\beta|^2 = 1$. Therefore, $\mathrm{SU}(2)$ is diffeomorphic to a three dimensional unit sphere S^3 in $\mathbb{C}^2$. In particular, $\mathrm{SU}(2)$ is simply connected, as we already established above.

Moreover, from 1.9, we see that $\mathrm{SU}(2)$ is isomorphic to $\mathrm{Sp}(1)$ as a Lie group.

We identify $\mathbb{R}^3$ with the space $\mathcal{H}$ of traceless selfadjoint 2×2 matrices via the map:

$$H : (x, y, z) \longmapsto \begin{pmatrix} x & y + iz \\ y - iz & -x \end{pmatrix} .$$

Then

$$\det H(x, y, z) = -(x^2 + y^2 + z^2),$$

i.e., it is the negative of the square of the distance from the origin to the point (x, y, z). Clearly, for any $T \in \mathrm{SU}(2)$ and $S \in \mathcal{H}$, the matrix $TST^* = TST^{-1}$ satisfies

$$(TST^*)^* = TS^*T^* = TST^* \text{ and } \operatorname{tr}(TST^*) = \operatorname{tr}(ST^*T) = \operatorname{tr}(S) = 0,$$

i.e., is again selfadjoint and traceless. Therefore, the map $\psi(T) : S \longmapsto TST^*$ is a representation of $\mathrm{SU}(2)$ on the real linear space $\mathcal{H}$. Clearly, $\det(TST^*) = \det(S)$, Hence, if we identify $\mathcal{H}$ with $\mathbb{R}^3$ using H, we see that the action of $\mathrm{SU}(2)$ on $\mathbb{R}^3$ is by orthogonal matrices. Therefore, we constructed a continuous homomorphism ψ of $\mathrm{SU}(2)$ in the group of $\mathrm{O}(3)$. Since $\mathrm{SU}(2)$ is connected, this is a homomorphism of $\mathrm{SU}(2)$ into $\mathrm{SO}(3)$.

Since we have

$$TH(x,y,z)T^* = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \begin{pmatrix} x & y + iz \\ y - iz & -x \end{pmatrix} \begin{pmatrix} \bar{\alpha} & -\beta \\ \bar{\beta} & \alpha \end{pmatrix} =$$
$$\begin{pmatrix} (|\alpha|^2 - |\beta|^2)x + 2\operatorname{Re}(\alpha\bar{\beta})y - 2\operatorname{Im}(\alpha\bar{\beta})z & -2\alpha\beta x + (\alpha^2 - \beta^2)y + i(\alpha^2 + \beta^2)z \\ -2\bar{\alpha}\bar{\beta}x + (\bar{\alpha}^2 - \bar{\beta}^2)y - i(\bar{\alpha}^2 + \bar{\beta}^2)z & -(|\alpha|^2 - |\beta|^2)x - 2\operatorname{Re}(\alpha\bar{\beta})y + 2\operatorname{Im}(\alpha\bar{\beta})z \end{pmatrix},$$

we see that

$$\psi(T) = \begin{pmatrix} |\alpha|^2 - |\beta|^2 & 2\operatorname{Re}(\alpha\bar{\beta}) & -2\operatorname{Im}(\alpha\bar{\beta}) \\ -2\operatorname{Re}(\alpha\beta) & \operatorname{Re}(\alpha^2 - \beta^2) & -\operatorname{Im}(\alpha^2 + \beta^2) \\ -2\operatorname{Im}(\alpha\beta) & \operatorname{Im}(\alpha^2 - \beta^2) & \operatorname{Re}(\alpha^2 + \beta^2) \end{pmatrix}.$$

Let T be in the kernel of ψ. Then $(1,1)$ coefficient of $\psi(T)$ has to be equal to 1, i.e., $|\alpha|^2 - |\beta|^2 = 1$. Since $|\alpha|^2 + |\beta|^2 = 1$, we see that $|\alpha| = 1$ and $\beta = 0$. Now, from the $(2,3)$ coefficient we see that $\operatorname{Im}(\alpha^2) = 0$ and from the $(2,2)$ coefficient we see that $\operatorname{Re}(\alpha^2) = 1$. It follows that $\alpha^2 = 1$ and $\alpha = \pm 1$. Hence, the kernel of ψ consists of matrices $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Therefore, the differential of ψ is injective. Since $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$ are three-dimensional, it follows that the differential of ψ is an isomorphism of tangent spaces at the identity. Since both groups are connected, it must be a covering projection by 1.6.6.

1.8.14. Lemma. *The fundamental group of* $\mathrm{SO}(3)$ *is* $\mathbb{Z}_2$. *Its universal covering group is* $\mathrm{SU}(2) = \mathrm{Sp}(1)$.

Clearly, $\mathrm{SO}(1) = \{1\}$ is simply connected. The group $\mathrm{SO}(2)$ is isomorphic to the circle S^1, and its fundamental group is $\mathbb{Z}$. On the other hand, from the isomorphism $\mathrm{SO}(n)/\mathrm{SO}(n-1)$ with S^{n-1}, the fact that spheres of dimension ≥ 2 are simply connected and 1.8.14, we see that the fundamental groups of $\mathrm{SO}(n)$, $n \geq 3$, are either trivial or equal to $\mathbb{Z}_2$.[1]

1.9. Appendix: Quaternions. Let $M_2(\mathbb{C})$ be the ring of 2×2 complex matrices. Denote by A the subset of all matrices of the form

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

where α and β are arbitrary complex numbers. Clearly, this subset is a four dimensional real linear subspace of $M_2(\mathbb{C})$.

In addition,

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \begin{pmatrix} \gamma & \delta \\ -\bar{\delta} & \bar{\gamma} \end{pmatrix} = \begin{pmatrix} \alpha\gamma - \beta\bar{\delta} & \alpha\delta + \beta\bar{\gamma} \\ -\bar{\beta}\gamma - \bar{\alpha}\bar{\delta} & -\bar{\beta}\delta + \bar{\alpha}\bar{\gamma} \end{pmatrix}$$

and we see that A is actually a real subalgebra of $M_2(\mathbb{C})$. The identity matrix I is the identity element in this algebra.

Moreover, the hermitian adjoint of the matrix

$$T = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$$

is the matrix

$$T^* = \begin{pmatrix} \bar{\alpha} & -\beta \\ \bar{\beta} & \alpha \end{pmatrix},$$

and the involution on $M_2(\mathbb{C})$ maps A into itself. Therefore, it induces involution on A.

Also, we have

$$\det T = \begin{vmatrix} \bar{\alpha} & \beta \\ -\bar{\beta} & \alpha \end{vmatrix} = |\alpha|^2 + |\beta|^2 \geq 0.$$

So, every nonzero element of A is invertible. The inverse of $T \neq 0$ is

$$T^{-1} = \frac{1}{\det T} \begin{pmatrix} \bar{\alpha} & -\beta \\ \bar{\beta} & \alpha \end{pmatrix} = \frac{1}{\det T} T^*,$$

i.e., it is in A, and A is a division algebra.

[1] One can actually show that they are all isomorphic to $\mathbb{Z}_2$.

We can pick a basis of A as a real linear space

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{i} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \mathbf{j} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathbf{k} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Then an element q of A is represented uniquely as

$$q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$$

for $a, b, c, d \in \mathbb{R}$. The element $\mathbf{1}$ is the identity in A and the rest of multiplication table is given by

$$\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = -\mathbf{1}$$

and

$$\mathbf{i} \cdot \mathbf{j} = -\mathbf{j} \cdot \mathbf{i} = \mathbf{k}, \quad \mathbf{i} \cdot \mathbf{k} = -\mathbf{k} \cdot \mathbf{i} = -\mathbf{j}, \quad \mathbf{j} \cdot \mathbf{k} = -\mathbf{k} \cdot \mathbf{j} = \mathbf{i}.$$

Therefore, A is as a real associative algebra with identity isomorphic to the algebra of quaternions $\mathbb{H}$.

The involution on A corresponds to the map $q \longmapsto \bar{q}$ on quaternions given by

$$\bar{q} = a\mathbf{1} - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}.$$

Therefore, $q \longmapsto \bar{q}$ is an involution on $\mathbb{H}$.

Moreover, we have

$$q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} = \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix},$$

and, since

$$\begin{vmatrix} a + bi & c + di \\ -c + di & a - bi \end{vmatrix} = a^2 + b^2 + c^2 + d^2,$$

we see that the function

$$|q| = \sqrt{a^2 + b^2 + c^2 + d^2}$$

on $\mathbb{H}$ satisfies $|q| \geq 0$ and $|q| = 0$ if and only if $q = 0$. Moreover,

$$|p + q| \leq |p| + |q|$$

for any $p, q \in \mathbb{H}$. We call this function the *norm* on $\mathbb{H}$.

Moreover, by the above formula we see that

$$q\bar{q} = \bar{q}q = |q|^2\mathbf{1}.$$

Hence, for any nonzero quaternion $q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ we have

$$q^{-1} = \frac{1}{|q|^2}\bar{q}.$$

Moreover, if q and q' are two quaternions given by

$$q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \text{ and } q' = a'\mathbf{1} + b'\mathbf{i} + c'\mathbf{j} + d'\mathbf{k},$$

we have

$$|qq'| = \det\left(\begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix}\begin{pmatrix} a' + b'i & c' + d'i \\ -c' + d'i & a' - b'i \end{pmatrix}\right)$$
$$= \begin{vmatrix} a + bi & c + di \\ -c + di & a - bi \end{vmatrix}\begin{vmatrix} a' + b'i & c' + d'i \\ -c' + d'i & a' - b'i \end{vmatrix} = |q||q'|.$$

Hence, we have the following result.

1.9.1. Lemma. *For any two quaternions $q, q' \in \mathbb{H}$ we have*

$$|qq'| = |q||q'|.$$

Let $M_n(\mathbb{H})$ be the set of all n-by-n quaternion matrices with standard matrix addition and multiplication. Then $M_n(\mathbb{H})$ is an associative algebra over $\mathbb{R}$. Its identity is the identity matrix. The group of all invertible matrices in $M_n(\mathbb{H})$ is the *general linear group* $\mathrm{GL}(n, \mathbb{H})$.

We view elements of $\mathbb{H}^n$ as column vectors and define the map $(T, v) \longmapsto Tv$ from $M_n(\mathbb{H}) \times \mathbb{H}^n$ into $\mathbb{H}^n$ given by matrix multiplication.

For any T in $M_n(\mathbb{H})$ we denote by T^* the matrix such that $(T^*)_{ij} = \bar{T}_{ji}$ for any $1 \leq i, j \leq n$. By direct calculation, one easily checks that $T \longmapsto T^*$ is an involution on $M_n(\mathbb{H})$. We denote by $\mathrm{Sp}(n)$ the subgroup of $\mathrm{GL}(n, \mathbb{H})$ consisting of matrices T such that $T^{-1} = T^*$. Such matrices are called *symplectic*, and $\mathrm{Sp}(n)$ is the *symplectic group*.

Let T be a symplectic matrix. Then $T^*T = I$, i.e.,

$$\sum_{k=1}^{n} T^*_{ik} T_{kj} = \sum_{k=1}^{n} \bar{T}_{ki} T_{kj} = \delta_{ij} \mathbf{1}$$

for all $1 \leq i, j \leq n$. In particular, we have

$$\sum_{k=1}^{n} |T_{ki}|^2 = 1$$

for any $1 \leq i \leq n$. This implies the following result.

1.9.2. LEMMA. *Let T be a symplectic matrix. Then $|T_{ij}| \leq 1$ for any $1 \leq i, j \leq n$.*

We define a norm $q \longmapsto \|q\|$ on $\mathbb{H}^n$ by

$$\|q\|^2 = \sum_{i=1}^{n} |q_i|^2.$$

Let T be a symplectic matrix in $M_n(\mathbb{H})$. For $v \in \mathbb{H}^n$ we have

$$\begin{aligned} \|Tv\|^2 \mathbf{1} &= \sum_{i=1}^{n} |(Tv)_i|^2 \mathbf{1} = \sum_{i=1}^{n} |\sum_{j=1}^{n} T_{ij} v_j|^2 \mathbf{1} \\ &= \sum_{i=1}^{n} \overline{\sum_{j=1}^{n} T_{ij} v_j} \sum_{k=1}^{n} T_{ik} v_k = \sum_{i,j,k=1}^{n} \bar{v}_j \bar{T}_{ij} T_{ik} v_k = \sum_{i,j,k=1}^{n} \bar{v}_j T^*_{ji} T_{ik} v_k \\ &= \sum_{j,k=1}^{n} \bar{v}_j \delta_{jk} v_k = \sum_{j=1}^{n} \bar{v}_j v_j = \|v\|^2 \mathbf{1}. \end{aligned}$$

Therefore, we have the following result.

1.9.3. LEMMA. *Let T be a symplectic matrix in $M_n(\mathbb{H})$. Then*

$$\|Tv\| = \|v\|$$

for any $v \in \mathbb{H}^n$.

2. Lie algebra of a Lie group

2.1. Lie algebras. A Lie algebra $\mathfrak{a}$ over a field k of characteristic 0 is a linear space over k with a bilinear operation $(x, y) \longmapsto [x, y]$ such that

(i) $[x, x] = 0$ for all $x \in \mathfrak{a}$;

(ii) (Jacobi identity) $[x,[y,z]]+[y,[z,x]]+[z,[x,y]]=0$ for all $x,y,z\in\mathfrak{a}$.

The operation $(x,y)\longmapsto[x,y]$ is called the commutator. The condition (i) implies that

$$0=[x+y,x+y]=[x,x]+[x,y]+[y,x]+[y,y]=[x,y]+[y,x]$$

i.e.

$$[x,y]=-[y,x]$$

for all $x,y\in\mathfrak{a}$.

A k-linear map φ between Lie algebra $\mathfrak{a}$ and $\mathfrak{b}$ is a *morphism of Lie algebras* if

$$\varphi([x,y])=[\varphi(x),\varphi(y)] \text{ for all } x,y\in\mathfrak{a}.$$

Lie algebras over k and morphisms of Lie algebras for the category of Lie algebras.

If A is an associative algebra, we can define $[S,T]=ST-TS$ for all $S,T\in A$. By direct calculation one can check that A with this commutator becomes a Lie algebra. This defines a functor from the category of associative algebras into the category of Lie algebras.

In particular, if V is a linear space over k and $\mathcal{L}(V)$ the algebra of all linear transformations on V, the commutator defines on $\mathcal{L}(V)$ a structure of a Lie algebra. This Lie algebra is denoted by $\mathfrak{gl}(V)$.

A Lie algebra homomorphism $\psi:\mathfrak{a}\longrightarrow\mathfrak{gl}(V)$ is called a *representation* of $\mathfrak{a}$ on V.

Let $\mathfrak{a}$ be a Lie algebra. For $x\in\mathfrak{a}$ we denote by $\operatorname{ad}(x)$ the linear transformation on $\mathfrak{a}$ defined by $\operatorname{ad}(x)(y)=[x,y]$ for all $y\in\mathfrak{a}$.

2.1.1. LEMMA. ad *is a representation of* $\mathfrak{a}$ *on* $\mathfrak{a}$.

PROOF. Let $x,y\in\mathfrak{a}$. Then, by the Jacobi identity, we have

$$\begin{aligned}\operatorname{ad}([x,y])(z)&=[[x,y],z]=-[z,[x,y]]=[x,[y,z]]+[y,[z,x]]\\&=\operatorname{ad}(x)(ad(y)(z))-\operatorname{ad}(y)(\operatorname{ad}(x)(z))=[\operatorname{ad}(x),\operatorname{ad}(y)](z)\end{aligned}$$

for any $z\in\mathfrak{a}$. □

This representation is called the *adjoint representation* of $\mathfrak{a}$.

Let $\mathfrak{b}$ be a linear subspace of $\mathfrak{a}$. If $x,y\in\mathfrak{b}$ imply that $[x,y]\in\mathfrak{b}$, the restriction of the commutator to $\mathfrak{b}$ defines a structure of Lie algebra on $\mathfrak{b}$. The Lie algebra $\mathfrak{b}$ is called the *Lie subalgebra* of $\mathfrak{a}$. Let $\mathfrak{b}$ be such that $x\in\mathfrak{a}$ and $y\in\mathfrak{b}$ imply that $[x,y]\in\mathfrak{b}$. Then the Lie subalgebra $\mathfrak{b}$ is an *ideal* in $\mathfrak{a}$.

Let $\mathfrak{a}$ be a Lie algebra and $\mathfrak{b}$ an ideal in $\mathfrak{a}$. Let $x,x'\in\mathfrak{a}$ be two representatives of the same coset modulo $\mathfrak{b}$. Also, let $y,y'\in\mathfrak{a}$ be two representatives of the same coset modulo $\mathfrak{b}$. Then

$$[x,y]-[x',y']=[x-x',y]+[x',y-y']\in\mathfrak{b},$$

i.e., $[x,y]$ and $[x',y']$ are in the same coset modulo $\mathfrak{b}$. Therefore,

$$(x+\mathfrak{b},y+\mathfrak{b})\longmapsto[x,y]+\mathfrak{b}$$

is a well defined bilinear operation on $\mathfrak{a}/\mathfrak{b}$. Clearly, $\mathfrak{a}/\mathfrak{b}$ is a Lie algebra with that operation. It is called the *quotient Lie algebra* $\mathfrak{a}/\mathfrak{b}$ of $\mathfrak{a}$ modulo the ideal $\mathfrak{b}$.

2.1.2. LEMMA. *Let* $\varphi:\mathfrak{a}\longrightarrow\mathfrak{b}$ *be a morphism of Lie algebras. Then:*

(i) *The kernel* $\ker\varphi$ *of* φ *is an ideal in* $\mathfrak{a}$.

(ii) *The image* $\operatorname{im}\varphi$ *of* φ *is a Lie subalgebra in* $\mathfrak{b}$.

Let $\mathfrak{a}$ and $\mathfrak{b}$ be two Lie algebras. Then the linear space $\mathfrak{a} \times \mathfrak{b}$ with the commutator

$$[(x, y), (x', y')] = ([x, x'], [y, y'])$$

for $x, x' \in \mathfrak{a}$ and $y, y' \in \mathfrak{b}$ is a Lie algebra – the *product* $\mathfrak{a} \times \mathfrak{b}$ of Lie algebras $\mathfrak{a}$ and $\mathfrak{b}$.

Let $\mathfrak{a}$ be a Lie algebra. The *center* $\mathfrak{c}$ of $\mathfrak{a}$ is

$$\mathfrak{c} = \{x \in \mathfrak{a} \mid [x, y] = 0 \text{ for all } y \in \mathfrak{a}\}.$$

Clearly, $\mathfrak{c}$ is an ideal in $\mathfrak{a}$.

A Lie algebra $\mathfrak{a}$ is *abelian* if $[x, y] = 0$ for all $x, y \in \mathfrak{a}$.

Let $\mathfrak{a}$ be a Lie algebra. We denote by $\mathfrak{a}^{opp}$ the *opposite Lie algebra* of $\mathfrak{a}$. It is the same linear space with the commutator $(x, y) \longmapsto [x, y]^\circ = -[x, y]$. Clearly, $\mathfrak{a}^{opp}$ is a Lie algebra. Moreover, $x \longmapsto -x$ is an isomorphism of $\mathfrak{a}$ with $\mathfrak{a}^{opp}$. Evidently, we have $(\mathfrak{a}^{opp})^{opp} = \mathfrak{a}$.

If $\dim \mathfrak{a} = 1$, $\mathfrak{a}$ has to be abelian.

If $\dim \mathfrak{a} = 2$, we can pick a basis (v_1, v_2) of $\mathfrak{a}$ and see that $[x, y]$ is proportional to $[v_1, v_2]$ for any $x, y \in \mathfrak{a}$. Therefore, we can assume that $[x, y]$ is proportional to e_1 for any $x, y \in \mathfrak{a}$. If $\mathfrak{a}$ is not abelian, there exists e_2 such that $[e_1, e_2] = e_1$. Therefore, we conclude that up to an isomorphism there exists a unique nonabelian two dimensional Lie algebra over k.

Finally, we quote the following fundamental theorem of Ado (which will be proven later) which says that every finite-dimensional Lie algebra has a faithful finite-dimensional representation.

2.1.3. THEOREM (Ado). *Let* $\mathfrak{g}$ *be a finite-dimensional Lie algebra over* k. *Then* $\mathfrak{g}$ *is isomorphic to a Lie subalgebra of* $\mathfrak{gl}(V)$ *for some finite-dimensional linear space* V *over* k.

2.2. Lie algebra of a Lie group. Let G be a Lie group. Let $T_1(G)$ be the tangent space to G at the identity 1. If $\varphi : G \longrightarrow H$ is a morphism of Lie groups, $T_1(\varphi)$ is a linear map from $T_1(G)$ into $T_1(H)$. Therefore, in this way we get a functor from the category of Lie groups into the category of finite-dimensional linear spaces over $\mathbb{R}$.

We want to show that these objects have additional structure which carries additional information about Lie groups.

For any $g \in G$, $\operatorname{Int}(g) : G \longrightarrow G$ given by $\operatorname{Int}(g)(h) = ghg^{-1}$ is an automorphism of G. Therefore, $\operatorname{Ad}(g) = T_1(\operatorname{Int}(g))$ is a linear automorphism of $T_1(G)$.

2.2.1. LEMMA. *The map* $\operatorname{Ad} : G \longrightarrow GL(T_1(G))$ *is a Lie group homomorphism.*

PROOF. We have

$$\begin{aligned}\operatorname{Ad}(gg') &= T_1(\operatorname{Int}(gg')) = T_1(\operatorname{Int}(g) \circ \operatorname{Int}(g')) \\ &= T_1(\operatorname{Int}(g)) \circ T_1(\operatorname{Int}(g')) = \operatorname{Ad}(g) \circ \operatorname{Ad}(g')\end{aligned}$$

for all $g, g' \in G$. Therefore, Ad is a group homomorphism. Clearly, it is also differentiable. □

Let $\varphi : G \longrightarrow H$ be a Lie group morphism. Then we have

$$\varphi(\operatorname{Int}(g)(g')) = \varphi(gg'g^{-1}) = \varphi(g)\varphi(g')\varphi(g)^{-1} = \operatorname{Int}(\varphi(g))(\varphi(g'))$$

for $g, g' \in G$. By differentiation at $1 \in G$ we get

$$T_1(\varphi) \circ T_1(\mathrm{Int}(g)) = T_1(\mathrm{Int}(\varphi(g)) \circ T_1(\varphi)$$

i.e.,

$$T_1(\varphi) \circ \mathrm{Ad}_G(g) = \mathrm{Ad}_H(\varphi(g)) \circ T_1(\varphi)$$

for any $g \in G$. Hence $T_1(\varphi)$ intertwines the group actions.

By differentiating the Lie group homomorphism $\mathrm{Ad} : G \longrightarrow \mathrm{GL}(T_1(G))$ we get a linear map $T_1(\mathrm{Ad})) : T_1(G) \longrightarrow \mathcal{L}(T_1(G))$. This map defines a bilinear map

$$(\xi, \eta) \longmapsto [\xi, \eta] = (T_1(\mathrm{Ad})(\xi))(\eta)$$

from $T_1(G) \times T_1(G) \longrightarrow T_1(G)$. We can view it as a bilinear operation on $T_1(G)$.

2.2.2. Lemma. *Let $\varphi : G \longrightarrow H$ be a Lie group morphism. Then*

$$T_1(\varphi)([\xi, \eta]) = [T_1(\varphi)(\xi), T_1(\varphi)(\eta)]$$

for any $\xi, \eta \in T_1(G)$.

Proof. Let $\varphi : G \longrightarrow H$ be a Lie group morphism. By the above formula, for any $\eta \in T_1(G)$ we have

$$T_1(\varphi)(\mathrm{Ad}_G(g)\eta) = \mathrm{Ad}_H(\varphi(g))(T_1(\varphi)(\eta))$$

for any $\eta \in T_1(G)$. By taking the differential of this map at $1 \in G$ and evaluating it on $\xi \in T_1(G)$, we get

$$T_1(\varphi)([\xi, \eta]) = [T_1(\varphi)(\xi), T_1(\varphi)(\eta)]$$

for any $\xi, \eta \in T_1(G)$. □

Therefore, $T_1(\varphi) : T_1(G) \longrightarrow T_1(H)$ is a morphism of the tangent spaces equipped with the above bilinear operation.

2.2.3. Lemma. *Let $\xi, \eta, \zeta \in T_1(G)$. Then*

$$[\xi, [\eta, \zeta]] = [[\xi, \eta], \zeta] + [\eta, [\xi, \zeta]].$$

Proof. If we apply 2.2.2 to the Lie group automorphism $\mathrm{Int}(g)$ we get

$$\mathrm{Ad}(g)([\eta, \zeta]) = [\mathrm{Ad}(g)(\eta), \mathrm{Ad}(g)(\zeta)].$$

for any $\eta, \zeta \in T_1(G)$. By taking the differential of these maps at $1 \in G$ and evaluating it at $\xi \in T_1(G)$, we get

$$\begin{aligned}[\xi, [\eta, \zeta]] &= T_1(\mathrm{Ad})(\xi)([\eta, \zeta]) \\ &= [T_1(\mathrm{Ad})(\xi)(\eta), \zeta] + [\eta, T_1(\mathrm{Ad})(\xi)(\zeta)] = [[\xi, \eta], \zeta] + [\eta, [\xi, \zeta]].\end{aligned}$$

□

We shall prove that $T_1(G)$ with this bilinear operation is a Lie algebra. We only have to establish that $[\cdot, \cdot]$ is anticommutative. Then Jacobi identity follows from 2.2.3.

We start first with a special case. Let $G = \mathrm{GL}(n, \mathbb{R})$. Then, we can identify $T_1(G)$ with the space $M_n(\mathbb{R})$. For small t, the line $t \longmapsto I + tT$ lies in $\mathrm{GL}(n, \mathbb{R})$. Moreover,

$$\mathrm{Int}(S)(I + tT) = S(I + tT)S^{-1} = I + tSTS^{-1}$$

and we have

$$\mathrm{Ad}(S)(T) = STS^{-1}$$

for any $T \in M_n(\mathbb{R})$ and $S \in \mathrm{GL}(n,\mathbb{R})$. Moreover, for small t we have

$$(I + tS)^{-1} = I - tS + t^2(\dots),$$

for any $S \in M_n(\mathbb{R})$, what yields to

$$\mathrm{Ad}(I + tS)(T) = (I + tS)T(I + tS)^{-1} = T + t(ST - TS) + t^2(\dots),$$

for small t. It follows that

$$T_1(\mathrm{Ad}(S))(T) = ST - TS = [S,T]$$

and the above bilinear operation is the natural commutator on $M_n(\mathbb{R})$. Therefore, $T_1(\mathrm{GL}(n,\mathbb{R}))$ is a Lie algebra.

Now we want to prove this for an arbitrary Lie group. This requires some preparation.

Consider first the multiplication map $m : G \times G \longrightarrow G$. By 1.8.1, its differential $T_{(1,1)}(m) : T_1(G) \times T_1(G) \longrightarrow T_1(G)$ at $(1,1)$ is equal to

$$T_{(1,1)}(m)(\xi,\eta) = \xi + \eta$$

for all $\xi, \eta \in T_1(G)$.

Since we have $m(g,\iota(g)) = 1$, it follows that

$$0 = T_{(1,1)}(m) \circ (I_{T_1(G)} \times T_1(\iota)) \circ T_1(\Delta) = I_{T_1(G)} + T_1(\iota),$$

where $\Delta : G \longrightarrow G \times G$ is the diagonal map. Hence, we have

$$T_1(\iota) = -I_{T_1(G)}.$$

Let M and N be two differentiable manifolds, $p \in M$ and $q \in N$. Let $X \in T_p(M)$ and $Y \in T_q(N)$. For $f \in C^\infty(M \times N)$ we denote by

$$f_X : n \longmapsto X(f(\cdot,n)) \text{ and } f^Y : m \longmapsto Y(f(m,\cdot))$$

are smooth functions in $C^\infty(N)$ and $C^\infty(M)$ respectively. In addition, $Y(f_X) = X(f^Y)$.

Let G be a Lie group and $\xi, \eta \in T_1(G)$. Put

$$(\xi * \eta)(f) = \xi((f \circ m)^\eta) = \eta((f \circ m)_\xi)$$

for any $f \in C^\infty(G)$. Then $\xi * \eta$ is a linear form on the real linear space $C^\infty(G)$. It is called the *convolution* of ξ and η.

2.2.4. Lemma. *For $\xi, \eta \in T_1(G)$ we have*

$$[\xi,\eta] = \xi * \eta - \eta * \xi.$$

Proof. Fix $\xi, \eta \in T_1(G)$. Let $f \in C^\infty(G)$. Then can consider the function $\omega : g \longmapsto (\mathrm{Ad}(g)\eta)(f)$. The differential of ω at 1 satisfies

$$d\omega_1(\xi) = (T_1(\mathrm{Ad})(\xi)\eta)(f) = [\xi,\eta](f).$$

On the other hand,

$$\omega(g) = (\mathrm{Ad}(g)\eta)(f) = (T_1(\mathrm{Int}(g))\eta)(f) = \eta(f \circ \mathrm{Int}(g))$$

for all $g \in G$. Therefore, if we put

$$F(g,h) = (f \circ \mathrm{Int}(g))(h) = f(ghg^{-1})$$

for all $g, h \in G$, it follows that

$$\omega(g) = \eta(f \circ \mathrm{Int}(g)) = \eta(F(g,\cdot)) = F^\eta(g)$$

and

$$d\omega_1(\xi) = \xi(F^\eta) = \eta(F_\xi).$$

On the other hand, if we define $\mu : G \times G \times G \longrightarrow G$ as $\mu(g, g', h) = ghg'$ for $g, g', h \in G$, we have

$$F(g,h) = (f \circ \mu)(g, g^{-1}, h) = (f \circ \mu)(g, \iota(g), h)$$

for $g, h \in G$. Therefore,

$$\begin{aligned} F_\xi(h) &= \xi((f \circ \mu)(\cdot, 1, h)) + \xi((f \circ \mu)(1, \iota(\cdot), h)) \\ &= \xi((f \circ m)(\cdot, h)) - \xi((f \circ m)(h, \cdot)) = (f \circ m)_\xi(h) - (f \circ m)^\xi(h), \end{aligned}$$

what finally leads to

$$d\omega_1(\xi) = \eta((f \circ m)_\xi) - \eta((f \circ m)^\xi) = (\xi * \eta)(f) - (\eta * \xi)(f).$$

□

In particular, the bilinear operation $(\xi, \eta) \longmapsto [\xi, \eta]$ on $T_1(G)$ is anticommutative.

This completes the proof that the tangent space $T_1(G)$ of a Lie group G equipped with the bilinear operation $(\xi, \eta) \longmapsto [\xi, \eta]$ is a Lie algebra. We call it te *Lie algebra of the Lie group* G and denote by $L(G)$.

Let $\varphi : G \longrightarrow H$ be a morphism of Lie groups. By 2.2.2, the linear map $T_1(\varphi) : T_1(G) \longrightarrow T_1(H)$ can be interpreted as a morphism of Lie algebra $L(G)$ into $L(H)$. Hence we shall denote it by $L(\varphi) : L(G) \longrightarrow L(H)$. In this way we get a functor from the category of Lie groups into the category of real Lie algebras.

Moreover, from the definition of the commutator we see that the following relation holds.

2.2.5. Lemma.

$$L(\mathrm{Ad}) = \mathrm{ad}\,.$$

Let G be a Lie group and $g \in G$. Then $\mathrm{Int}(g)$ is an automorphism of G. Therefore, $\mathrm{Ad}(g) = L(\mathrm{Int}(g))$ is an automorphism of $L(G)$. Therefore, the adjoint representation $\mathrm{Ad} : G \longrightarrow \mathrm{GL}(L(G))$ is a homomorphism of G into the group $\mathrm{Aut}(L(G))$ of automorphisms of $L(G)$.

There is another realization of the Lie algebra of G which is somethimes useful. It is a consequence of the following discussion.

The algebra $\mathrm{End}(C^\infty(G))$ is an associative algebra with identity. Therefore, with the commutator $[A, B] = A \circ B - B \circ A$ it is a real Lie algebra.

A *vector field* X on G is an element of $\mathrm{End}(C^\infty(G))$ which is also a *derivation* of $C^\infty(G)$, i.e., it satisfies

$$X(\varphi\psi) = \varphi X(\psi) + X(\varphi)\psi$$

for all $\varphi, \psi \in C^\infty(G)$.

We claim that the linear space $\mathcal{T}(G)$ of all vector fields on G is a Lie sublagebra of $\operatorname{End}(C^\infty(G))$. Let $X, Y \in \mathcal{T}(G)$. Then we have

$$\begin{aligned}
&[X,Y](\varphi\psi) = X(Y(\varphi\psi)) - Y(X(\varphi\psi)) \\
&= X(\varphi Y(\psi)) + X(Y(\varphi)\psi) - Y(\varphi X(\psi)) - Y(X(\varphi)\psi) = X(\varphi)Y(\psi) + \varphi X(Y(\psi)) \\
&\quad + X(Y(\varphi))\psi + Y(\varphi)X(\psi) - Y(\varphi)X(\psi) - \varphi Y(X(\psi)) - X(\varphi)Y(\psi) \\
&= \varphi X(Y(\psi)) - \varphi Y(X(\psi)) + X(Y(\varphi))\psi - Y(X(\varphi))\psi = \varphi[X,Y](\psi) - [X,Y](\varphi)\psi
\end{aligned}$$

for all $\varphi, \psi \in C^\infty(G)$. Therefore, $[X, Y]$ is a vector field on G. It follows that $\mathcal{T}(G)$ is a Lie subalgebra of $\operatorname{End}(C^\infty(G))$.

Let X be a vector field on G. Let $g \in G$. Then $f \longmapsto X(f)(g)$ is a tangent vector X_g in $T_g(G)$ which we call the value of X at g.

The vector field X is *left-invariant* if $X_{gh} = T_h(\gamma(g))X_h$ for any $g, h \in G$. This implies that for any $f \in C^\infty(G)$, we have

$$X(f)(gh) = X(f \circ \gamma(g))(h)$$

for all $g, h \in G$, i.e.,

$$X(f) \circ \gamma(g) = X(f \circ \gamma(g))$$

for all $g \in G$. It is clear that the last property of X is equivalent to the left-invariance.

Let X and Y be two left-invariant vector fields on G. Then

$$\begin{aligned}
&[X,Y](f \circ \gamma(g)) = X(Y(f \circ \gamma(g))) - Y(X(f \circ \gamma(g))) \\
&\quad = X(Y(f) \circ \gamma(g)) - Y(X(f) \circ \gamma(g)) = X(Y(f)) \circ \gamma(g) - Y(X(f)) \circ \gamma(g) \\
&\qquad\qquad = [X,Y](f) \circ \gamma(g)
\end{aligned}$$

for all $g \in G$, i.e., the vector field $[X, Y]$ is also left-invariant.

Therefore, left-invariant vector fields form a Lie subalgebra $\mathcal{L}(G)$ of $\mathcal{T}(G)$.

2.2.6. Lemma. *The map $X \longmapsto X_1$ is a linear isomorphism of $\mathcal{L}(G)$ onto $T_1(G)$.*

Proof. If X is left-invariant, $X_g = T_1(\gamma(g))X_1$ for any $g \in G$, i.e., the map $X \longmapsto X_1$ is injective.

On the other hand, for any $\xi \in T_1(G)$, the map $f \longmapsto \xi(f \circ \gamma(\cdot))$ is a left-invariant vector field on G. □

Let $\xi, \eta \in T_1(G)$. Then, by 2.2.6, there exist left invariant vector fields X and Y on G such that $X_1 = \xi$ and $Y_1 = \eta$.

2.2.7. Lemma. *We have*

$$[X,Y]_1 = [\xi, \eta].$$

Proof. To prove this, it is enough to establish that

$$(\xi * \eta)(f) = \xi(Y(f))$$

for any $f \in C^\infty(G)$. Since Y is left-invariant, we have

$$Y(f)(g) = Y_g(f) = \eta(f \circ \gamma(g)) = \eta((f \circ m)(g, \cdot)) = (f \circ m)^\eta(g)$$

for any $g \in G$. Therefore, we have

$$\xi(Y(f)) = \xi((f \circ m)^\eta) = (\xi * \eta)(f)$$

for any $f \in C^\infty(G)$. □

Therefore, we see that the linear isomorphism $X \longmapsto X_1$ of $\mathcal{L}(G)$ onto $T_1(G)$ also preserves the commutators, i.e., it is a Lie algebra isomorphism of $\mathcal{L}(G)$ with $L(G)$.

Let G be a Lie group and G^{opp} the opposite Lie group. Then the map $\iota : g \longmapsto g^{-1}$ is an isomorphism from G onto G^{opp}. As we remarked already, $L(\iota) = -1_{L(G)}$ and it defines an isomorphism of $L(G)$ onto $L(G^{opp})$. Therefore, we have $L(G^{opp}) = L(G)^{opp}$.

We say that a vector field X on G is *right-invariant* if

$$X(f \circ \delta(g)) = X(f) \circ \delta(g)$$

for all $f \in C^\infty(G)$ and $g \in G$.

Let $\gamma^o(g)$ be the left translation by $g \in G^{opp}$. Then

$$\gamma^o(g)(h) = g \circ h = hg = \delta(g^{-1})(h).$$

for any $h \in G$. Therefore, a right-invariant vector field on G is a left-invariant vector field on G^{opp}. This in turn implies that all right-invariant vector fields on G form a Lie algebra which we denote by $\mathcal{R}(G)$. Moreover, $X \longmapsto X_1$ is an isomorphism of $\mathcal{R}(G)$ onto $L(G)^{opp}$. Therefore, for two right-invariant vector fields X and Y on G such that $\xi = X_1$ and $\eta = Y_1$, we have

$$[\xi, \eta] = -[X, Y]_1.$$

This gives an interpretation of the commutator in $L(G)$ in terms of right-invariant vector fields.

Let H be a Lie subgroup of a Lie group G. Then the natural inclusion $i : H \longrightarrow G$ is a Lie group morphism. Therefore, the natural inclusion $L(i) : L(H) \longrightarrow L(G)$ is a Lie algebra morphism, i.e., we can view $L(H)$ as a Lie subalgebra of $L(G)$.

2.2.8. LEMMA. *Let H be a normal Lie subgroup of a Lie group G. Then $L(H)$ is an ideal in $L(G)$.*

PROOF. For any $g \in G$ we have $\mathrm{Int}(g)(H) = H$. Therefore, $\mathrm{Ad}(g)(L(H)) = L(H)$ for any $g \in G$. By differentiation, from 2.2.5 we conclude that $\mathrm{ad}(\xi)(L(H)) \subset L(H)$ for any $\xi \in L(G)$. □

2.2.9. LEMMA. *Let $\varphi : G \longrightarrow H$ be a morphism of Lie groups. Then $L(\ker \varphi) = \ker L(\varphi)$.*

PROOF. This is just a reformulation of 1.1.5.(ii). □

2.2.10. LEMMA. *Let G be a Lie group and H its normal Lie subgroup. Denote by $p : G \longrightarrow G/H$ the canonical projection. Then $L(p) : L(G) \longrightarrow L(G/H)$ induces an isomorphism of $L(G)/L(H)$ with $L(G/H)$.*

PROOF. By 2.2.8, $L(H)$ is an ideal in $L(G)$. Since the canonical projection p is a submersion, $L(p)$ is surjective. Moreover, by 2.2.9, we have $\ker L(p) = L(H)$. □

2.2.11. LEMMA. *Let $\varphi : G \longrightarrow H$ be a morphism of connected Lie groups. Then the following statements are equivalent:*

(i) *φ is a covering projection;*
(ii) *$L(\varphi) : L(G) \longrightarrow L(H)$ is an isomorphism of Lie algebras.*

PROOF. This follows immediately from 1.6.6. □

2.2.12. PROPOSITION. *Let $\varphi : G \longrightarrow H$ be a morphism of Lie groups. Let K be a Lie subgroup of H. Then, $\varphi^{-1}(K)$ is a Lie subgroup of G and*

$$L(\varphi^{-1}(K)) = L(\varphi)^{-1}(L(K)).$$

PROOF. Let H/K be the left coset space of H. Let $p : H \longrightarrow H/K$ be the quotient projection. Then, K is equal to the fiber over the identity coset in H/K. Hence, since p is a submersion, by 1.1.4.4, $L(K) = \ker T_1(p)$.

The group H acts differentiably on H/K. Therefore, the composition of this action with φ defines a differentiable action of G on H/K. The stabilizer at the K-coset of 1 is equal to $\varphi^{-1}(K)$. Therefore, by 1.1.4, $\varphi^{-1}(K)$ is a Lie subgroup of G and

$$L(\varphi^{-1}(K)) = \{\xi \in L(G) \mid T_1(p \circ \varphi)(\xi) = 0\} = \{\xi \in L(G) \mid T_1(\varphi)(\xi) \in L(K)\}.$$

□

Let G and H be two Lie groups. Then $G \times H$ is a Lie group.

2.2.13. LEMMA. *$L(G \times H) = L(G) \times L(H)$.*

Let Δ be the diagonal in $G \times G$. Then Δ is a Lie subgroup of $G \times G$. Clearly, the map $\alpha : g \longmapsto (g, g)$ is an isomorphism of G onto Δ. Let H and H' be two Lie subgroups of G. Then $H \times H'$ is a Lie subgroup of $G \times G$. Moreover, $\alpha^{-1}(H \times H') = H \cap H'$. Therefore, by 2.2.12, we have the following result.

2.2.14. LEMMA. *Let H and H' be two Lie subgroups of G. Then $H \cap H'$ is a Lie subgroup of G.*

2.2.15. LEMMA. *Let $\varphi : G \longrightarrow H$ and $\psi : G \longrightarrow H$ be two Lie group morphisms. Then*

$$K = \{g \in G \mid \varphi(g) = \psi(g)\}$$

is a Lie subgroup of G and

$$L(K) = \{\xi \in L(G) \mid L(\varphi)(\xi) = L(\psi)(\xi)\}.$$

PROOF. We consider the Lie group morphism $\Phi : G \longrightarrow H \times H$ given by $\Phi(g) = (\varphi(g), \psi(g))$ for all $g \in G$. Clearly, $L(\Phi) : L(G) \longrightarrow L(H) \times L(H)$ is given by $L(\Phi)(\xi) = (L(\varphi)(\xi), L(\psi)(\xi))$ for $\xi \in L(G)$. The Lie algebra of the diagonal Δ in $H \times H$ is the diagonal in $L(H) \times L(H)$. Therefore, by 2.2.12,

$$K = \Phi^{-1}(\Delta)$$

is a Lie subgroup of G and its Lie algebra is equal to

$$L(\Phi)^{-1}(L(\Delta)) = \{\xi \in L(G) \mid L(\varphi)(\xi) = L(\psi)(\xi)\}.$$

□

Let G and H be two Lie groups. In general, we cannot say anything about the map $\varphi \longmapsto L(\varphi)$ from $\mathrm{Hom}(G, H)$ into $\mathrm{Hom}(L(G), L(H))$.

2.2.16. PROPOSITION. *Let G and H be Lie groups. Assume that G is connected. Then the map $\varphi \longmapsto L(\varphi)$ from $\mathrm{Hom}(G, H)$ into $\mathrm{Hom}(L(G), L(H))$ is injective.*

PROOF. Let $\varphi : G \longrightarrow H$ and $\psi : G \longrightarrow H$ be two Lie group morphisms such that $L(\varphi) = L(\psi)$. Then, by 2.2.15, $K = \{g \in G \mid \varphi(g) = \psi(g)\}$ is a Lie subgroup of G. Moreover, the Lie algebra $L(K)$ of K is equal to $L(G)$. It follows that K contains a neighborhood of 1 in G. Since it is a subgroup of G, and G is connected, it must be equal to G by 1.5.1. Therefore, $\varphi = \psi$. □

Of course, even if G is connected, the map $\varphi \longmapsto L(\varphi)$ from $\mathrm{Hom}(G, H)$ into $\mathrm{Hom}(L(G), L(H))$ is not bijective in general. For example, if $G = \mathbb{R}/\mathbb{Z}$ and $H = \mathbb{R}$, the set $\mathrm{Hom}(G, H)$ consists of the trivial morphism only, while $\mathrm{Hom}(L(G), L(H))$ is the space of all linear endomorphisms of $\mathbb{R}$.

We are going to prove later that if G is in addition simply connected, the map $\varphi \longmapsto L(\varphi)$ from $\mathrm{Hom}(G, H)$ into $\mathrm{Hom}(L(G), L(H))$ is bijective.

2.2.17. LEMMA. *Let G be a connected Lie group. Then*

(i) *the center Z of G is a Lie subgroup;*
(ii) $Z = \ker \mathrm{Ad}$*;*
(iii) *$L(Z)$ is the center of $L(G)$.*

PROOF. Clearly (ii) implies (i).

Let $z \in Z$. Then $\mathrm{Int}(z) = id_G$ and $\mathrm{Ad}(z) = L(\mathrm{Int}(z)) = 1$. Assume that $\mathrm{Ad}(g) = 1$ for $g \in G$. Then $L(\mathrm{Int}(g)) = L(id_G)$, and by 2.2.16, we see that $\mathrm{Int}(g) = id_G$, i.e., $g \in Z$. This proves (ii).

By 2.2.9, we have $L(Z) = L(\ker \mathrm{Ad}) = \ker L(\mathrm{Ad}) = \ker \mathrm{ad}$. Clearly, $\ker \mathrm{ad}$ is the center of $L(G)$. □

2.2.18. LEMMA. *Let G be a connected Lie group. Then the following statements are equivalent:*

(i) *G is abelian;*
(ii) *$L(G)$ is abelian.*

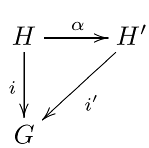
PROOF. (i)⇒(ii) If G is abelian, it is equal to its center. Therefore, by 2.2.17, $L(G)$ is equal to its center, i.e., it is abelian.

(ii)⇒(i) If $L(G)$ is abelian, by 2.2.17, the Lie algebra of the center Z of G is equal to $L(G)$. Therefore, Z contains a neighborhood of 1 in G. Hence Z is an open subgroup of G and, since G is connected, it is equal to G. □

2.3. From Lie algebras to Lie groups.

2.3.1. LEMMA. *Let G be a Lie group. Let $\mathfrak{h}$ be a Lie subalgebra of the Lie algebra $L(G)$ of G.*

(i) *There exists a connected Lie group H and an injective Lie group morphism $i : H \longrightarrow G$ such that $L(i) : L(H) \longrightarrow L(G)$ is an isomorphism of $L(H)$ onto $\mathfrak{h}$.*
(ii) *The pair (H, i) is unique up to an isomorphism, i.e., if (H', i') is another such pair, there exists a Lie group isomorphism $\alpha : H \longrightarrow H'$ such that the diagram*



commutes.

The proof of this lemma consists of several steps.

Let $T(G)$ be the tangent bundle of G. Let E vector subbundle of $T(G)$ such that the fiber E_g at $g \in G$ is equal to $T_1(\gamma(g))\mathfrak{h}$. Let $(\xi_1, \xi_2, \dots, \xi_m)$ be a basis of $\mathfrak{h}$. Denote by $X_1, X_2, \dots, X_m$ the left-invariant vector fields on G such that the value of X_i at 1 is equal to ξ_i for $1 \le i \le m$. Then the values of X_i, $1 \le i \le m$, at $g \in G$ span the fiber E_g. In particular, E is a trivial vector bundle on G.

Since, $\mathfrak{h}$ is a subalgebra, there exist $c_{ijk} \in \mathbb{R}$, $1 \le i, j, k \le m$, such that

$$[\xi_i, \xi_j] = \sum_{k=1}^{m} c_{ijk} \xi_k$$

for all $1 \le i, j \le m$. Therefore, we also have

$$[X_i, X_j] = \sum_{k=1}^{m} c_{ijk} X_k$$

for all $1 \le i, j \le m$.

A smooth vector field Y on G is a section of E if and only if $X_g \in E_g$ for all $g \in G$, i.e., if $X = \sum_{i=1}^{m} e_i X_i$ for some $e_i \in C^\infty(G)$. Let Z be another such vector field. Then we have $Z = \sum_{i=1}^{m} f_i X_i$ for some $f_i \in C^\infty(G)$.

Hence, we have

$$\begin{aligned}[Y, Z] = \sum_{i,j=1}^{m} [e_i X_i, f_j X_j] &= \sum_{i,j=1}^{m} (e_i X_i(f_j) X_j - f_j X_j(e_i) X_i + e_i f_j [X_i, X_j]) \\ &= \sum_{i,j=1}^{m} (e_i X_i(f_j) - f_i X_i(e_j)) X_j + \sum_{i,j,k=1}^{m} c_{ijk} e_i f_j X_k,\end{aligned}$$

i.e., $[X, Y]_g \in E_g$ for any $g \in G$. Therefore it follows that E is involutive.

By 1.3.2.1, E determines an integral foliation (L, i) of G which we call the *left foliation* attached to $\mathfrak{h}$.

Let H be the leaf of this foliation through $1 \in G$. We claim that H is a Lie group.

Let $g \in G$. Then $i_g = \gamma(g) \circ i : L \longrightarrow G$ is again an integral manifold. Therefore, by 1.3.2.1, i_g induces a diffeomorphism $j_g : L \longrightarrow L$. Hence, $j_g(H)$ is a leaf through $g \in G$. In particular, if $g \in H$, we see that $j_g(H) = H$. Therefore, the left multiplication by $g \in H$ induces a diffeomorphism of H onto H. Moreover, its inverse is $j_{g^{-1}} : H \longrightarrow H$. Hence, $j_{g^{-1}}(1) = g^{-1} \in H$. It follows that H is a subgroup of G.

In addition, the map $\mu : H \times H \longrightarrow G$ given by $\mu(g, h) = gh$ for $g, h \in H$, is differentiable and its image is equal to the leaf H. Since H is connected, it lies in the identity component of G. Hence, without any loss of generality we can assume that G is connected. Therefore, by 1.5.2, G is a separable manifold. By 1.3.3.4, it follows that H is a separable manifold. Hence, by 1.3.3.6, we conclude that the map $\mu : H \times H \longrightarrow H$ is differentiable. It follows that H is a Lie group. This completes the proof of (i).

If (H', i') is another such pair, it is an integral manifold for the left foliation attached to $\mathfrak{h}$. It follows that there exists $\alpha : H' \longrightarrow L$ which is a diffeomorphism onto an open submanifold of L. Since H' is connected and $i'(1) = 1$, $\alpha(H')$ must be an open subgroup of H. This in turn implies that $\alpha(H') = H$. Therefore, (ii) follows.

2.4. Additional properties of the Lie algebra functor. Let G and H be Lie groups. Assume in addition that G is connected. We already established in

2.2.16 that the map the functor L induces from $\mathrm{Hom}(G,H)$ into $\mathrm{Hom}(L(G),L(H))$ is injective.

First, let $\varphi : G \longrightarrow H$ be a Lie group morphism. Then we can consider its graph $\Gamma_\varphi = \{(g,\varphi(g)) \in G \times H \mid g \in G\}$ in $G \times H$. By 1.1.4.3, it is a Lie subgroup of $G \times H$. The natural morphism $\lambda : g \longmapsto (g,\varphi(g))$ is a Lie group isomorphism of G with Γ_φ. Its inverse is the restriction of the projection to the first factor.

Moreover, its Lie algebra $L(\Gamma_\varphi)$ is the image of $L(\lambda) : L(G) \longrightarrow L(G) \times L(H)$. Since $L(\lambda) : \xi \longmapsto (\xi, L(\varphi)(\xi))$, $\xi \in L(G)$, we see that $L(\Gamma_\varphi) = \{(\xi, L(\varphi)(\xi)) \in L(G) \times L(H) \mid \xi \in L(G)\}$, i.e., it is equal to the graph of the Lie algebra morphism $L(\varphi)$ in $L(G) \times L(H)$.

2.4.1. Proposition. *Let G be a simply connected, connected Lie group. Let H be another Lie group and $\Phi : L(G) \longrightarrow L(H)$ a Lie algebra morphism. Then there exists a Lie group morphism $\varphi : G \longrightarrow H$ such that $L(\varphi) = \Phi$.*

Proof. Let $L(G) \times L(H)$ be the product Lie algebra of $L(G)$ and $L(H)$. Then the graph $\Gamma_\Phi = \{(\xi, \Phi(\xi)) \in L(G) \times L(H) \mid \xi \in L(G)\}$ of Φ is a Lie subalgebra of $L(G) \times L(H)$. The map $\alpha : L(G) \longrightarrow L(G) \times L(H)$ given by $\alpha(\xi) = (\xi, \Phi(\xi))$ is a Lie algebra isomorphism from $L(G)$ into Γ_Φ. Its inverse is given by the canonical projection to the first factor in $L(G) \times L(H)$. On the other hand, Φ is the composition of α with the canonical projection to the second factor.

By 2.3.1, there exists a connected Lie group K and an injective Lie group morphism $i : K \longrightarrow G \times H$ such that $L(i) : L(K) \longrightarrow L(G) \times L(H)$ is an isomorphism of $L(K)$ onto Γ_Φ. Let $p : G \times H \longrightarrow G$ be the canonical projection to the first factor. Then it is a Lie group morphism, and $L(p) : L(G) \times L(H) \longrightarrow L(G)$ is also the canonical projection to the first factor. The composition $p \circ i : K \longrightarrow G$ is a Lie group morphism of connected Lie groups. Moreover, since the canonical projection to the first factor is an isomorphism of Γ_Ψ onto $L(G)$, $L(p \circ i) = L(p) \circ L(i)$ is an isomorphism of the Lie algebra $L(K)$ onto $L(G)$. By 2.2.11, $p \circ i$ is a covering projection. Since G is simply connected, $p \circ i$ is an isomorphism of Lie groups. Therefore, its inverse $\beta : G \longrightarrow K$ is a Lie group morphism. Clearly, $L(\beta)$ is the composition of α with the isomorphism $L(i)^{-1}$.

Let $q : G \times H \longrightarrow H$ be the canonical projection to the second factor. Then, $q \circ i \circ \beta : G \longrightarrow H$ is a Lie group morphism. Its differential is equal to

$$L(q \circ i \circ \beta) = L(q) \circ L(i) \circ L(\beta) = L(q) \circ \alpha = \Phi. \quad \square$$

This has the following obvious consequence.

2.4.2. Corollary. *Let G be a simply connected, connected Lie group. Let H be another Lie group. Then, Then the map induced by the functor L from* $\mathrm{Hom}(G,H)$ *into* $\mathrm{Hom}(L(G),L(H))$ *is bijective.*

In other words, the functor L from the category $\mathcal{S}imply\mathcal{C}onn\mathcal{L}ie$ of simply connected, connected Lie groups into the category of finite-dimensional real Lie algebras $\mathcal{L}ie\mathcal{A}lg$ is fully faithful.

On the other hand, Ado's theorem has the following consequence.

2.4.3. Theorem. *Let $\mathfrak{g}$ be a finite-dimensional real Lie algebra. Then there exists a simply connected, connected Lie group G such that $L(G)$ is isomorphic to $\mathfrak{g}$.*

PROOF. By 2.1.3, there exists a finite-dimensional real linear space V such that $\mathfrak{g}$ is isomorphic to a Lie subalgebra of $\mathfrak{gl}(V)$. Since $\mathfrak{gl}(V)$ is the Lie algebra of $\mathrm{GL}(V)$, by 2.3.1 we conclude that there exists a connected Lie group with the Lie algebra isomorphic to $\mathfrak{g}$. Therefore, taking its universal covering Lie group for G completes the proof. □

This implies that the Lie algebra functor L from the category $\mathcal{S}imply\mathcal{C}onn\mathcal{L}ie$ into $\mathcal{L}ie\mathcal{A}lg$ is also essentially onto. Therefore, we have the following result.

2.4.4. THEOREM. *The Lie algebra functor L is an equivalence of the category $\mathcal{S}imply\mathcal{C}onn\mathcal{L}ie$ of simply connected, connected Lie groups with the category $\mathcal{L}ie\mathcal{A}lg$ of finite-dimensional real Lie algebras.*

2.5. Discrete subgroups of $\mathbb{R}^n$. Let V be an n-dimensional linear space considered as an additive Lie group. We want to describe all discrete subgroups in V.

Let D be a discrete subgroup in V. The elements of D span a linear subspace W of V. We say that $\dim W$ is the *rank* of D.

2.5.1. THEOREM. *Let D be a discrete subgroup of V of rank r. Then there exists a linearly independent set of vectors $a_1, a_2, \dots, a_r$ in V such that $\mathbb{Z}^r \ni (n_1, n_2, \dots, n_r) \longmapsto n_1a_1 + n_2a_2 + \dots + n_ra_r$ is an isomorphism of $\mathbb{Z}^r$ onto D.*

We first observe that without any loss of generality we can assume that $r = n$.

We start the proof with the following weaker result. Since D has rank n, there exists a linearly independent set $b_1, b_2, \dots, b_n$ contained in D.

2.5.2. LEMMA. *There exists a positive integer d such that D is contained in the discrete subgroup D' of V generated by $\frac{1}{d}b_1, \frac{1}{d}b_2, \dots, \frac{1}{d}b_n$.*

PROOF. Let

$$\Omega = \{v \in V \mid v = \sum_{i=1}^{n} \omega_i b_i \text{ with } 0 \le \omega_i \le 1 \text{ for } 1 \le i \le n\}.$$

Then Ω is a compact subset of V and $D \cap \Omega$ is a finite set. Clearly, $D \cap \Omega$ contains $b_1, b_2, \dots, b_n$.

Let $v \in D$. Then $v = \sum_{i=1}^n \alpha_i b_i$. Let $u = \sum_{i=1}^n [\alpha_i] b_i \in D$. It follows that $v - u = \sum_{i=1}^n (\alpha_i - [\alpha_i]) b_i \in D \cap \Omega$. Therefore, D is generated by the elements of $D \cap \Omega$.

Let $v \in D \cap \Omega$. Applying the above argument to mv, $m \in \mathbb{N}$, we see that

$$\sum_{i=1}^{n} (m\alpha_i - [m\alpha_i]) b_i \in D \cap \Omega.$$

Since the set $D \cap \Omega$ is finite, there exist $m, m' \in \mathbb{N}$, such that $m \neq m'$ and

$$m\alpha_i - [m\alpha_i] = m'\alpha_i - [m'\alpha_i]$$

for all $1 \le i \le n$. Therefore,

$$(m - m')\alpha_i = [m\alpha_i] - [m'\alpha_i] \in \mathbb{Z},$$

for all $1 \le i \le n$, i.e., α_i are rational numbers.

It follows that the coordinates of all vectors in $D \cap \Omega$ with respect to $b_1, b_2, \dots, b_n$ are rational. Since $D \cap \Omega$ is finite, the coordinates of these points all lie in $\frac{1}{d}\mathbb{Z}$ for sufficiently large $d \in \mathbb{N}$.

This implies that D is contained in the subgroup generated by $\frac{1}{d}b_1, \frac{1}{d}b_2, \dots, \frac{1}{d}b_n$. □

Fix a linearly independent set $b_1, b_2, \dots, b_n$ of vectors in D. Let $d \in \mathbb{N}$ be an integer which satisfies the conditions of the preceding lemma. Let $c_i = \frac{1}{d}b_i$, $1 \le i \le n$. Then, an element $v \in D$ can be represented uniquely as $v = \sum_{i=1}^n m_i c_i$ where $m_i \in \mathbb{Z}$. It follows that for any linearly independent set $v_1, v_2, \dots, v_n$ contained in D we have $v_i = \sum_{j=1}^n m_{ij} c_j$ where $m_{ij} \in \mathbb{Z}$ for all $1 \le i, j \le n$. Define the function

$$\Delta(v_1, v_2, \dots, v_n) = \begin{vmatrix} m_{11} & m_{12} & \dots & m_{1n} \\ m_{21} & m_{22} & \dots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \dots & m_{nn} \end{vmatrix}$$

for any such linearly independent n-tuple $v_1, v_2, \dots, v_n$. Clearly $\Delta(v_1, v_2, \dots, v_n) \in \mathbb{Z}$. Moreover, since the n-tuple is linearly independent, $\Delta(v_1, v_2, \dots, v_n) \neq 0$. Therefore, there exists an n-tuple $d_1, d_2, \dots, d_n$ such that the absolute value of $\Delta(d_1, d_2, \dots, d_n)$ is minimal.

2.5.3. Lemma. *The map* $(m_1, m_2, \dots, m_n) \longmapsto m_1 d_1 + m_2 d_2 + \dots + m_n d_n$ *is an isomorphism of* $\mathbb{Z}^n$ *onto* D.

Proof. Denote by D'' the discrete subgroup generated by $d_1, d_2, \dots, d_n$. Clearly, $D'' \subset D$.

Let $v \in D$. Then $v = \sum_{i=1}^n \alpha_i d_i$ where $\alpha_i \in \mathbb{R}$, $1 \le i \le n$. In addition, we have $u = \sum_{i=1}^n [\alpha_i] d_i \in D''$. Therefore, $w = v - u = \sum_{i=1}^n (\alpha_i - [\alpha_i]) d_i \in D$. By the construction $w = \sum_{i=1}^n e_i d_i$ with $0 \le e_i < 1$ for all $1 \le i \le n$.

Assume that $w \neq 0$. Then the set $w, d_1, d_2, \dots, d_n$ is linearly dependent. After relabeling, we can assume that $e_1 > 0$. This implies that $w, d_2, \dots, d_n$ is a linearly independent set of vectors in D. Clearly,

$$w = \sum_{i=1}^n e_i d_i = \sum_{i,j=1}^n e_i m_{ij} b_j.$$

Therefore,

$$\Delta(w, d_2, \dots, d_n) = \begin{vmatrix} \sum_{i=1}^n e_i m_{i1} & \sum_{i=1}^n e_i m_{i2} & \dots & \sum_{i=1}^n e_i m_{in} \\ m_{21} & m_{22} & \dots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \dots & m_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} e_1 m_{11} & e_1 m_{i2} & \dots & e_1 m_{in} \\ m_{21} & m_{22} & \dots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \dots & m_{nn} \end{vmatrix} = e_1 \Delta(d_1, d_2, \dots, d_n).$$

Since $0 < e_1 < 1$, we have a contradiction with the minimality of $\Delta(d_1, d_2, \dots, d_n)$.

Hence, we must have $w = 0$, and $v = u \in D''$. This implies that $D = D''$. □

This completes the proof of 2.5.1.

2.6. Classification of connected abelian Lie groups. Let G be a connected abelian Lie group. Then, by 2.2.18, the Lie algebra $L(G)$ of G is abelian. Therefore, it is isomorphic to $\mathbb{R}^n$ with the trivial commutator for $n = \dim G$.

This Lie algebra is the Lie algebra of the additive Lie group $\mathbb{R}^n$. Since $\mathbb{R}^n$ is simply connected, by 2.4.1, there exists a Lie group morphism $\varphi : \mathbb{R}^n \longrightarrow G$ such that $L(\varphi)$ is a Lie algebra isomorphism of $\mathbb{R}^n$ onto $L(G)$. By 2.2.11, $\varphi : \mathbb{R}^n \longrightarrow G$ is a covering projection. This immediately implies the following result.

2.6.1. Proposition. *Let G be a simply connected, connected abelian Lie group. Then G is isomorphic to $\mathbb{R}^n$ for $n = \dim G$.*

If G is not simply connected, the kernel of φ is a discrete subgroup D of $\mathbb{R}^n$ and $G = \mathbb{R}^n/D$.

Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Then $\mathbb{T}$ is a one-dimensional connected compact abelian Lie group. The product $\mathbb{T}^r$ of r copies of $\mathbb{T}$ is an r-dimensional connected compact abelian Lie group which we call the *r-dimensional torus*.

2.6.2. Theorem. *Let G be an n-dimensional connected abelian Lie group. Then there exists $0 \le r \le n$ such that G is isomorphic to $\mathbb{T}^r \times \mathbb{R}^{n-r}$.*

Proof. This follows from 2.5.1 and the above discussion. □

2.6.3. Corollary. *Let G be a one-dimensional connected Lie group. Then G is isomorphic to either $\mathbb{R}$ or $\mathbb{T}$.*

Proof. Let $L(G)$ be the Lie algebra of G. Then $L(G)$ is a one-dimensional Lie algebra. Therefore, it must be abelian. By 2.2.18, G is an abelian Lie group. □

2.7. Induced structure on subgroups. Let G be a Lie group and H a subgroup of G. Let $\mathfrak{h}$ be the set of all $\xi \in L(G)$ such that there exist

(i) an open interval $I \subset \mathbb{R}$ containing 0;
(ii) a smooth curve $\Gamma : I \longrightarrow G$ such that $\Gamma(0) = 1$ and $\Gamma(I) \subset H$;
(iii) $T_0(\Gamma)(1) = \xi$.

2.7.1. Lemma. *Let H be a subgroup of a Lie group G. Then*

(i) *The subset $\mathfrak{h}$ of $L(G)$ is a Lie subalgebra.*
(ii) $\mathrm{Ad}(h)(\mathfrak{h}) = \mathfrak{h}$ *for each $h \in H$.*

Proof. First, if we put $\Gamma(t) = 1$ for all $t \in \mathbb{R}$, we see that $\Gamma(t) \in H$ for all $t \in \mathbb{R}$ and $T_0(\Gamma)(1) = 0$. Hence, $0 \in \mathfrak{h}$.

Let $\Gamma_1 : I_1 \longrightarrow G$ and $\Gamma_2 : I_2 \longrightarrow G$ be two curves in G such that $\Gamma_1(0) = \Gamma_2(0) = 1$, $\Gamma_1(I_1) \subset H$, $\Gamma_2(I_2) \subset H$ and $\xi_1 = T_0(\Gamma_1)(1)$, $\xi_2 = T_0(\Gamma_2)(1)$. Put $I = I_1 \cap I_2$. Then $\Gamma : I \longrightarrow G$ given by $\Gamma(t) = \Gamma_1(t) \cdot \Gamma_2(t)$ for $t \in I$, is a smooth curve in G. Moreover, $\Gamma(I) \subset H$ and $\Gamma(0) = 1$. Finally,

$$\begin{aligned} T_0(\Gamma)(1) = T_0(m \circ (\Gamma_1 \times \Gamma_2))(1) &= T_{(1,1)}(m)(T_0(\Gamma_1)(1), T_0(\Gamma_2)(1)) \\ &= T_{(1,1)}(m)(\xi_1, \xi_2) = \xi_1 + \xi_2. \end{aligned}$$

Hence, $\mathfrak{h}$ is closed under addition.

Let $\lambda \in \mathbb{R}^*$. Then $\Gamma_\lambda(t) = \Gamma_1(\lambda t)$ for $t \in I_\lambda = \frac{1}{\lambda} I_1$ is a smooth curve in G. Clearly, $\Gamma_\lambda(0) = \Gamma_1(0) = 1$ and $\Gamma_\lambda(I_\lambda) = \Gamma_1(I_1) \subset H$. Also, we have

$$T_0(\Gamma_\lambda)(1) = T_0(\Gamma_1)(\lambda) = \lambda T_0(\Gamma_1)(1) = \lambda \xi_1.$$

Therefore, $\lambda \xi_1 \in \mathfrak{h}$. It follows that $\mathfrak{h}$ is a linear subspace of $L(G)$.

If $h \in H$, $\Gamma_h : I \longrightarrow G$ defined by $\Gamma_h(t) = \operatorname{Int}(h)(\Gamma_1(t))$ is a smooth curve in G. Clearly, $\Gamma_h(0) = 1$ and $\Gamma_h(I_1) = \operatorname{Int}(h)(\Gamma_1(I_1)) \subset \operatorname{Int}(h)(H) = H$. Moreover, we have

$$T_0(\Gamma_h)(1) = T_1(\operatorname{Int}(h))(T_0(\Gamma_1)(1)) = L(\operatorname{Int}(h))(\xi_1) = \operatorname{Ad}(h)(\xi_1).$$

Therefore, $\operatorname{Ad}(h)(\mathfrak{h}) \subset \mathfrak{h}$. This proves (ii).

Finally, by (ii), for any $t \in I$, we have $\operatorname{Ad}(\Gamma_1(t))(\xi_2) \in \mathfrak{h}$. Therefore, $t \longmapsto \operatorname{Ad}(\Gamma_1(t))(\xi_2)$ is a smooth curve in $\mathfrak{h}$, and its tangent vector at 0 is also in $\mathfrak{h}$. This tangent vector is equal to

$$\begin{aligned}(T_0(\operatorname{Ad}\circ\Gamma_1)(1))(\xi_2) = ((T_1(\operatorname{Ad}) \circ T_0(\Gamma_1))(1))(\xi_2) &= (L(\operatorname{Ad})(\xi_1))(\xi_2) \\ &= \operatorname{ad}(\xi_1)(\xi_2) = [\xi_1, \xi_2].\end{aligned}$$

Therefore, $\mathfrak{h}$ is a Lie subalgebra of $L(G)$. □

We say that $\mathfrak{h}$ is the Lie subalgebra *tangent* to the subgroup H.

2.7.2. THEOREM. *Let G be a Lie group and H its subgroup. Then:*

(i) *On the set H there exists a unique structure of a differentiable manifold such that for any differentiable manifold M and map $f : M \longrightarrow H$, f is a differentiable map from M into H if and only if it is a differentiable map from M into G.*
(ii) *With this differentiable structure on H:*
  (a) *H is a Lie group;*
  (b) *the canonical injection $i : H \longrightarrow G$ is a morphism of Lie groups;*
  (c) *$L(i)$ is an isomorphism of $L(H)$ onto the Lie subalgebra $\mathfrak{h}$ tangent to H.*

We say that this Lie group structure on H is *induced* by the Lie group structure of G.

PROOF. Let (L, i) be the left foliation attached to $\mathfrak{h}$ and E the corresponding involutive vector subbundle of the tangent bundle $T(G)$.

Let M be a differentiable manifold and $f : M \longrightarrow G$ a differentiable map such that $f(M) \subset H$. Let $m \in M$ and $\xi \in T_m(M)$. Then there exists and open interval $I \subset \mathbb{R}$, $0 \in I$, and a smooth curve $\Gamma : I \longrightarrow M$ such that $\Gamma(0) = m$ and $T_0(\Gamma)(1) = \xi$. Then $f \circ \Gamma : I \longrightarrow G$ is a smooth curve in G, $(f \circ \Gamma)(0) = f(m)$ and $T_0(f \circ \Gamma)(1) = T_m(f)\xi$. It follows that $\Gamma_m = \gamma(f(m)^{-1}) \circ f \circ \Gamma : I \longrightarrow G$ is a smooth curve in G such that $\Gamma_m(0) = \gamma(f(m)^{-1})(f(m)) = 1$ and $T_0(\Gamma_m)(1) = T_{f(m)}(\gamma(f(m)^{-1})T_m(f)\xi$. Since $f(m) \subset H$ and H is a subgroup of G, it follows that $\Gamma_m(I) \subset H$. Hence, $T_0(\Gamma_m)(1) \in \mathfrak{h}$, i.e., we have $T_m(f)\xi \in T_1(\gamma(f(m)))\mathfrak{h} = E_{f(m)}$. Therefore, we see that

$$T_m(f)(T_m(M)) \subset E_{f(m)}, \text{ for any } m \in M.$$

Assume that $m_0 \in M$ is such that $f(m_0) = 1$. Let $c = (U, \varphi, n)$ be a chart centered at 1 such that $\varphi(U) = V \times W$ where V and W are connected open subsets in $\mathbb{R}^{n-l}$ and $\mathbb{R}^l$ respectively, such that $\varphi^{-1}(\{v\} \times W)$ are integral manifolds for E. Let O be an open connected neighborhood of m_0 such that $f(O) \subset U$. Denote by p the projection to the first factor in $\mathbb{R}^{n-l} \times \mathbb{R}^l$. Then, by the first part of the proof, we have

$$\begin{aligned}\operatorname{im} T_m(p\circ\varphi\circ f) &= (T_{\varphi(f(m))}(p)\circ T_{f(m)}(\varphi)\circ T_m(f))(T_m(M))\\ &\subset (T_{\varphi(f(m))}(p)\circ T_{f(m)}(\varphi))(E_{f(m)}) = (T_{\varphi(f(m))}(p)(\{0\}\times\mathbb{R}^l) = \{0\}.\end{aligned}$$

Hence, we have $T_m(p\circ\varphi\circ f) = 0$ for any $m \in O$. Since O is connected, $p\circ\varphi\circ f$ is constant on O. This in turn implies that $f(O) \subset \varphi^{-1}(\{0\}\times W)$. Therefore, $f(O)$ is contained in the leaf H_0 of L through $1 \in G$. Moreover, $f : O \longrightarrow H_0$ is a differentiable map.

As we proved in the proof of 2.3.1, H_0 is a Lie group, the canonical inclusion $j : H_0 \longrightarrow G$ is a morphism of Lie groups and $L(j) : L(H_0) \longrightarrow L(G)$ is an isomorphism of $L(H_0)$ onto $\mathfrak{h}$. Let $(\xi_1, \xi_2, \ldots, \xi_l)$ be a basis of $\mathfrak{h}$. Denote by $\Gamma_1, \Gamma_2, \ldots, \Gamma_l : I \longrightarrow G$ the corresponding smooth curves such that $\Gamma_i(I) \subset H$, $\Gamma_i(0) = 1$ and $T_0(\Gamma_i)(1) = \xi_i$ for $1 \le i \le l$. We define

$$F(t_1, t_2, \ldots, t_l) = \Gamma_1(t_1)\cdot\Gamma_2(t_2)\ldots\Gamma_l(t_l)$$

for $(t_1, t_2, \ldots, t_l) \in I^l$. Then $F : I^l \longrightarrow G$ is differentiable and $F(I^l) \subset H$. Since $F(0) = 1$, by the preceding part of the proof, there exists a neighborhood O of 0 in I^l such that $F|_O : O \longrightarrow H_0$ is a differentiable map. Clearly, if we denote by $e_1, e_2, \ldots, e_l$ the canonical basis of $\mathbb{R}^l$, we have $T_0(F)(e_i) = \xi_i$ for all $1 \le i \le l$. Therefore, F is a local diffeomorphism at 0. In particular, it is an open map. It follows that $H \cap H_0$ contains a neighborhood of 1 in H_0. Since H_0 is a connected Lie group, by 1.5.1, H_0 is contained in H.

Let $g \in G$. Then, as we proved in the proof of 2.3.1, $\gamma(g) : L \longrightarrow L$ is a differentiable map which permutes the leaves of L. If $h \in H$, we see that $\gamma(h)(H_0)$ is a leaf of L through h. Since $H_0 \subset H$, it follows that $\gamma(h)(H_0) \subset H$. Therefore, H is a union of leaves of L. We consider H to be equipped with the corresponding differentiable structure (as an open submanifold of L).

Let $f : M \longrightarrow H$ a map. If $f : M \longrightarrow H$ is differentiable, $f : M \longrightarrow G$ is also differentiable.

Conversely, if $f : M \longrightarrow G$ is differentiable, $\gamma(h)\circ f : M \longrightarrow G$ is differentiable for any $h \in H$. Fix $m \in M$. Then $f_m = \gamma(f(m)^{-1})\circ f : M \longrightarrow G$ is a differentiable map and $f_m(m) = \gamma(f(m)^{-1})(f(m)) = 1$. Therefore, by the above argument, there exists a neighborhood O of m such that f_m is a differentiable map from O into H_0. This implies that $f = \gamma(f(m))\circ f_m : M \longrightarrow H$ is differentiable at m. Therefore $f : M \longrightarrow H$ is differentiable. Hence, the differentiable structure on H satisfies the universal property in (i).

Assume that H is equipped with the differentiable structure satisfying the universal property from (i). Then, the identity map on H is differentiable. This in turn implies that the inclusion map $i : H \longrightarrow G$ is also differentiable, i.e., the assertion (b) in (ii) holds.

Now we prove the uniqueness part of (i). Assume that there exists another differentiable structure on H with the same universal property. Denote by $\mathbf{H}$ the corresponding manifold. Then $\mathbf{H} \longrightarrow G$ is differentiable, hence the identity map $\mathbf{H} \longrightarrow H$ is differentiable. Reversing the roles, we see that the identity map $H \longrightarrow \mathbf{H}$ is also differentiable. Hence, the differentiable structures on H and $\mathbf{H}$ are identical. This completes the proof of (i).

The multiplication map $m : H \times H \longrightarrow G$ is differentiable. Moreover, its image is in H. Therefore, by (i), $m : H \times H \longrightarrow H$ is differentiable and H is a Lie group.

Clearly, H_0 is the identity component of H. Therefore, we have $L(H) = L(H_0)$. From 2.3.1 we conclude that $L(i) : L(H) \longrightarrow L(G)$ is injective and its image is equal to $\mathfrak{h}$. Therefore, the Lie group structure on H satisfies (ii). □

The induced structure is the obvious one in the case of Lie subgroups.

2.7.3. LEMMA. *Let H be a Lie subgroup of G. Then the induced structure on H is equal to its natural differentiable structure.*

PROOF. It is clear that the differentiable structure of a submanifold has the universal property of the induced structure. By the uniqueness, they have to be equal. □

2.7.4. PROPOSITION. *Let G be a Lie group and H a subgroup of G. On the subgroup H there exists at most one structure of a Lie group with countably many components such that the canonical injection is a morphism of Lie groups.*

If such structure of Lie group with countably many components exists on H, it is equal to the induced structure.

PROOF. Assume that H has a structure of a Lie group with countably many components such that the canonical inclusion $i : H \longrightarrow G$ is a Lie group morphism. Denote $\mathbf{H}$ the Lie group H with induced structure on it. Then, by 2.7.2.(i), the identity map from $H \longrightarrow \mathbf{H}$ is a morphism of Lie groups. Since H has countably many components, by 1.5.7, this morphism must be an isomorphism. □

The next result is just a special case of the above result.

2.7.5. COROLLARY. *Let G be a Lie group and H a subgroup of G. There exists at most one structure of connected Lie group on H such that the canonical injection is a morphism of Lie groups. If such structure exists, it is equal to the induced structure on H.*

Let G be a Lie group. An *integral subgroup* of G is a subgroup H with a structure of connected Lie group such that the canonical inclusion is a Lie group morphism. This structure must be equal to the induced structure.

Let G be a Lie group and H an integral subgroup of G. We identify $L(i)L(H)$ with its image in $L(G)$ under $L(i)$. Then $L(H)$ is the Lie algebra tangent to H.

2.7.6. THEOREM. *Let G be a Lie group. The map $H \longmapsto L(H)$ is a bijection from the set of all integral subgroups of G onto the set of all Lie subalgebras of $L(G)$.*

This bijection is order preserving, i.e., $H_1 \subset H_2$ if and only if $L(H_1) \subset L(H_2)$.

PROOF. Let H and H' be two integral subgroups such that $L(H) = L(H')$. Then, by 2.3.1.(ii), we see that $H = H'$. Therefore, the map from integral subgroups of G into Lie subalgebras of $L(G)$ is injective.

On the other hand, 2.3.1.(i), implies that the this map is also surjective.

It remains to prove that this bijection preserves the inclusions. Clearly, if H_1 and H_2 are two integral subgroups such that $H_1 \subset H_2$, their tangent Lie algebras satisfy $L(H_1) \subset L(H_2)$.

On the other hand, assume that H_1 and H_2 are two integral subgroups of G such that $L(H_1) \subset L(H_2)$. Then $L(H_1)$ is a Lie subalgebra of $L(H_2)$. Therefore, by the first part of the proof, there exists an integral subgroup H' of H_2 such that

$L(H') = L(H_1)$. Clearly, H' is an integral subgroup of G, and by the first part of the proof $H' = H_1$. □

2.7.7. Lemma. *Let G be a Lie group and H_1 and H_2 two integral subgroups of G. For any $g \in G$ the following assertions are equivalent:*

(i) $gH_1g^{-1} = H_2$;
(ii) $\operatorname{Ad}(g)(L(H_1)) = L(H_2)$.

Proof. Clearly, $\operatorname{Int}(g)$ is a Lie group automorphism of G. Therefore, it induces a bijection on the set of all integral subgroups of G. Since $L(\operatorname{Int}(g)) = \operatorname{Ad}(g)$, this bijection corresponds to the bijection induced by $\operatorname{Ad}(g)$ on the set of all Lie subalgebras of $L(G)$. □

2.7.8. Lemma. *Let G be a Lie group and H an integral subgroup of G. Then the following conditions are equivalent:*

(i) *H is a normal subgroup of G;*
(ii) *$L(H)$ is an ideal in $L(G)$ invariant under $\operatorname{Ad}(G)$.*

Proof. From 2.7.7 we immediately see that H is normal if and only if $L(H)$ is invariant under $\operatorname{Ad}(G)$. By differentiation, this implies that $\operatorname{ad}(\xi)(L(H)) \subset L(H)$ for any $\xi \in L(G)$. Hence $L(H)$ is an ideal in $L(G)$. □

2.8. Lie subgroups of $\mathbb{R}^n$. Let V be an n-dimensional linear space considered as an additive Lie group. We want to describe all Lie subgroups in V.

We start with a technical lemma.

2.8.1. Lemma. *Let G be a Lie group, H a Lie subgroup and N a normal Lie subgroup of G contained in H. Then H/N is a Lie subgroup of G/N.*

Proof. Clearly, the natural map $j : H/N \longrightarrow G/N$ is injective. Therefore, by 1.1.5, it must be an immersion. By definition of the quotient topology, it is also a homeomorphism onto its image. Hence, by 1.1.4.2, the image of j is a Lie subgroup and j is a diffeomorphism of H/N onto $j(H/N)$. □

2.8.2. Theorem. *Let H be a Lie subgroup of V. Then there exists a linearly independent set $a_1, a_2, \ldots, a_r$ in V such that*

$$\mathbb{R}^k \times \mathbb{Z}^{r-k} \ni (\alpha_1, \ldots, \alpha_k, m_{k+1}, \ldots, m_r) \longmapsto \alpha_1 a_1 + \cdots + \alpha_k a_k + m_{k+1} a_k + \cdots + m_r a_r$$

is an isomorphism of $\mathbb{R}^k \times \mathbb{Z}^{r-k}$ onto H.

Proof. Let $L(H)$ be the Lie subalgebra of $L(V) = V$ corresponding to H. Then $L(H)$ is a subspace of V, and therefore a connected Lie subgroup of V. Since its Lie algebra is identified with $L(H)$, by 2.7.6, we conclude that the identity component H_0 of H is equal to this subspace. Let $k = \dim H_0$. We can pick a basis $a_1, a_2, \ldots, a_k$ of H_0 as a linear subspace of V.

Then V/H_0 is a Lie group isomorphic to $\mathbb{R}^{n-k}$. By 2.8.1, H/H_0 is a Lie subgroup of V/H_0. Moreover, it is a discrete subgroup. Hence, by 2.5.1, it is isomorphic to $\mathbb{Z}^{r-k}$ for some $r - k \leq \dim(V/H_0) = n - k$. More precisely, there exist $a_{k+1}, \ldots, a_r$ in H such that their images in V/H_0 are linearly independent and generate H/H_0.

The image H' of the map $\mathbb{R}^k \times \mathbb{Z}^{r-k} \ni (\alpha_1, \ldots, \alpha_k, m_{k+1}, \ldots, m_r) \longmapsto \alpha_1 a_1 + \cdots + \alpha_k a_k + m_{k+1} a_k + \cdots + m_r a_r$ is a Lie subgroup contained in H. It also contains H_0. On the other hand, H'/H_0 is the discrete subgroup in V/H_0 generated by the images of $a_{k+1}, \ldots, a_r$, i.e., it is equal to H/H_0. Therefore, $H' = H$. □

2.9. Exponential map. In this section we construct a differentiable map from the Lie algebra $L(G)$ of a Lie group G into G, which generalizes the exponential function $\exp : \mathbb{R} \longrightarrow \mathbb{R}_+^*$.

2.9.1. THEOREM. *Let G be a Lie group and $L(G)$ its Lie algebra. Then there exists a unique differentiable map $\varphi : L(G) \longrightarrow G$ with the following properties:*

(i) $\varphi(0) = 1$;
(ii) $T_0(\varphi) = 1_{L(G)}$;
(iii) $\varphi((t+s)\xi) = \varphi(t\xi)\varphi(s\xi)$ *for every* $t, s \in \mathbb{R}$ *and* $\xi \in L(G)$.

PROOF. We first prove the uniqueness part. Let φ_1 and φ_2 be two maps having the properties (i), (ii) and (iii). Take $\xi \in L(G)$. Then, because of (iii), $\psi_i(t) = \varphi_i(t\xi)$, $t \in \mathbb{R}$, are Lie group morphisms of $\mathbb{R}$ into G for $i = 1, 2$. Because of (ii), $T_0(\psi_i)(1) = \xi$, for $i = 1, 2$; hence, $L(\psi_1) = L(\psi_2)$. Since $\mathbb{R}$ is connected, by 2.2.16, it follows that $\psi_1 = \psi_2$. This implies that $\varphi_1(\xi) = \varphi_2(\xi)$. since ξ was arbitrary, it follows that $\varphi_1 = \varphi_2$.

It remains to show the existence. Let $\xi \in L(G)$. By 2.4.2, since $\mathbb{R}$ is a simply connected, connected Lie group, the morphism $t \longmapsto t\xi$ from $\mathbb{R}$ into $L(G)$ determines a unique Lie group morphism $f_\xi : \mathbb{R} \longrightarrow G$ such that $L(f_\xi)(1) = \xi$.

Let $s \in \mathbb{R}$. Then $c_s : t \longmapsto st$, $t \in \mathbb{R}$, is a Lie group homomorphism of $\mathbb{R}$ into itself. Clearly, $L(c_s) : t \longmapsto st$, $t \in \mathbb{R}$. Therefore, the composition $f_\xi \circ c_s$ is a Lie group morphism of $\mathbb{R}$ into G with the differential

$$L(f_\xi \circ c_s)(1) = L(f_\xi)(L(c_s)(1)) = T_0(f_\xi)(s) = sT_0(f_\xi)(1).$$

Therefore, $L(f_\xi \circ c_s) = L(f_{s\xi})$, and by 2.2.16, we have

$$f_\xi(st) = (f_\xi \circ c_s)(t) = f_{s\xi}(t)$$

for all $t \in \mathbb{R}$.

Consider the map $\varphi(\xi) = f_\xi(1)$ for $\xi \in L(G)$. Clearly, $\varphi(0) = f_0(1) = 1$. Hence, φ satisfies (i).

In addition, by the above calculation, for $t, s \in \mathbb{R}$ and $\xi \in L(G)$, we have

$$\varphi((t+s)\xi) = f_{(t+s)\xi}(1) = f_\xi(t+s) = f_\xi(t)f_\xi(s) = f_{t\xi}(1)f_{s\xi}(1) = \varphi(t\xi)\varphi(s\xi).$$

Therefore, (iii) also holds.

It remains to prove the differentiablity of φ and (ii).

First we prove that the function φ is differentiable in a neighborhood of $0 \in L(G)$. Clearly,

$$(\gamma_G(f_\xi(t)) \circ f_\xi)(s) = f_\xi(t)f_\xi(s) = f_\xi(t+s) = (f_\xi \circ \gamma_{\mathbb{R}}(t))(s)$$

for any $t, s \in \mathbb{R}$. Therefore,

$$\begin{aligned} T_t(f_\xi)(1) &= T_t(f_\xi)(T_0(\gamma_{\mathbb{R}}(t))(1)) = T_0(f_\xi \circ \gamma_{\mathbb{R}}(t))(1) \\ &= T_0(\gamma_G(f_\xi(t)) \circ f_\xi)(1) = T_1(\gamma_G(f_\xi(t)))(T_0(f_\xi)(1)) = T_1(\gamma_G(f_\xi(t)))\xi \end{aligned}$$

for any $t \in \mathbb{R}$.

Let (U, ψ, n) be a chart on G centered at 1. Denote by $D_1, D_2, \dots, D_n$ the vector fields on U which correspond to $\partial_1, \partial_2, \dots, \partial_n$ on $\psi(U)$ under the diffeomorphism ψ. Then $\xi_1 = D_{1,1}, \xi_2 = D_{2,1}, \dots, \xi_n = D_{n,1}$ form a basis of $T_1(G)$. Moreover,

$$T_1(\gamma_G(g))\xi_i = \sum_{j=1}^{n} (F_{ij} \circ \psi)(g) D_{j,g}$$

for any $g \in U$; where $F_{ij} : \psi(U) \longrightarrow \mathbb{R}$ are smooth functions. For $\xi = \sum_{i=1}^n x_i \xi_i \in L(G)$ there exists $\epsilon(x_1, x_2, \dots, x_n) > 0$ such that

$$|t| < \epsilon(x_1, x_2, \dots, x_n) \text{ implies } \varphi(t\xi) \in U.$$

We denote by $\psi_i(u)$, $1 \le i \le n$, the coordinates of $\psi(u)$ for $u \in U$, and put

$$f_i(t; x_1, x_2, \dots, x_n) = \psi_i(\varphi(t\xi))$$

for $|t| < \epsilon(x_1, x_2, \dots, x_n)$. Then, by the above calculation, we have

$$\begin{aligned} \frac{df_j}{dt} &= T_t(f_j)(1) = T_t(\psi_j \circ f_\xi)(1) = T_{\varphi(t\xi)}(\psi_j) T_t(f_\xi)(1) \\ &= T_{\varphi(t\xi)}(\psi_j) T_1(\gamma_G(f_\xi(t)))\xi = \sum_{i=1}^n x_i F_{ij}(f_1(t; x_1, \dots, x_n), \dots, f_n(t; x_1, \dots, x_n)) \end{aligned}$$

for every $|t| < \epsilon(x_1, x_2, \dots, x_n)$ and $x_1, x_2, \dots, x_n \in \mathbb{R}$. In addition, we have

$$f_i(0; x_1, x_2, \dots, x_n) = 0$$

for $1 \le i \le n$. If we consider the first order system of differential equations

$$\frac{df_j}{dt} = \sum_{i=1}^n x_i F_{ij}(f_1(t; x_1, \dots, x_n), \dots, f_n(t; x_1, \dots, x_n))$$

with the initial conditions

$$f_i(0; x_1, x_2, \dots, x_n) = 0$$

for $1 \le i \le n$, it follows that this Cauchy problem has unique solution on $(-\delta, \delta)$ with parameters $|x_i| < \epsilon$, $1 \le i \le n$, for some $\epsilon, \delta > 0$. Moreover, the solutions f_i, $1 \le i \le n$, are smooth functions in $|t| < \delta$ and $|x_i| < \epsilon$, $1 \le i \le n$. Therefore, if we put $V = \{\xi \in L(G) \mid \xi = \sum_{i=1}^n x_i \xi_i, |x_i| < \delta\epsilon\}$, V is an open neighborhood of 0 in $L(G)$ and the function φ is differentiable on V.

On the other hand, by (iii) we have

$$\varphi(\xi) = \varphi\left(\frac{1}{n}\xi\right)^n$$

for any $n \in \mathbb{N}$. Therefore, the differentiability of φ on V implies the differentiability on nV for any $n \in \mathbb{N}$. Hence φ is differentiable on $L(G)$.

Finally, if ξ is in $L(G)$, we have the differentiable map $t \longmapsto \varphi(t\xi) = f_\xi(t)$ on $\mathbb{R}$. Moreover,

$$\xi = L(f_\xi)(1) = T_0(f_\xi)(1) = T_0(\varphi)(\xi),$$

i.e., $T_0(\varphi) = 1_{L(G)}$, and (ii) holds. □

The map $\varphi : L(G) \longrightarrow G$ is called the *exponential map* and denoted by $\exp_G$ (or just exp).

Let G be the multiplicative group of positive real numbers $\mathbb{R}_+^*$. Then its Lie algebra is equal to $\mathbb{R}$. Clearly, the function $t \longmapsto e^t$ satisfies the properties (i), (ii) and (iii) of 2.9.1. Therefore, in this example we have $\varphi(t) = e^t$ for $t \in \mathbb{R}$.

2.9.2. Corollary. (i) *Exponential map* $\exp_G : L(G) \longrightarrow G$ *is a local diffeomorphism at* $0 \in L(G)$.

(ii) *For every* $\xi \in L(G)$, $\psi : t \longmapsto \exp(t\xi)$ *is the unique Lie group morphism of* $\mathbb{R}$ *into* G *such that* $L(\psi)(1) = \xi$.

For every $\xi \in L(G)$, $\{\exp(t\xi) \mid t \in \mathbb{R}\}$ is an integral subgroup of G which we call *one-parameter subgroup* attached to ξ. From 2.6.3, we see that one-parameter subgroups are isomorphic to either $\mathbb{R}$ or $\mathbb{T}$.

2.9.3. PROPOSITION. *Let G and H be two Lie groups and $\varphi : G \longrightarrow H$ a morphism of Lie groups. Then*

(i) $\varphi \circ \exp_G = \exp_H \circ L(\varphi)$;
(ii) *if G is an integral subgroup of H, we have* $\exp_G = \exp_H |_{L(G)}$.

PROOF. Clearly, (ii) is a special case of (i).

To prove (i) we remark that $\psi_1 : t \longmapsto \varphi(\exp_G(t\xi))$ and $\psi_2 : t \longmapsto \exp_H(tL(\varphi)\xi)$ are two Lie group morphisms of $\mathbb{R}$ into H. Also, we have

$$L(\psi_1)(1) = L(\varphi)\xi = L(\psi_2)(1),$$

i.e., $L(\psi_1) = L(\psi_2)$. By 2.2.16, it follows that $\psi_1 = \psi_2$. In particular, we have

$$\varphi(\exp_G(\xi)) = \psi_1(1) = \psi_2(1) = \exp_H(L(\varphi)\xi). \quad \square$$

Let $G = \mathrm{GL}(V)$. Then $L(G)$ is the Lie algebra $\mathcal{L}(V)$ of all linear endomorphisms on V. For any linear transformation T on V, the series $\sum_{n=0}^{\infty} \frac{1}{n!} T^n$ converges to a regular linear transformation on V. Therefore, this defines a real analytic map $T \longmapsto e^T$ from $\mathcal{L}(V)$ into $\mathrm{GL}(V)$. Clearly, this map satisfies the properties (i), (ii) and (iii) from 2.9.1. Hence $\exp(T) = e^T$ for $T \in \mathcal{L}(G)$.

2.9.4. COROLLARY. (i) *Let $\xi \in L(G)$. Then*

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

(ii) *Let $g \in G$. Then*

$$g(\exp\xi)g^{-1} = \exp(\mathrm{Ad}(g)(\xi))$$

for all $\xi \in L(G)$.

PROOF. (i) The adjoint representation Ad is a Lie group morphism of G into $\mathrm{GL}(L(G))$. Therefore, by 2.9.3 and the above discussion, we have

$$\mathrm{Ad}(\exp(\xi)) = e^{L(\mathrm{Ad})\xi}.$$

The final statement follows from 2.2.5.

(ii) $\mathrm{Int}(g)$ is an automorphism of G, hence, by 2.9.3, we have

$$g(\exp\xi)g^{-1} = \mathrm{Int}(g)(\exp(\xi)) = \exp(L(\mathrm{Int}(g))\xi) = \exp(\mathrm{Ad}(g)\xi)$$

for all $\xi \in L(G)$. $\square$

2.9.5. COROLLARY. *Let G be a Lie group and H an integral subgroup of G. Then, the following statements are equivalent for $\xi \in L(G)$:*

(i) $\xi \in L(H)$;
(ii) $\exp_G(t\xi) \in H$ *for all* $t \in \mathbb{R}$.

PROOF. By 2.9.3.(ii), we see that $\xi \in L(H)$ implies that $\exp_G(t\xi) \in H$ for all $t \in \mathbb{R}$.

If $\exp_G(t\xi) \in H$ for all $t \in \mathbb{R}$, then ξ is in the Lie algebra tangent to H. Hence, by 2.7.6, we see that $\xi \in L(H)$. $\square$

Clearly, the image of $\exp : L(G) \longrightarrow G$ is in the identity component of G. On the other hand, exp in general is neither injective nor surjective. The answer is simple only in the case of connected abelian Lie groups.

2.9.6. PROPOSITION. *Let G be a connected Lie group. Then the following assertions are equivalent:*

(i) *the group G is abelian;*
(ii) $\exp : L(G) \longrightarrow G$ *is a Lie group morphism of the additive group $L(G)$ into G.*

If these conditions are satisfied, $\exp : L(G) \longrightarrow G$ *is a covering projection.*

PROOF. Assume that G is a simply connected abelian Lie group. Then $L(G)$ is an abelian Lie algebra by 2.2.18. In addition, by 2.6.1, G is isomorphic to $\mathbb{R}^n$ for $n = \dim G$. Moreover, $L(G)$ is also isomorphic to $\mathbb{R}^n$ as an abelian Lie algebra. Clearly, the identity map on $\mathbb{R}^n$ satisfies the conditions of 2.9.1. Therefore, exp is the identity map in this case, so it is clearly a Lie group morphism.

If G is an arbitrary connected abelian Lie group, its universal cover is isomorphic to $\mathbb{R}^n$ for $n = \dim G$. Let $p : \mathbb{R}^n \longrightarrow G$ be the covering projection. Then, by 2.9.3 and the first part of the proof, we have $p = \exp_G$. It follows that $\exp_G$ is a Lie group morphism and the covering projection.

If $\exp : L(G) \longrightarrow G$ is a Lie group morphism, its image is a subgroup of G. By 2.9.2.(i), it contains an open neighborhood of 1 in G. Since G is connected, by 1.5.1, we see that exp is surjective. Therefore, G has to be abelian. □

2.9.7. LEMMA. *Let G be a connected Lie group and H an integral subgroup of G. Then the following conditions are equivalent:*

(i) *H is a normal subgroup of G;*
(ii) *$L(H)$ is an ideal in $L(G)$.*

PROOF. Assume that H is a normal subgroup in G. Then by 2.7.8, $L(H)$ is an ideal in $L(G)$.

If $L(H)$ is an ideal in $L(G)$, by 2.9.4, we have

$$\operatorname{Ad}(\exp(\xi))(L(H)) = e^{\operatorname{ad}(\xi)}(L(H)) = L(H)$$

for any $\xi \in L(G)$. By 2.9.1, there exists a neighborhood U of 1 in G such that $\operatorname{Ad}(g)(L(H)) = L(H)$ for all $g \in U$. Since G is connected, by 1.5.1, it follows that $\operatorname{Ad}(g)(L(H)) = L(H)$ for all $g \in G$. Hence, by 2.7.8, H is a normal subgroup. □

2.10. Some examples. First we consider the group of affine transformations of the space $\mathbb{R}^n$, $n \in \mathbb{N}$. For $A \in \mathrm{GL}(n,\mathbb{R})$, $a \in \mathbb{R}^n$, we define the affine transformation

$$\alpha_{A,a}(x) = Ax + a, \ x \in \mathbb{R}^n.$$

Clearly, for $A, B \in \mathrm{GL}(n,\mathbb{R})$ and $a, b \in \mathbb{R}^n$, we have

$$\alpha_{A,a} \circ \alpha_{B,b}(x) = \alpha_{A,a}(Bx + b) = ABx + Ab + a = \alpha_{Ab+a,AB}(x)$$

for all $x \in \mathbb{R}^n$. Therefore, the group of all affine transformations of the real line can be identified with the the manifold $\mathbb{R}^n \times \mathrm{GL}(n,\mathbb{R})$ with the operation

$$(a, A) \cdot (b, B) = (Ab + a, AB).$$

This is clearly a Lie group G, which we call the group of affine transformations of $\mathbb{R}^n$.

We define a map $\pi : (a, A) \longmapsto \begin{pmatrix} A & a \\ 0 & 1 \end{pmatrix}$ of G into $\mathrm{GL}(n+1, \mathbb{R})$. Clearly, we have

$$\pi(a, A) \circ \pi(b, B) = \begin{pmatrix} A & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} B & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} AB & Ab + a \\ 0 & 1 \end{pmatrix} = \pi((A, a) \cdot (B, b)),$$

i.e., π is a representation of G. The image of π is the subgroup H of $\mathrm{GL}(n+1, \mathbb{R})$ which is the intersection of the open submanifold $\mathrm{GL}(n+1, \mathbb{R})$ of the space $M_{n+1}(\mathbb{R})$ of all $(n+1) \times (n+1)$ real matrices with the closed submanifold of all matrices having the second row equal to $(0 \dots 0\ 1)$. Therefore, H is a Lie subgroup of $\mathrm{GL}(n+1, \mathbb{R})$. Since π is injective, by 1.5.7, π is an isomorphism of G onto H.

Therefore, the Lie algebra $L(G)$ of G is isomorphic to the Lie algebra $L(H)$ of H. On the other hand, the Lie algebra $L(H)$ is the subalgebra of the Lie algebra $M_{n+1}(\mathbb{R})$ consisting of all matrices with with last row equal to zero.

Consider now the case $n = 1$. Then G is diffeomorphic to $\mathbb{R} \times \mathbb{R}^*$. Therefore, it has two components, and the identity component G_0 is simply connected. The Lie algebra of G is spanned by the vectors

$$e_1 = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \text{ and } e_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

By direct calculation we check that $[e_1, e_2] = e_1$. Therefore, the Lie algebra $L(G)$ is isomorphic to the unique two-dimensional nonabelian Lie algebra which we discussed in 2.1.

Let (a, b) be in the center of G_0. Then

$$(a, b) \cdot (c, d) = (a + bc, bd)$$

is equal to

$$(c, d) \cdot (a, b) = (c + da, bd)$$

for all $c \in \mathbb{R}$ and $d \in \mathbb{R}_+^*$. This implies that $a + bc = c + da$ for all $c \in \mathbb{R}$ and $d \in \mathbb{R}_+^*$. This is possible only if $a = 0$ and $b = 1$. Therefore, the center of G_0 (and of G) is trivial. This implies that, up to a Lie group isomorphism, G_0 is the unique connected Lie group with Lie algebra isomorphic to $L(G)$.

Combining this with the above discussion, we get the following result.

2.10.1. Lemma. *The connected component of the group of affine transformations of the real line is (up to an isomorphism) the unique connected 2-dimensional nonabelian Lie group.*

Combined with 2.6.2, this completes the classification of all connected Lie groups of dimension ≤ 2.

2.10.2. Proposition. *Any connected Lie group G of dimension 2 is isomorphic to one of the following Lie groups:*

(i) *real plane $\mathbb{R}^2$;*
(ii) *two-dimensional torus $\mathbb{T}^2$;*
(iii) *the product $\mathbb{R} \times \mathbb{T}$;*
(iv) *the connected component G_0 of the group of affine motions of the real line.*

By direct calculation we see that

$$\operatorname{Ad}(a,b)e_1 = \begin{pmatrix} b & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b & a \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} b & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b^{-1} & -ab^{-1} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -b \\ 0 & 0 \end{pmatrix} = be_1;$$

and

$$\operatorname{Ad}(a,b)e_2 = \begin{pmatrix} b & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b & a \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} b & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b^{-1} & -ab^{-1} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -a \\ 0 & 0 \end{pmatrix} = ae_1 + e_2.$$

Therefore, in the basis (e_1, e_2) the adjoint representation of G is equal to the representation π.

Now we return to the general case. We want to calculate the exponential map for G. The Lie algebra $L(G)$ can be viewed as the Lie subalgebra of $M_{n+1}(\mathbb{R})$ consisting of all matrices with last row equal to 0. Therefore, an element of $L(G)$ can be written as $\begin{pmatrix} T & v \\ 0 & 0 \end{pmatrix}$ where $T \in M_n(\mathbb{R})$ and $v \in \mathbb{R}^n$. Since G is a Lie subgroup of $\mathrm{GL}(n+1,\mathbb{R})$ its exponential map is given by the usual exponential function on $M_{n+1}(\mathbb{R})$. Therefore, me have

$$\exp \begin{pmatrix} T & v \\ 0 & 0 \end{pmatrix} = \sum_{p=0}^{\infty} \frac{1}{p!} \begin{pmatrix} T & v \\ 0 & 0 \end{pmatrix}^p .$$

By induction in p we see that

$$\begin{pmatrix} T & v \\ 0 & 0 \end{pmatrix}^p = \begin{pmatrix} T^p & T^{p-1}v \\ 0 & 0 \end{pmatrix}$$

for any $p \in \mathbb{N}$. Let

$$f(t) = \sum_{p=0} \frac{t^p}{(p+1)!}$$

for any $t \in \mathbb{C}$. Then, f is an entire function, and for $t \neq 0$ we have $f(t) = \frac{e^t - 1}{t}$. With this notation we have

$$\exp \begin{pmatrix} T & v \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} e^T & f(T)v \\ 0 & 1 \end{pmatrix}.$$

In particular, returning to the case $n = 1$, we see that

$$\exp \begin{pmatrix} t & v \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} e^t & f(t)v \\ 0 & 1 \end{pmatrix},$$

for any $t, v \in \mathbb{R}$. On the other hand, the identity component G_0 of G consists of matrices $\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$ where $a > 0$. If $t = 0$, $f(0) = 1$, and we have

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

If $t \neq 0$, we have

$$\exp \begin{pmatrix} t & v \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} e^t & \frac{(e^t-1)v}{t} \\ 0 & 1 \end{pmatrix} .$$

From these formulae it is easy to see that the exponential map is a diffeomorphism of $L(G)$ onto G_0.

From the above discussion we conclude that the exponential map is a diffeomorphism for all simply connected, connected Lie groups of dimension 1 and 2.

Consider now the group G for $n = 2$. This is the group of affine transformations of the plane $\mathbb{R}^2$. Let H be the subgroup of G consisting of all affine transformations which preserve the euclidean distance in $\mathbb{R}^2$. This is the group of *euclidean motions* of $\mathbb{R}^2$. From the above discussion, we see that H consists of all matrices $\begin{pmatrix} A & a \\ 0 & 1 \end{pmatrix}$ where $A \in \mathrm{O}(2)$ and $a \in \mathbb{R}^2$. Therefore, H is diffeomorphic to $\mathbb{R}^2 \times \mathrm{O}(2)$. By 1.8.4, H has two connected components. Its identity component H_0 is the group of orientation preserving euclidean motions consisting of all matrices of the form $\begin{pmatrix} A & a \\ 0 & 1 \end{pmatrix}$ where $A \in \mathrm{SO}(2)$ and $a \in \mathbb{R}^2$. Therefore, its fundamental group is isomorphic to $\mathbb{Z}$.

Consider the manifold $\tilde{H} = \mathbb{R}^3$ with multiplication

$$(x, y, \varphi) \cdot (x', y', \varphi') = (x + x' \cos\varphi + y' \sin\varphi, y - x' \sin\varphi + y' \cos\varphi, \varphi + \varphi').$$

By direct calculation, one can check that this is a Lie group. Moreover, the mapping $\Phi : \tilde{H} \longrightarrow H_0$ given by

$$\Phi(x, y, \varphi) = \begin{pmatrix} \cos\varphi & \sin\varphi & x \\ -\sin\varphi & \cos\varphi & y \\ 0 & 0 & 1 \end{pmatrix}$$

is a Lie group morphism. The kernel of Φ is $(0, 0, 2\pi k)$, $k \in \mathbb{Z}$, and Φ is surjective. Therefore, Φ is a covering projection. It follows that $\tilde{H}$ is the universal cover of H_0.

The Lie algebra of H is spanned by matrices

$$e_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ e_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \ e_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} .$$

Then, we have

$$[e_1, e_2] = 0, \ [e_3, e_1] = -e_2, \ [e_3, e_2] = e_1$$

and these relations determine $L(H)$ completely.

Now we consider the exponential map $\exp : L(H) \longrightarrow H$. By induction we see that

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^p = \begin{cases} (-1)^{\frac{p}{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{if } p \text{ is even;} \\ (-1)^{\frac{(p-1)}{2}} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & \text{if } p \text{ is odd.} \end{cases}$$

Therefore, we have

$$e^{\begin{pmatrix} 0 & \varphi \\ -\varphi & 0 \end{pmatrix}} = \left(\sum_{p=0}^{\infty} \frac{(-1)^p}{(2p)!}\varphi^{2p}\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \left(\sum_{p=0}^{\infty} \frac{(-1)^p}{(2p+1)!}\varphi^{2p+1}\right)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}.$$

Analogously, if $\varphi \neq 0$, we have

$$\begin{aligned} f\begin{pmatrix} 0 & \varphi \\ -\varphi & 0 \end{pmatrix} &= \left(\sum_{p=0}^{\infty} \frac{(-1)^p}{(2p+1)!}\varphi^{2p}\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \left(\sum_{p=0}^{\infty} \frac{(-1)^p}{(2p+2)!}\varphi^{2p+1}\right)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= \frac{\sin\varphi}{\varphi}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1-\cos\varphi}{\varphi}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \frac{1}{\varphi}\begin{pmatrix} \sin\varphi & 1-\cos\varphi \\ \cos\varphi - 1 & \sin\varphi \end{pmatrix}. \end{aligned}$$

Hence, by above calculation, we have

$$\exp_H \begin{pmatrix} 0 & 0 & x \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & x \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

and, if $\varphi \neq 0$,

$$\exp_H \begin{pmatrix} 0 & \varphi & x \\ -\varphi & 0 & y \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi & \frac{1}{\varphi}(x\sin\varphi + y(1-\cos\varphi)) \\ -\sin\varphi & \cos\varphi & \frac{1}{\varphi}(x(\cos\varphi - 1) + y\sin\varphi) \\ 0 & 0 & 1 \end{pmatrix}.$$

From this one easily sees that

$$\exp_{\tilde{H}} \begin{pmatrix} 0 & 0 & x \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} = (x, y, 0)$$

and

$$\exp_{\tilde{H}} \begin{pmatrix} 0 & \varphi & x \\ -\varphi & 0 & y \\ 0 & 0 & 0 \end{pmatrix} = \left(\frac{1}{\varphi}(x\sin\varphi + y(1-\cos\varphi)), \frac{1}{\varphi}(x(\cos\varphi - 1) + y\sin\varphi), \varphi\right)$$

for $\varphi \neq 0$.

From this we immediately deduce that for $k \in \mathbb{Z}^*$, we have

$$\exp_{\tilde{H}} \begin{pmatrix} 0 & 2\pi k & x \\ -2\pi k & 0 & y \\ 0 & 0 & 0 \end{pmatrix} = (0, 0, 2\pi k).$$

Hence, the exponential map is neither injective nor surjective for $\tilde{H}$.

2.11. Cartan's theorem. In this section we prove the following fundamental result in the theory of Lie groups.

2.11.1. THEOREM (E. Cartan). *A closed subgroup of a Lie group is a Lie subgroup.*

PROOF. Let G be a Lie group and H its closed subgroup. Let $\mathfrak{h}$ be the Lie subalgebra of $L(G)$ which is tangent to H. Denote by H_0 the integral subgroup of G attached to $\mathfrak{h}$. Then, by 2.7.2, $H_0 \subset H$.

Let $\mathfrak{k}$ be a complementary linear subspace to $\mathfrak{h}$ in $L(G)$. Then the map $\mathfrak{h} \times \mathfrak{k} \ni (\xi_1, \xi_2) \longmapsto \exp(\xi_1) \cdot \exp(\xi_2)$ is a differentiable map from $\mathfrak{h} \times \mathfrak{k}$ into G. By 2.9.1, the differential of this map at $(0,0)$ is equal to $(\xi_1, \xi_2) \longmapsto \xi_1 + \xi_2$. Hence the map is a local diffeomorphism at $(0,0)$. There exist open symmetric convex neighborhoods U_1 and U_2 of 0 in $\mathfrak{h}$ and $\mathfrak{k}$ respectively, such that $(\xi_1, \xi_2) \longmapsto \exp(\xi_1) \cdot \exp(\xi_2)$ is a diffeomorphism of $U_1 \times U_2$ onto an open neighborhood U of 1 in G.

Clearly, $\exp(U_1) \subset H_0$. We claim that there is a neighborhood $U_2' \subset U_2$ of $0 \in \mathfrak{k}$ such that

$$H \cap \exp(U_1) \cdot \exp(U_2') = \exp(U_1).$$

Assume the opposite. Then there exist sequences (ξ_n) in U_1 and (η_n) in $U_2 - \{0\}$ such that $\eta_n \longrightarrow 0$ and $\exp(\xi_n) \cdot \exp(\eta_n) \in H$. Since we have $\exp(\eta_n) = \exp(-\xi_n)\exp(\xi_n)\exp(\eta_n)$, we see that $\exp(\eta_n) \in H$ for all $n \in \mathbb{N}$. Taking possibly a subsequence, we can find $\lambda_n \in \mathbb{R} - \{0\}$, $n \in \mathbb{N}$, such that $\lambda_n^{-1}\eta_n \longrightarrow \eta \in \mathfrak{k} - \{0\}$ as $n \to \infty$. For example, if we take a norm $\|\cdot\|$ on $\mathfrak{k}$, we can put $\lambda_n = \|\eta_n\|$. Clearly, we must have $\lambda_n \longrightarrow 0$ as $n \to \infty$. Let $\lambda \in \mathbb{R}$. Let k_n be the largest integer less than or equal to $\lambda\lambda_n^{-1}$. Then $|\lambda - \lambda_n k_n| \longrightarrow 0$ as $n \to \infty$. Therefore, by the continuity of the exponential map, we have

$$\begin{aligned}\exp(\lambda\eta) = \exp\left(\lambda \lim_{n\to\infty} \lambda_n^{-1}\eta_n\right) &= \lim_{n\to\infty} \exp(\lambda\lambda_n^{-1}\eta_n) \\ &= \lim_{n\to\infty} \left(\exp((\lambda - k_n\lambda_n)\lambda_n^{-1}\eta_n) \cdot \exp(k_n\eta_n)\right).\end{aligned}$$

On the other hand, by the same reasoning, we have

$$\lim_{n\to\infty} \exp((\lambda - k_n\lambda_n)\lambda_n^{-1}\eta_n) = \exp\left(\lim_{n\to\infty}((\lambda - k_n\lambda_n)\lambda_n^{-1}\eta_n)\right) = \exp(0) = 1.$$

Therefore, we have

$$\begin{aligned}\exp(\lambda\eta) &= \lim_{n\to\infty} \left(\exp((\lambda - k_n\lambda_n)\lambda_n^{-1}\eta_n) \cdot \exp(k_n\eta_n)\right) \\ &= \lim_{n\to\infty} \exp(-(\lambda - k_n\lambda_n)\lambda_n^{-1}\eta_n) \cdot \lim_{n\to\infty} \left(\exp((\lambda - k_n\lambda_n)\lambda_n^{-1}\eta_n) \cdot \exp(k_n\eta_n)\right) \\ &= \lim_{n\to\infty} \exp(k_n\eta_n) = \lim_{n\to\infty} \exp(\eta_n)^{k_n} \in H,\end{aligned}$$

since $k_n \in \mathbb{Z}$, $\exp(\eta_n) \in H$ and H is closed in G.

It follows that $\exp(\lambda\eta) \in H$ for all $\lambda \in \mathbb{R}$. Hence, $\eta \in \mathfrak{h}$, which is impossible. Therefore, we have a contradiction.

Therefore, we established that there exists a neighborhood U_2' of 0 in $\mathfrak{k}$ such that $H \cap \exp(U_1)\exp(U_2') = \exp(U_1)$. Hence, there exists an open neighborhood O of the identity such that $H \cap O$ is a submanifold. Since H is a subgroup of G, any $h \in H$ has such neighborhood $\gamma(h)(O)$. Therefore, H is a submanifold of G and a Lie subgroup of G. □

Cartan's theorem has the following consequence.

2.11.2. THEOREM. *Let G and H be two Lie groups and $\varphi : G \longrightarrow H$ a continuous group homomorphism. Then φ is differentiable, i.e., it is a Lie group morphism.*

PROOF. Since φ is continuous, the graph Γ_φ of φ in $G \times H$ is a closed subgroup. Therefore, Γ_f is a Lie subgroup in $G \times H$. Hence, the restriction p of the projection

$G \times H \longrightarrow G$ to Γ_φ is a morphism of Lie groups. Clearly, it is a homeomorphism. Since it is injective, by 1.1.5 and 2.2.9, it must be an immersion. On the other hand, since it is open, it must be a local diffeomorphism. It follows that p is a diffeomorphism, and an isomorphism of Lie groups. On the other hand, φ is the composition of the inverse of p with the projection to the second factor in $G \times H$. This implies that φ is a Lie group morphism. □

2.11.3. COROLLARY. *Let G be a locally compact group. Then on G there exists at most one structure of a Lie group (compatible with the topology of G).*

2.12. A categorical interpretation. Let $\mathcal{L}ie$ be the category of Lie groups and $\mathcal{T}op\mathcal{G}rp$ the category of topological groups. Then we have the natural forgetful functor For : $\mathcal{L}ie \longrightarrow \mathcal{T}op\mathcal{G}rp$. By 2.11.2 this functor is fully faithful. Moreover, by 2.11.3, this functor is an isomorphism of the category $\mathcal{L}ie$ with the full subcategory of $\mathcal{T}op\mathcal{G}rp$ consisting of topological groups which admit a compatible Lie group structure.

The following property distinguishes Lie groups among topological groups.

2.12.1. PROPOSITION. *Let G be a Lie group. Then there exists a neighborhood U of 1 in G with the following property: If H is a subgroup of G contained in U, H is trivial, i.e., $H = \{1\}$.*

We say that Lie groups do not admit small subgroups.

PROOF. Let U be an open neighborhood of 1 in G and V a bounded open convex neighborhood of 0 in $L(G)$ such that $\exp : V \longrightarrow U$ is a diffeomorphism. Let $V' \subset \frac{1}{2}V \subset V$ be another neighborhood of 0 in $L(G)$. Then $U' = \exp(V')$ is an open neighborhood of 1 in G. Let H be a subgroup of G contained in U'. Let $h \in H$. Then $h = \exp(\xi)$ for some $\xi \in V'$. Hence, we have $h^2 = \exp(\xi)^2 = \exp(2\xi) \in H$. Moreover, $h^2 \in H$ and $h^2 = \exp(\eta)$ for some $\eta \in V'$. It follows that $\exp(\eta) = \exp(2\xi)$ for $2\xi, \eta \in V$. Since exp is injective on V, we must have $2\xi = \eta$. Hence, $\xi \in \frac{1}{2}V'$. It follows that $H \subset \exp\left(\frac{1}{2}V'\right)$. By induction we get that $H \subset \exp\left(\frac{1}{2^n}V'\right)$ for any $n \in \mathbb{N}$. Since V' is bounded, this implies that $H = \{1\}$. □

2.12.2. EXAMPLE. In contrast to 2.12.1, there exist compact groups with small subgroups. For example, let $C = \mathbb{Z}/2\mathbb{Z}$ be the cyclic group of order two, and G the infinite product of countably many copies of C. Then G is a compact group. On the other hand, by the definition of topology on G, there exists a fundamental system of open neighborhoods of 1 in G consisting of subgroups of finite index in G.

This proves that the full subcategory of $\mathcal{T}op\mathcal{G}rp$ consisting of all locally compact groups is strictly larger than $\mathcal{L}ie$.

On the other hand, a connected locally compact group without small subgroups is a Lie group. In particular, a topological group which is a topological manifold has no small subgroups and therefore is a Lie group. This gives the positive answer to Hilbert's fifth problem.

2.13. Closures of one-parameter subgroups. Let G be a Lie group. Let H be a subgroup of G. By continuity of multiplication and inversion in G, the closure $\bar{H}$ of H is a closed subgroup of G. By Cartan's theorem 2.11.1, $\bar{H}$ is a Lie subgroup of G.

Let $\xi \in L(G)$ and H the corresponding one-parameter subgroup $\{\exp(t\xi) \mid t \in \mathbb{R}\}$. Then, by 2.6.3, H is isomorphic to $\mathbb{R}$ or $\mathbb{T}$. In the second case, H is compact, and therefore closed in G.

We want to study the closure $\bar{H}$ of H in the first case. Since H is connected and abelian, $\bar{H}$ must be a connected abelian Lie group. Hence, by 2.6.2, $\bar{H}$ is isomorphic to a product $\mathbb{T}^p \times \mathbb{R}^q$ for some $p, q \in \mathbb{Z}_+$.

The universal cover of $\bar{H}$ is isomorphic to $\mathbb{R}^{p+q}$. The Lie algebra $L(\bar{H})$ can also be identified with $\mathbb{R}^{p+q}$ and the exponential map $\exp : \mathbb{R}^{p+q} \longrightarrow \bar{H}$ is the covering projection by 2.9.6. We can assume that the kernel of this covering projection is $\mathbb{Z}^p \times \{0\}$. Since $\xi \in L(H) \subset L(\bar{H})$, ξ determines a line in $L(\bar{H})$. Let $e_1, e_2, \dots, e_{p+q}$ denote the canonical basis of $\mathbb{R}^{p+q}$. Then $e_1, e_2, \dots, e_p$ and the line $\{t\xi \mid t \in \mathbb{R}\}$ generate a subgroup K of $\mathbb{R}^{p+q}$. Let U be a nonempty open subset of $\mathbb{R}^{p+q}$. Since the projection of $\mathbb{R}^{p+q}$ onto $\bar{H}$ is open, the image V of U is a nonempty open set in $\bar{H}$. Since H is dense in $\bar{H}$, V must intersect H. It follows that K intersects U. Hence, K is dense in $\mathbb{R}^{p+q}$. This first implies that $e_1, e_2, \dots, e_p$ and ξ must span $\mathbb{R}^{p+q}$. Hence, $q \leq 1$. On the other hand, if $q = 1$, ξ is linearly independent from $e_1, e_2, \dots, e_p$. In this case, K is closed in $\mathbb{R}^{p+1}$. Since it is also dense in $\mathbb{R}^{p+1}$, it must be equal to $\mathbb{R}^{p+1}$. This is possible only if $p = 0$ and $\bar{H}$ is one-dimensional. Since H is one-dimensional too, by 2.7.6, we see that $H = \bar{H}$. Therefore, we established the following result.

2.13.1. Proposition. *Let H be a one-parameter subgroup in a Lie group G. Then, either H is a Lie subgroup isomorphic to $\mathbb{R}$ or $\bar{H}$ is a Lie group isomorphic to $\mathbb{T}^n$ for some $n \in \mathbb{N}$.*

Now we want to show that any torus $\mathbb{T}^n$ can be obtained in this way.

2.13.2. Proposition. *Let $n \in \mathbb{N}$. There exists a one-parameter subgroup dense in $\mathbb{T}^n$.*

Proof. As we remarked above, it is enough to show that for any $n \in \mathbb{N}$ there exists a line L in $\mathbb{R}^n$ such that it and $e_1, e_2, \dots, e_n$ generate a dense subgroup H in $\mathbb{R}^n$.

Let L be an arbitrary line in $\mathbb{R}^n$ and H the subgroup generated by L and $e_1, e_2, \dots, e_n$. Then $\bar{H}$ is a closed subgroup of $\mathbb{R}^n$. By Cartan's theorem, $\bar{H}$ is a Lie subgroup of $\mathbb{R}^n$. Therefore, by 2.8.2, there exists a basis $a_1, a_2, \dots, a_n$ such that $(\alpha_1, \dots, \alpha_r, m_{r+1}, \dots, m_n) \longmapsto \alpha_1 a_1 + \cdots + \alpha_r a_r + m_{r+1} a_{r+1} + \cdots + m_n a_n$ is an isomorphism of $\mathbb{R}^r \times \mathbb{Z}^{n-r}$ onto $\bar{H}$. If $\bar{H}$ is different from $\mathbb{R}^n$, we have $r < n$. Let f be the linear form on $\mathbb{R}^n$ defined by $f(a_i) = 0$ for $1 \leq i \leq n-1$ and $f(a_n) = 1$. Then f is a nonzero linear form on $\mathbb{R}^n$ satisfying $f(\bar{H}) \subset \mathbb{Z}$.

Therefore, if H is not dense in $\mathbb{R}^n$, there exists a nontrivial linear form f on $\mathbb{R}^n$ such that $f(H) \subset \mathbb{Z}$.

Let $c_i = f(e_i)$ for $1 \leq i \leq n$. Since $e_1, e_2, \dots, e_n \in H$, we must have $c_i \in \mathbb{Z}$ for $1 \leq i \leq n$. Let $\xi = (\theta_1, \theta_2, \dots, \theta_n)$ be a nonzero element of L. Since f takes integral values on L, it must be equal to 0 on L. Therefore, we must have

$$c_1\theta_1 + c_2\theta_2 + \cdots + c_n\theta_n = 0.$$

Therefore, if a nontrivial f exists, $\theta_1, \theta_2, \dots, \theta_n$ must be linearly dependent over $\mathbb{Q}$. Since $\mathbb{R}$ is infinite dimensional linear space over $\mathbb{Q}$, we see that we can always find ξ such that $\theta_1, \theta_2, \dots, \theta_n$ are linearly independent over $\mathbb{Q}$. In this case H must be dense in $\mathbb{R}^n$. Hence, the corresponding one-parameter subgroup is dense in $\mathbb{T}^n$. □

3. Haar measures on Lie groups

3.1. Existence of Haar measure. In this section we prove the existence of left invariant positive measures on Lie groups. They generalize the counting measure on a finite group G. The main result is the following theorem.

3.1.1. THEOREM. *Let G be a Lie group.*

(i) *There exists a nonzero left invariant positive measure μ on G.*
(ii) *Let ν be another left invariant measure on G. Then there exists $c \in \mathbb{C}$ such that $\nu = c\mu$.*

Therefore, any nonzero positive left invariant measure on G is of the form $c \cdot \mu$, $c > 0$. Such measure is called a *left Haar measure* on G.

Since a left Haar measure μ on G is left invariant, its support $\operatorname{supp}(\mu)$ must be a left invariant subset of G. Therefore, since μ is nonzero, $\operatorname{supp}(\mu)$ has to be equal to G. In particular, the measure $\mu(U)$ of a a nonempty open set U in G must be positive.

PROOF. Let $n = \dim G$. Then $\bigwedge^n T_1(G)^*$ is one-dimensional linear space. A nonzero n-form Ω in $\bigwedge^n T_1(G)^*$ determines a differentiable n-form ω on G by

$$\omega_g(T_1(\gamma(g))\xi_1 \wedge (T_1(\gamma(g))\xi_2 \wedge \cdots \wedge (T_1(\gamma(g))\xi_n) = \Omega(\xi_1 \wedge \xi_2 \wedge \cdots \wedge \xi_n)$$

for all $\xi_1, \xi_2, \ldots, \xi_n \in T_1(G)$. Clearly, ω satisfies $\gamma(g)^*\omega = \omega$ for any $g \in G$, i.e., this form is left invariant. The corresponding positive measure $|\omega|$ is a nonzero left invariant measure on G. This proves (i).

Now we prove the uniqueness of left invariant measures. Let μ be a nonzero positive left invariant measure on G. Let ν be another left invariant measure on G. Let $\varphi \in C_0(G)$ such that $\mu(\varphi) \neq 0$. Then we can define the function

$$F_\varphi(g) = \frac{1}{\mu(\varphi)} \int \varphi(hg)\, d\nu(h).$$

This is a continuous function on G.

For any $\psi \in C_0(G)$, we have

$$\begin{aligned}
\mu(\varphi) \int \psi(h^{-1})\, d\nu(h) &= \int \left(\int \varphi(g)\psi(h^{-1})\, d\nu(h) \right) d\mu(g) \\
&= \int \left(\int \varphi(g)\psi((g^{-1}h)^{-1})\, d\nu(h) \right) d\mu(g) = \int \left(\int \varphi(g)\psi(h^{-1}g)\, d\mu(g) \right) d\nu(h) \\
&= \int \left(\int \varphi(hg)\psi(g)\, d\mu(g) \right) d\nu(h) = \int \psi(g) \left(\int \varphi(hg)\, d\nu(h) \right) d\mu(g) \\
&= \mu(\varphi) \int \psi(g) F_\varphi(g)\, d\mu(g).
\end{aligned}$$

It follows that

$$\int \psi(h^{-1})\, d\nu(h) = \int \psi(g) F_\varphi(g)\, d\mu(g)$$

for any $\psi \in C_0(G)$. Since left side is independent of φ, we conclude that for $\varphi, \varphi' \in C_0(G)$ such that $\mu(\varphi) \neq 0$ and $\mu(\varphi') \neq 0$, we have

$$\int \psi(g) F_\varphi(g)\, d\mu(g) = \int \psi(g) F_{\varphi'}(g)\, d\mu(g).$$

Therefore, we have

$$\int \psi(g)(F_\varphi(g) - F_{\varphi'}(g))\, d\mu(g) = 0$$

for any $\psi \in C_0(G)$. Hence, the measure $(F_\varphi - F_{\varphi'})\mu$ is equal to zero. This is possible only if the set $S = \{g \in G \mid (F_\varphi - F_{\varphi'})(g) \neq 0\}$ is a set of measure zero with respect to μ. On the other hand, since $F_\varphi - F_{\varphi'}$ is continuous, the set S is open. It follows that this set must be empty, i.e., $F_{\varphi'} = F_\varphi$.

Hence, the function F_φ is independent of φ, and we can denote it by F. From its definition we get

$$F(1) \int \varphi(g)\, d\mu(g) = \int \varphi(g)\, d\nu(g)$$

for any $\varphi \in C_0(G)$ such that $\mu(\varphi) \neq 0$. The complement of $\{\varphi \in C_0(G) \mid \mu(\varphi) = 0\}$ spans the space $C_0(G)$. Therefore, the above identity holds on $C_0(G)$, i.e., $\nu = F(1)\mu$. This proves the part (ii) of the theorem. □

3.2. Modular function. Let G be a Lie group and μ a left Haar measure on G. Let τ be an automorphism of the Lie group G. Then $\nu_\tau : \varphi \longmapsto \int \varphi(\tau(g))\, d\mu(g)$ is a positive measure on G. In addition, for any $\varphi \in C_o(G)$, we have

$$\begin{aligned}\int \varphi(hg)\, d\nu_\tau(g) = \int \varphi(h\tau(g))\, d\mu(g) &= \int \varphi(\tau(\tau^{-1}(h)g))\, d\mu(g)\\ &= \int \varphi(\tau(g))\, d\mu(g) = \int \varphi(g)\, d\nu_\tau(g),\end{aligned}$$

i.e., the measure ν_τ is left invariant. Therefore, there exists a positive number $\mathrm{mod}(\tau)$ such that $\mathrm{mod}(\tau)\nu_\tau = \mu$, i.e.,

$$\mathrm{mod}(\tau) \int \varphi(\tau(g))\, d\mu(g) = \int \varphi(g)\, d\mu(g).$$

for all $\varphi \in C_0(G)$. Equivalently, we have

$$\mu(\tau(S)) = \mathrm{mod}(\tau)\mu(S)$$

for any measurable set S in G.

3.2.1. Lemma. *The function* mod *is a homomorphism of the group* $\mathrm{Aut}(G)$ *of automorphisms of* G *into the multiplicative group* $\mathbb{R}_+^*$ *of positive real numbers.*

Proof. Let $\sigma, \tau \in \mathrm{Aut}(G)$. Then, for any measurable set S in G, we have

$$\begin{aligned}\mathrm{mod}(\sigma \circ \tau)\mu(S) = \mu((\sigma \circ \tau)(S)) &= \mu(\sigma(\tau(S)))\\ &= \mathrm{mod}(\sigma)\mu(\tau(S)) = \mathrm{mod}(\sigma)\,\mathrm{mod}(\tau)\mu(S),\end{aligned}$$

i.e.,

$$\mathrm{mod}(\sigma \circ \tau) = \mathrm{mod}(\sigma)\,\mathrm{mod}(\tau). \quad \square$$

Clearly $\mathrm{Int} : G \longrightarrow \mathrm{Aut}(G)$ is a group homomorphism. Therefore, by composition with mod we get the group homomorphism $\mathrm{mod} \circ \mathrm{Int}$ of G into $\mathbb{R}_+^*$. Clearly, the function Δ, defined by

$$\Delta(g) = \Delta_G(g) = \mathrm{mod}(\mathrm{Int}(g))^{-1}$$

from G into $\mathbb{R}_+^*$, is a group homomorphism. It is called the *modular function* of G. By the above formulas, we have

$$\int \varphi(hg^{-1})\,d\mu(h) = \int \varphi(ghg^{-1})\,d\mu(h) = \Delta(g)\int \varphi(h)\,d\mu(g)$$

for any $\varphi \in C_0(G)$. Equivalently,

$$\mu(Sg) = \Delta(g)\mu(S)$$

for any $g \in G$ and measurable set S in G. Therefore, a left Haar measure is right invariant if and only if $\Delta_G = 1$.

3.2.2. Proposition. *Let G be a Lie group. Then:*

(i) *The modular function $\Delta : G \longrightarrow \mathbb{R}_+^*$ is a Lie group homomorphism.*

(ii) *For any $g \in G$, we have*

$$\Delta_G(g) = |\det \mathrm{Ad}(g)|^{-1}.$$

Proof. Let $n = \dim G$. Let ω be a nonzero left invariant differential n-form on G. Then ω is completely determined by its value at 1. Clearly,

$$(\mathrm{Int}(g) \circ \gamma(h))(k) = ghkg^{-1} = (\gamma(ghg^{-1}) \circ \mathrm{Int}(g))(k)$$

for any $k \in G$. Hence, for any $h \in G$, we have

$$\begin{aligned}\gamma(g)^*(\mathrm{Int}(h)^*\omega) = (\mathrm{Int}(h) \circ \gamma(g))^*\omega = (\gamma(hgh^{-1}) \circ \mathrm{Int}(h))^*\omega \\ = \mathrm{Int}(h)^*(\gamma(hgh^{-1})^*\omega) = \mathrm{Int}(h)^*\omega\end{aligned}$$

for all $g \in G$, i.e., $\mathrm{Int}(h)\omega$ is a left invariant differential n-form on G. Therefore, it must be proportional to ω.

On the other hand,

$$\begin{aligned}(\mathrm{Int}(g)^*\omega)(\xi_1 \wedge \xi_2 \wedge \cdots \wedge \xi_n) = \omega(T_1(\mathrm{Int}(g))\xi_1 \wedge T_1(\mathrm{Int}(g))\xi_2 \wedge \cdots \wedge T_1(\mathrm{Int}(g))\xi_n) \\ = \det(T_1(\mathrm{Int}(g)))\omega(\xi_1 \wedge \xi_2 \wedge \cdots \wedge \xi_n)\end{aligned}$$

for any $\xi_1, \xi_2, \ldots, \xi_n \in T_1(G)$. It follows that

$$\mathrm{Int}(g)^*\omega = \det(T_1(\mathrm{Int}(g))\omega = \det(\mathrm{Ad}(g))\omega$$

for any $g \in G$. Therefore, we have

$$|\mathrm{Int}(g)^*\omega| = |\det \mathrm{Ad}(g)| \cdot |\omega|$$

for any $g \in G$.

Let μ be the left Haar measure attached to ω. Then, by 1.4.2.1, for any $\varphi \in C_0(G)$, we have

$$\begin{aligned}\int \varphi(h)\,d\mu(h) = \int \varphi\,|\omega| = \int (\varphi\circ\mathrm{Int}(g))\,|\mathrm{Int}(g)^*\omega| = |\det \mathrm{Ad}(g)| \int (\varphi\circ\mathrm{Int}(g))\,|\omega| \\ = |\det \mathrm{Ad}(g)| \int \varphi(ghg^{-1})\,d\mu(h) = |\det \mathrm{Ad}(g)|\,\Delta(g)\int \varphi(h)\,d\mu(h).\end{aligned}$$

Hence, (ii) follows.

From (ii) it follows that Δ is differentiable. This establishes (i). □

A Lie group G is called *unimodular* if $\Delta_G = 1$. As we remarked above, a left Haar measure on a unimodular Lie group is also right invariant, i.e., it is *biinvariant*.

Clearly, abelian Lie groups are unimodular. In addition, we have the following result.

3.2.3. PROPOSITION. *Let G be a compact Lie group. Then G is unimodular.*

PROOF. If G is compact, the image $\Delta(G)$ of G is a compact subgroup of $\mathbb{R}_+^*$. Therefore, it must be equal to $\{1\}$. □

3.2.4. EXAMPLE. Let G be the Lie group of affine transformations of the line studied in 2.10. We established there that

$$\operatorname{Ad}(a,b) = \begin{pmatrix} b & a \\ 0 & 1 \end{pmatrix}$$

for any $(a,b) \in G$. Therefore, we have

$$\Delta(a,b) = |\det \operatorname{Ad}(a,b)|^{-1} = |b|.$$

It follows that G is not unimodular.

3.3. Volume of compact Lie groups. In this section we prove the following characterization of compact Lie groups.

3.3.1. THEOREM. *Let G be a Lie group and $\mu(G)$ a left Haar measure on G. Then, the following conditions are equivalent:*

(i) *The group G is compact.*
(ii) *$\mu(G)$ is finite.*

PROOF. If G is compact, $\mu(G) < \infty$.

Assume that $\mu(G) < \infty$. Let V be a compact neighborhood of 1 in G. Then $\mu(V) > 0$.

Let $\mathcal{S}$ be the family of finite sets $\{g_1, g_2, \dots, g_n\}$ such that $g_i V \cap g_j V = \emptyset$ for all $i \neq j$, $1 \leq i, j \leq n$. Then we have

$$\mu(G) \geq \mu\left(\bigcup_{i=1}^{n} g_i V\right) = \sum_{i=1}^{n} \mu(g_i V) = n\mu(V),$$

and $n \leq \frac{\mu(G)}{\mu(V)}$. It follows that the elements of $\mathcal{S}$ have bounded cardinality. In particular, there exist elements in S of maximal cardinality n_0. Let $\{g_1, g_2, \dots, g_{n_0}\}$ be such element in $\mathcal{S}$. Let $g \in G$. Then $\{g, g_1, \dots, g_{n_0}\}$ is not in $\mathcal{S}$. Therefore, there exists $1 \leq i \leq n_0$ such that $gV \cap g_i V \neq \emptyset$. Therefore, $g \in g_i V V^{-1}$. Since g was completely arbitrary, it follows that $G = \bigcup_{i=1}^{n_0} g_i V V^{-1}$. Hence G is a union of compact sets, i.e., G is compact. □

CHAPTER 3

Compact Lie groups

1. Compact Lie groups

1.1. Lie algebra of a compact Lie group. Let G be a compact Lie group and $L(G)$ its Lie algebra. By 2.3.1.1, there exists a left Haar measure on G. By 2.3.2.3, this measure is biinvariant. Moreover, by 2.3.3.1, the volume of G is finite. therefore, we can select the biinvariant Haar measure μ on G such that $\mu(G) = 1$.

1.1.1. LEMMA. *Let G be a compact group.*

(i) *The Lie algebra $L(G)$ admits an inner product such that the image of* $\mathrm{Ad} : G \longrightarrow \mathrm{GL}(L(G))$ *is a closed subgroup of* $\mathrm{O}(L(G))$.

(ii) *With respect to this inner product,* $\mathrm{ad}(\xi)$, $\xi \in L(G)$, *are skew symmetric linear transformations.*

PROOF. Since G is compact, the image $\mathrm{Ad}(G) \subset \mathrm{GL}(L(G))$ is compact and therefore closed.

Let $(\xi, \eta) \longmapsto (\xi|\eta)$ be an arbitrary inner product on $L(G)$. Then we define another inner product on $L(G)$ by

$$[\xi|\eta] = \int_G (\mathrm{Ad}(g)\xi|\,\mathrm{Ad}(g)\eta)\, d\mu(g),$$

for $\xi, \eta \in L(G)$ Clearly, we have

$$\begin{aligned}[\mathrm{Ad}(g)\xi|\,\mathrm{Ad}(g)\eta] &= \int_G (\mathrm{Ad}(h)\,\mathrm{Ad}(g)\xi|\,\mathrm{Ad}(h)\,\mathrm{Ad}(g)\eta)\, d\mu(g) \\ &= \int_G (\mathrm{Ad}(hg)\xi|\,\mathrm{Ad}(hg)\eta)\, d\mu(g) = \int_G (\mathrm{Ad}(g)\xi|\,\mathrm{Ad}(g)\eta)\, d\mu(g) = [\xi|\eta]\end{aligned}$$

for all $\xi, \eta \in L(G)$ and $g \in G$. Therefore, $\mathrm{Ad}(g) \in \mathrm{O}(L(G))$ for all $g \in G$. This proves (i).

(ii) follows immediately the description of Lie algebra of the orthogonal group in 2.1.8. □

1.2. Tori in compact Lie groups. By 2.2.6.2 a compact connected abelian n-dimensional Lie group is isomorphic to a torus $\mathbb{T}^n$. Therefore, we are going to call it a *torus*.

Let G be a compact Lie group and T a torus in G. Then the Lie algebra $L(T)$ of T is an abelian Lie subalgebra of $L(G)$.

We consider the set of all subgroups of G and the set of all Lie subalgebras of $L(G)$ equipped with the partial ordering given by inclusion.

1.2.1. LEMMA. *Let G be a compact Lie group.*

(i) *Any abelian Lie subalgebra of $L(G)$ is contained in a maximal abelian Lie subalgebra.*

(ii) *Any torus in G is contained in a maximal torus.*
(iii) *An integral subgroup T is a maximal torus in G if and only if $L(T)$ is a maximal abelian Lie subalgebra of $L(G)$.*
(iv) *the map $T \longmapsto L(T)$ is a bijection between maximal tori in G and maximal abelian Lie subalgebras in $L(G)$.*

Proof. (i) is obvious, since $L(G)$ is finite-dimensional.

Let $\mathfrak{h}$ be an abelian Lie subalgebra of $L(G)$. Denote by H the integral subgroup of G corresponding to $\mathfrak{h}$. Then the closure $\bar{H}$ of H is a compact connected abelian subgroup of G. By Cartan's theorem 2.2.11.1, it is a torus in G. Hence, its Lie algebra $L(\bar{H})$ is an abelian Lie subalgebra of $L(G)$ containing $L(H)$.

If $\mathfrak{h}$ is a maximal abelian Lie subalgebra of $L(G)$, $L(\bar{H}) = \mathfrak{h}$, i.e., $H = \bar{H}$ by 2.2.7.6. It follows that H is a torus. Assume that H' is a torus containing H. Then its Lie algebra $L(H')$ is an abelian Lie subalgebra of $L(G)$ and $L(H') \supset \mathfrak{h}$. By the maximality of $\mathfrak{h}$, it follows that $L(H') = \mathfrak{h}$, and $H' = H$ by 2.2.7.6. Therefore, H is a maximal torus in G.

It follows that the bijection from Lie subalgebras into integral subgroups maps maximal abelian Lie subalgebras into maximal tori.

If T is a maximal torus in G, its Lie algebra $L(T)$ is contained in a maximal abelian Lie subalgebra $\mathfrak{h}$. The maximal torus H corresponding to $\mathfrak{h}$ must contain T by 2.2.7.6, hence $T = H$ and $L(T) = \mathfrak{h}$ is a maximal abelian Lie algebra. This completes the proof of (iii) and (iv).

Let T be a torus in G. By (i), its Lie algebra $L(T)$ is contained in a maximal abelian Lie subalgebra $\mathfrak{h}$ of $L(G)$. The corresponding integral subgroup H is a maximal torus in G, and by 2.2.7.6, $T \subset H$. This proves (ii). □

Let G be a compact Lie group and T a torus in G. For any $g \in G$, $\operatorname{Int}(g)(T) = gTg^{-1}$ is a torus in G, i.e., $\operatorname{Int}(g)$ permutes tori in G. Clearly, this action preserves the inclusion relations, therefore $\operatorname{Int}(g)$ permutes maximal tori in G. Hence, G acts by inner automorphisms on the set of all maximal tori in G.

Analogously, for any abelian Lie subalgebra $\mathfrak{h}$ of $L(G)$, the Lie algebra $\operatorname{Ad}(g)(\mathfrak{h})$ is also an abelian Lie subalgebra. Therefore, $\operatorname{Ad}(g)$ permutes all abelian Lie subalgebras in $L(G)$. Since this action also preserves the inclusion relations, $\operatorname{Ad}(g)$ also permutes all maximal abelian Lie subalgebras of $L(G)$.

1.2.2. Theorem. *Let G be a compact Lie group. Then*

(i) *The group G acts transitively on the set of all maximal tori in G, i.e., all maximal tori are conjugate.*
(ii) *The group G acts transitively on the set of all maximal abelian Lie subalgebras in $L(G)$, i.e., all maximal abelian Lie subalgebras are conjugate.*

By 1.2.1, the statements (i) and (ii) are equivalent.

This implies that all maximal tori in G have same dimension. Also, all maximal abelian Lie subalgebras in G have same dimension. Finally, by 1.2.1, these two numbers are equal. This number is called the *rank* of G.

The proof of the theorem is based on the following lemma.

1.2.3. Lemma. *Let G be a compact Lie group. Let $\xi, \eta \in L(G)$. Then there exists $g \in G$ such that $[\operatorname{Ad}(g)\xi, \eta] = 0$.*

PROOF. By 1.1.1, $L(G)$ admits an $\mathrm{Ad}(G)$-invariant inner product. Consider the function

$$G \ni g \longmapsto F(g) = (\mathrm{Ad}(g)\xi|\eta).$$

Clearly, this is a smooth function on G. Since G is compact, F must have a stationary point in G. If g_0 is a stationary point of F, the function $t \longmapsto F(\exp(t\zeta)g_0)$ has a stationary point at $t = 0$ for any $\zeta \in L(G)$. On the other hand, we have

$$\begin{aligned} F(\exp(t\zeta)g_0) &= (\mathrm{Ad}(\exp(t\zeta)g_0)\xi|\eta) \\ &= (\mathrm{Ad}(\exp(t\zeta))\,\mathrm{Ad}(g_0)\xi|\eta) = \left(e^{t\,\mathrm{ad}(\zeta)}\,\mathrm{Ad}(g_0)\xi\middle|\eta\right) \end{aligned}$$

by 2.2.9.4. Therefore, since $\mathrm{ad}(\zeta)$ and $\mathrm{ad}(\eta)$ are antisymmetric by 1.1.1, we have

$$\begin{aligned} 0 = \left.\frac{dF(\exp(t\zeta)g_0)}{dt}\right|_{t=0} &= (\mathrm{ad}(\zeta)\,\mathrm{Ad}(g_0)\xi|\eta) = -(\mathrm{Ad}(g_0)\xi|\,\mathrm{ad}(\zeta)\eta) \\ &= (\mathrm{Ad}(g_0)\xi|\,\mathrm{ad}(\eta)\zeta) = -(\mathrm{ad}(\eta)\,\mathrm{Ad}(g_0)\xi|\zeta) = ([\mathrm{Ad}(g_0)\xi,\eta]|\zeta) \end{aligned}$$

for all $\zeta \in L(G)$. It follows that $[\mathrm{Ad}(g_0)\xi, \eta] = 0$. □

Now we can prove 1.2.2. Let T and T' be two maximal tori in G. Let $L(T)$ and $L(T')$ be their Lie algebras. Then, by 2.2.13.2, there exist $\xi \in L(T)$ and $\eta \in L(T')$ such that the corresponding one-parameter subgroups are dense in T, resp. T'. By 1.2.3, There exists $g \in G$ such that $[\mathrm{Ad}(g)\xi, \eta] = 0$. Therefore, $\mathrm{Ad}(g)\xi$ and η span an abelian Lie subalgebra. Moreover, by 2.2.9.3, $\exp(t\,\mathrm{Ad}(g)\xi)$ and $\exp(s\eta)$ are in the corresponding integral subgroup H for all $t, s \in \mathbb{R}$. By 2.2.2.18, H is an abelian Lie group.

It follows that

$$\exp(t\,\mathrm{Ad}(g)\xi)\exp(s\eta) = \exp(s\eta)\exp(t\,\mathrm{Ad}(g)\xi)$$

for all $t, s \in \mathbb{R}$. Therefore, by 2.2.9.4, we have

$$g\exp(t\xi)g^{-1}\exp(s\eta) = \exp(s\eta)g\exp(t\xi)g^{-1}$$

for all $t, s \in \mathbb{R}$. Since one-parameter subgroups corresponding to ξ and η are dense in T, resp. T', by continuity we have

$$gtg^{-1}t' = t'gtg^{-1}$$

for all $t \in T$ and $t' \in T'$. Clearly, $T_g = gTg^{-1}$ is a maximal torus in G, and its elements commute with elements of T'. The differentiable map $\nu : T' \times T_g \longrightarrow G$ given by $\nu(t, s) = ts$, $t \in T'$, $s \in T_g$, is a Lie group morphism. Therefore, its image S is a connected compact abelian subgroup of G. By Cartan's theorem, 2.2.11.1, S is a torus in G. Since S contains the maximal tori T' and T_g, we see that $T' = S = T_g$. Hence, $gTg^{-1} = T'$, and (i) in 1.2.2 follows.

1.3. Surjectivity of the exponential map. In this section we prove the following basic result.

1.3.1. THEOREM. *Let G be a connected compact Lie group. The exponential map* $\exp : L(G) \longrightarrow G$ *is surjective.*

Let G be a connected compact Lie group and T a maximal torus in G. We claim first that the following two statements are equivalent

(i) the exponential map $\exp : L(G) \longrightarrow G$ is surjective.

(ii) every element of G lies in a conjugate of T, i.e., the map $\varphi : G \times T \longrightarrow G$ given by $\varphi(g,t) = gtg^{-1}$ is surjective.

If (i) holds, for any $g \in G$ we have $g = \exp(\xi)$ for some $\xi \in L(G)$. Therefore, g lies in the one-parameter subgroup $\{\exp(t\xi) \mid t \in \mathbb{R}\}$. The closure of this one-parameter subgroup is a torus in G. Therefore, by 1.2.1, it is contained in a maximal torus T' in G. By 1.2.2, $T' = hTh^{-1}$ for some $h \in G$. Therefore, $g \in T' = hTh^{-1}$ and $g = \varphi(h,t)$ for some $t \in T$.

On the other hand, if (ii) holds, any $g \in G$ is of the form $g = hth^{-1}$ for some $h \in G$ and $t \in T$. Therefore, g is in the maximal torus hTh^{-1} in G. By 2.2.9.6, g is in the image of the exponential map.

It follows that to prove 1.3.1 it is enough to establish (ii).

Let $X = \varphi(G \times T)$. Then X is a nonempty compact subset of G. Since G is connected to prove that X is equal to G it is enough to prove that X is open. Since X is invariant under conjugation by elements of G, it is enough to show that X is a neighborhood of any $t \in T$.

We prove the statement by induction in $\dim G - \dim T \geq 0$. If $\dim G - \dim T = 0$, we have $\dim G = \dim T$ and $G = T$ since G is connected. In this case the assertion is evident.

Therefore, we can assume that $\dim G - \dim T > 0$.

Let $t \in T$. Let H be the centralizer of t in G, i.e., $H = \{g \in G \mid gt = tg\}$. Then, we have $H = \{g \in G \mid \mathrm{Int}(t)g = g\}$. By 2.2.2.15, it follows that H is a Lie subgroup and

$$L(H) = \{\xi \in L(G) \mid \mathrm{Ad}(t)\xi = \xi\}.$$

Clearly, H is a compact Lie subgroup. Let H_0 be the identity component of H. Then $T \subset H_0$, and H_0 is a compact connected Lie group. Evidently, T is a maximal torus in H_0.

Clearly, there are two possibilities: either t is in the center Z of G or t is not in the center of G.

Assume first that $t \in Z$. Let T' be a maximal torus in G. Then, by 1.2.2, we have $T' = hTh^{-1}$ for some $h \in G$. Therefore, $t = hth^{-1} \in T'$. It follows that t is contained in all maximal tori in G. Let $\xi \in L(G)$. Then, ξ is in some maximal abelian Lie subalgebra of $L(G)$, and by 1.2.1, $\exp(\xi)$ is in the corresponding maximal torus T''. Since $t \in T''$, we conclude that $t \cdot \exp(\xi) \in T''$. By 1.2.2, there exists $k \in G$ such that $T'' = kTk^{-1}$, hence it follows that $t\exp(\xi) \in kTk^{-1} \subset X$. Since the exponential map is a local diffeomorphism at 0 by 2.2.9.1, we conclude that $\{t\exp(\xi) \mid \xi \in L(G)\}$ is a neighborhood of t in G.

It remains to treat the case $t \notin Z$. In this case, by 2.2.2.17, we have $\mathrm{Ad}(t) \neq 1_{L(G)}$. It follows that $L(H) \neq L(G)$. In particular, we have $\dim H_0 = \dim H < \dim G$. Hence, we have $\dim H_0 - \dim T < \dim G - \dim T$. By the induction assumption, we have

$$H_0 = \{ht'h^{-1} \mid h \in H_0, t' \in T\}.$$

Therefore, we have

$$X = \{ghg^{-1} \mid g \in G, h \in H_0\}.$$

Hence, to prove that X is a neighborhood of t it is enough to show that the map $\psi : G \times H_0 \longrightarrow G$ defined by $\psi(g,h) = ghg^{-1}$ is a submersion at $(1,t)$. Since the exponential map is a local diffeomorphism at 0 by 2.2.9.1, it is enough to show that the map

$$L(G) \times L(H) \ni (\xi, \eta) \longmapsto \psi(\exp\xi, t\exp\eta)$$

is a submersion at $(0,0)$. This in turn is equivalent to

$$L(G) \times L(H) \ni (\xi, \eta) \longmapsto t^{-1}\psi(\exp\xi, t\exp\eta)$$

being a submersion at $(0,0)$. On the other hand, by 2.2.9.4, we have

$$t^{-1}\psi(\exp\xi, t\exp\eta) = t^{-1}\exp(\xi)t\exp(\eta)\exp(-\xi) = \exp(\mathrm{Ad}(t^{-1})\xi)\exp(\eta)\exp(-\xi).$$

Therefore, the differential of this map at $(0,0)$ is

$$\alpha : (\xi, \eta) \longmapsto \mathrm{Ad}(t^{-1})\xi + \eta - \xi = (\mathrm{Ad}(t^{-1}) - I)\xi + \eta.$$

As we remarked in 1.1.1, there exists an inner product on G such that $\mathrm{Ad}(t^{-1})$ is an orthogonal transformation. Therefore, $L(H)^\perp$ is invariant for $\mathrm{Ad}(t^{-1})$. Hence, it is invariant for $\mathrm{Ad}(t^{-1}) - I$ too. Let $\xi \in L(H)^\perp$ be in the kernel of $\mathrm{Ad}(t^{-1}) - I$. Then, as we remarked before, ξ is in $L(H)$ too. It follows that $\xi \in L(H) \cap L(H)^\perp = \{0\}$. Therefore, $\mathrm{Ad}(t^{-1}) - I$ induces an isomorphism of $L(H)^\perp$. It follows that $\alpha(L(G) \times L(H)) \supset L(H)^\perp \oplus L(H) = L(G)$. Hence, ψ is a submersion at $(1,t)$. This completes the proof of the induction step, and finishes the proof of 1.3.1.

As we mentioned at the beginning of the proof of 1.3.1, this also establishes the following result.

1.3.2. Corollary. *Let G be a connected compact Lie group and T a maximal torus in G. Then the differentiable map $\varphi : G \times T \longrightarrow G$ given by $\varphi(g,t) = gtg^{-1}$ is surjective.*

1.3.3. Corollary. *Let G be a connected compact Lie group. Then any $g \in G$ lies in a maximal torus.*

Proof. By 1.3.2, for $g \in G$, there exists $h \in G$ and $t \in T$ such that $g = hth^{-1}$. It follows that g is in the maximal torus hTh^{-1}. □

1.4. Centralizers of tori.

1.4.1. Theorem. *Let G be a connected compact Lie group and T a torus in G. Let*

$$C = \{g \in G \mid gt = tg \text{ for all } t \in T\}$$

be the centralizer of T. Then:

(i) *C is connected Lie subgroup of G containing T.*
(ii) *If, in addition, T is a maximal torus in G, we have $C = T$.*

Proof. Clearly, the centralizer of T is a compact subgroup of G containing T. Hence, by Cartan's theorem, 2.2.11.1, it follows that C is a compact Lie subgroup of G.

Let $t \in C$. First we show that t and T lie in a torus in G. Let H be the centralizer of t in G. By 1.3.3, t is in a maximal torus T' in G. Clearly, $T' \subset H$. Therefore, T' is in the connected component H_0 of H. In particular, this implies that $t \in H_0$. Hence, t and T are in H_0. Since H_0 is a compact Lie group, by 1.2.1, T is contained in a maximal torus S in H_0. By 1.3.2, there exists $h \in H_0$ such that $t \in hSh^{-1}$. It follows that $t = h^{-1}th \in S$. Hence, the torus S satisfies the above requirement.

If T is a maximal torus in G, we must have $S = T$. This implies that $t \in T$. Since $t \in C$ was arbitrary, it follows that $C = T$ in that case. This completes the proof of (ii).

In general situation, S is in C. Since S is connected, it is in the identity component C_0 of C. Hence, t is in C_0. Since $t \in C$ was arbitrary, it follows that $C_0 = C$. This proves (i). □

1.4.2. Corollary. *Let G be a connected compact Lie group and Z the center of G. Then Z is equal to the intersection of all maximal tori in G.*

Proof. Let T be a maximal torus in G and $z \in Z$. Then z centralizes T, and $z \in T$ by 1.4.1.(ii). Hence, we have $Z \subset T$. It follows that Z is contained in the intersection of all maximal tori in G.

Conversely, let z be in the intersection of all maximal tori in G. Let $g \in G$. By 1.3.3, there exists a maximal torus T in G which contains g. Hence, z commutes with g. It follows that z is in the center Z of G. □

1.5. Normalizers of maximal tori. Let T be an n-dimensional torus and let $\tilde{T}$ be its universal covering group. Then, by 2.2.9.6, $\tilde{T}$ can be identified with $\mathbb{R}^n$ and T with $\mathbb{R}^n/\mathbb{Z}^n$. The projection map $\tilde{T} \longrightarrow T$ corresponds to the natural projection $\mathbb{R}^n \longrightarrow \mathbb{R}^n/\mathbb{Z}^n = \mathbb{T}^n$. Since the exponential map on $\mathbb{R}^n$ is the identity, we can also identify $L(T)$ with $\mathbb{R}^n$ and the covering map $\mathbb{R}^n \longrightarrow \mathbb{T}^n$ corresponds to the exponential map. Let α be an automorphism of T, Then $L(\alpha)$ is an automorphism of $L(T)$, i.e., $L(\alpha) \in GL(L(T))$. By 2.2.9.3, $\alpha \circ \exp = \exp \circ L(\alpha)$. Hence, the action of α on T is induced by the action of $L(\alpha)$ on $L(T)$. This implies that $L(\alpha)$ must map the lattice $\ker \exp$ into itself. Since the same argument applies to α^{-1}, it follows that $L(\alpha)$ is a bijection of $\ker \exp$.

With our identification, $L(\alpha)$ corresponds to an element of $\mathrm{GL}(n, \mathbb{Z})$, the subgroup of $\mathrm{GL}(n, \mathbb{R})$ consisting of all matrices which map $\mathbb{Z}^n$ onto itself. A matrix A is in $\mathrm{GL}(n, \mathbb{Z})$ if and only if A and A^{-1} are in $M_n(\mathbb{Z})$, i.e., their matrix entries are integers. This is equivalent to $A \in M_n(\mathbb{Z})$ and $\det A = \pm 1$. Clearly, $\mathrm{GL}(n, \mathbb{Z})$ is a discrete subgroup of $\mathrm{GL}(n, \mathbb{R})$.

Let T be a torus in a compact group G. Let g be an element of G which normalizes T, i.e., such that $gTg^{-1} = T$. Then, $t \longmapsto gtg^{-1}$ is an automorphism of T.

1.5.1. Lemma. *Let T be a torus in a connected compact Lie group G. Let $N = \{g \in G \mid gTg^{-1} = T\}$ be the normalizer of T. Then N is a Lie subgroup of G and its identity component is the centralizer of T.*

Proof. Clearly, N is a closed subgroup of G. Therefore, G is a Lie subgroup by 2.2.11.1. Let N_0 be the identity component of N. Let C be the centralizer of T. Then $C \subset N$. Moreover, by 1.4.1.(i), we see that $C \subset N_0$.

Let $n \in N$. Hence, $\mathrm{Int}(n)$ induces an automorphism of T. Therefore, its differential $\mathrm{Ad}(n)|_{L(T)}$ is in a discrete subgroup of $\mathrm{GL}(L(T))$. It follows that $n \longmapsto \mathrm{Ad}(n)\big|_{L(T)}$ is a Lie group homomorphism of N into a discrete subgroup of $\mathrm{GL}(L(T))$. Therefore, the identity component N_0 of N maps into the identity, i.e., $\mathrm{Ad}(n)\big|_{L(T)} = I$ for any $n \in N_0$. Moreover, by 2.2.2.16, it follows that $\mathrm{Int}(n)\big|_T = id_T$ for any $n \in N_0$, i.e., $N_0 \subset C$. This in turn implies that $N_0 = C$. □

1.5.2. Theorem. *Let T be a maximal torus in a connected compact Lie group G. Let N be the normalizer of T. Then the identity component N_0 of N is equal to T.*

Moreover N/T is a finite group.

PROOF. By 1.5.1, N_0 is equal to the centralizer of T. By 1.4.1.(ii), the centralizer of T is equal to T. This proves the first statement.

Since N is a compact group, the discrete group N/T is also compact. Therefore, N/T must be finite. □

The group $W = N/T$ is called the *Weyl group* of the pair (G, T).

1.5.3. EXAMPLE. Let $G = \mathrm{SU}(2)$. Then G is a simply connected, connected compact Lie group. Let T be the subgroup consisting of diagonal matrices in G, i.e.,

$$T = \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{pmatrix} \,\middle|\, |\alpha| = 1 \right\}.$$

Clearly, T is a one-dimensional torus in G. An element $g \in G$ can be written as

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \text{ with } |a|^2 + |b|^2 = 1.$$

Therefore, we have

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{pmatrix} \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}^{-1} = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{pmatrix} \begin{pmatrix} \bar{a} & -b \\ \bar{b} & a \end{pmatrix}$$
$$= \begin{pmatrix} \alpha|a|^2 + \bar{\alpha}|b|^2 & (\bar{\alpha} - \alpha)ab \\ (\bar{\alpha} - \alpha)\bar{a}\bar{b} & \bar{\alpha}|a|^2 + \alpha|b|^2 \end{pmatrix}$$

for any α, $|\alpha| = 1$. If g is in the normalizer of T, we must have $ab = 0$. Therefore, either $a = 0$ or $b = 0$. Clearly, $b = 0$ implies that $g \in T$. On the other hand, if $a = 0$, we have

$$g = \begin{pmatrix} 0 & -b \\ \bar{b} & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \bar{b} & 0 \\ 0 & b \end{pmatrix}$$

with $|b| = 1$. Therefore, we have

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

It follows that the connected component of N is equal to T. Hence, T is a maximal torus in G, and the rank of G is equal to 1. On the other hand, the Weyl group of (G, T) is isomorphic to the two-element group. The nontrivial element of W is represented by $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$.

1.6. Universal covering groups of connected compact Lie groups. Let G be a connected compact Lie group. We want to describe the structure of the universal covering group $\tilde{G}$ of G.

We start with some technical preparation.

1.6.1. LEMMA. *Let G be a connected Lie group and C a discrete central subgroup of G such that G/C is compact. Then there exists a compact neighborhood D of* 1 *in G such that*

$$\mathrm{int}(D) \cdot C = G.$$

PROOF. Let U be an open neighborhood of 1 in G such that its closure $\bar{U}$ is compact. Since the natural projection $p : G \longrightarrow G/C$ is open, $p(U)$ is an open neighborhood of 1 in G/C. Therefore, the translates $\gamma(k)(p(U))$, $k \in G/C$, form an open cover of G/C. Since G/C is compact, there exist $k_1 = 1, k_2, \dots, k_p \in G/C$ such that

$$G/C = \bigcup_{i=1}^{p} \gamma(k_i)(p(U)).$$

Let $g_1 = 1, g_2, \dots, g_p \in G$ be such that $p(g_i) = k_i$ for $1 \le i \le p$. Then

$$D = \bigcup_{i=1}^{p} g_i \bar{U}$$

is a compact set in G. In addition,

$$p(\operatorname{int}(D)) \supset p\left(\bigcup_{i=1}^{p} g_i U\right) = G/C,$$

hence we have $\operatorname{int}(D) \cdot C = G$. □

1.6.2. Corollary. *Let G be a connected Lie group and C a discrete central subgroup of G such that G/C is compact. The group G is compact if and only if C is finite.*

Proof. Since C is a discrete subgroup, if G is compact, C must be finite.

Conversely, if C is finite, by 1.6.1, G is a union of finitely many compact sets. Therefore, G is compact. □

1.6.3. Lemma. *Let G be a connected Lie group and C a discrete central subgroup of G such that G/C is compact. Then C is finitely generated.*

Proof. By 1.6.1, there exists a compact neighborhood D of 1 in G such that the translates $\gamma(c)\operatorname{int}(D)$ cover G. Since D^2 is a compact set in G, it is covered by finitely many such translates, i.e.,

$$D^2 \subset Dc_1 \cup Dc_2 \cup \cdots \cup Dc_m$$

for some $c_1, c_2, \dots, c_m \in C$. Let Γ be the subgroup of C generated by $c_1, c_2, \dots, c_m$. Then, as we remarked, $D^2 \subset D \cdot \Gamma$. We claim that $D^n \subset D \cdot \Gamma$. We prove this statement by induction in n. Assume that the statement holds for n. Then

$$D^{n+1} = D \cdot D^n \subset D^2 \cdot \Gamma \subset D \cdot \Gamma.$$

Since G is connected, by 2.1.5.1 we have

$$G = \bigcup_{n=1}^{\infty} D^n \subset D \cdot \Gamma.$$

Therefore, every element of $c \in C$ is of the form $c = db$ with $d \in D$ and $b \in \Gamma$. This implies that $d \in D \cap C$. Hence, C is generated by $D \cap C$ and $c_1, c_2, \dots, c_m$. Since D is compact and C discrete, $D \cap C$ is finite. □

Let G be a connected compact Lie group and $\tilde{G}$ its universal covering group. Let $p : \tilde{G} \longrightarrow G$ be the canonical projection and $C = \ker p$. By the results from 2.1.6 we know that C is isomorphic to the fundamental group of G. Hence we have the following consequence.

1.6.4. Corollary. *The fundamental group of a connected compact Lie group is finitely generated.*

Let G be a connected Lie group and C a discrete central subgroup such that G/C is compact. Denote by $\operatorname{Hom}(G,\mathbb{R})$ the set of Lie group morphisms of G into $\mathbb{R}$, and by $\operatorname{Hom}(C,\mathbb{R})$ the set of all group morphisms of C into $\mathbb{R}$. For $\varphi,\psi \in \operatorname{Hom}(G,\mathbb{R})$ we define $\varphi+\psi : G \longrightarrow \mathbb{R}$ by $(\varphi+\psi)(g) = \varphi(g)+\psi(g)$ for all $g \in G$. For any $g,h \in G$ we have

$$(\varphi+\psi)(gh) = \varphi(gh)+\psi(gh) = \varphi(g)+\varphi(h)+\psi(g)+\psi(h) = (\varphi+\psi)(g)+(\varphi+\psi)(h).$$

Therefore, $\varphi+\psi$ is a Lie group morphism in $\operatorname{Hom}(G,\mathbb{R})$. It follows that $\operatorname{Hom}(G,\mathbb{R})$ is an abelian group. Analogously, $\operatorname{Hom}(C,\mathbb{R})$ is also an abelian group.

Let $\varphi : G \longrightarrow \mathbb{R}$ be a Lie group morphism. Then its restriction to C is a group morphism $\operatorname{res}(\varphi) : C \longrightarrow \mathbb{R}$. Moreover, the restriction map $\operatorname{res} : \operatorname{Hom}(G,\mathbb{R}) \longrightarrow \operatorname{Hom}(C,\mathbb{R})$ is a group homomorphism.

The following observation is elementary.

1.6.5. Lemma. *The restriction map* $\operatorname{res} : \operatorname{Hom}(G,\mathbb{R}) \longrightarrow \operatorname{Hom}(C,\mathbb{R})$ *is an injective homomorphism.*

Proof. Let φ be in the kernel of res. Hence, φ factors through a Lie group morphism $\rho : G/C \longrightarrow \mathbb{R}$. Since G/C is compact, $\rho(G/C)$ is a compact subgroup of $\mathbb{R}$, i.e., $\rho(G/C) = \{0\}$. Therefore $\rho = 0$ and $\varphi = 0$. $\square$

The main result of this section is the following result.

1.6.6. Theorem. *The restriction map* $\operatorname{res} : \operatorname{Hom}(G,\mathbb{R}) \longrightarrow \operatorname{Hom}(C,\mathbb{R})$ *is an isomorphism.*

To prove this theorem, we need to establish the following extension result.

1.6.7. Lemma. *Let G be a connected Lie group and C a discrete central subgroup of G such that G/C is compact. Let $\varphi : C \longrightarrow \mathbb{R}$ be a group homomorphism. Then φ extends to a Lie group homomorphism of G into $\mathbb{R}$.*

Proof. Let D be a compact set satisfying the conditions of 1.6.1. Let r_1 be a positive continuous function on G with compact support such that $r_1|_D = 1$. We put

$$r_2(g) = \sum_{c\in C} r_1(cg)$$

for any $g \in G$. Let U be a compact symmetric neighborhood of 1 in G. Then gU is a neighborhood of $g \in G$. Moreover, $h \longmapsto r_1(ch)$ is zero on gU if $\operatorname{supp}(r_1)\cap cgU = \emptyset$. This is equivalent to $c \notin \operatorname{supp}(r_1)Ug^{-1}$. Since the set $\operatorname{supp}(r_1)Ug^{-1}$ is compact, the function $h \longmapsto r_1(ch)$ is nonzero on gU for finitely many $c \in C$ only. Hence, r_2 is a continuous function on gU.

It follows that r_2 is a continuous function on G constant on C-cosets. Any $g \in G$ can be represented as $g = dc$ with $d \in D$ and $c \in C$. Hence, we have

$$r_2(g) = r_2(cd) = \sum_{c'\in C} r_1(c'cd) = \sum_{c'\in C} r_1(c'd) \geq r_1(d) = 1.$$

Therefore, $r_2(g) > 0$ for any $g \in G$. Hence, we can define

$$r(g) = \frac{r_1(g)}{r_2(g)} \text{ for any } g \in G.$$

This is a positive continuous function on G with compact support. Moreover,

$$\sum_{c\in C} r(cg) = \frac{1}{r_2(g)} \sum_{c\in C} r_1(cg) = 1$$

for any $g \in G$.

Therefore, we constructed a continuous function $r : G \longrightarrow \mathbb{R}$ satisfying

(1) $\operatorname{supp} r$ is compact;
(2) $r(g) \geq 0$ for all $g \in G$;
(3) $\sum_{c\in C} r(cg) = 1$ for any $g \in G$.

Now, define $\psi : G \longrightarrow \mathbb{R}$ by

$$\psi(g) = \sum_{c\in C} \varphi(c) r(c^{-1}g)$$

for any $g \in G$. As before, we conclude that ψ is a continuous function on G and

$$\begin{aligned}\psi(cg) = \sum_{b\in C} \varphi(b) r(b^{-1}cg) &= \sum_{b\in C} \varphi(cb) r(b^{-1}g) = \sum_{b\in C} (\varphi(c) + \varphi(b)) r(b^{-1}g) \\ &= \varphi(c) \sum_{b\in C} r(b^{-1}g) + \sum_{b\in C} \varphi(b) r(b^{-1}g) = \varphi(c) + \psi(g)\end{aligned}$$

for all $c \in C$ and $g \in G$.

Define

$$\Phi(g) = \psi(g) - \psi(1) \text{ for } g \in G.$$

If $g = c$, from the above relations we get

$$\Phi(c) = \psi(c) - \psi(1) = \varphi(c) + \psi(1) - \psi(1) = \varphi(c)$$

for all $c \in C$. Therefore, the function Φ extends φ to G.

Moreover, we have

$$\Phi(cg) = \psi(cg) - \psi(1) = \varphi(c) + \psi(g) - \psi(1) = \Phi(c) + \Phi(g)$$

for all $c \in C$ and $g \in G$.

Now define

$$F(x; g) = \Phi(xg) - \Phi(x) \text{ for } x, g \in G.$$

Then, we have

$$F(x; c) = \Phi(xc) - \Phi(x) = \Phi(c) = \varphi(c)$$

for any $x \in G$ and $c \in C$; and

$$F(x; gg') = \Phi(xgg') - \Phi(x) = \Phi(xgg') - \Phi(xg) + \Phi(xg) - \Phi(x) = F(xg; g') + F(x; g)$$

for all $x, g, g' \in G$. Also, we have

$$F(cx; g) = \Phi(cxg) - \Phi(cx) = \Phi(c) + \Phi(xg) - \Phi(c) - \Phi(x) = \Phi(xg) - \Phi(x) = F(x; g),$$

i.e., $F : G \times G \longrightarrow \mathbb{R}$ factors through $G/C \times G$. Let $\tilde{F}$ be the continuous function from $G/C \times G$ into $\mathbb{R}$ induced by F.

Since G/C is a compact Lie group, it admits a biinvariant Haar measure μ such that $\mu(G/C) = 1$.

Therefore, we can define

$$\Psi(g) = \int_{G/C} \tilde{F}(y; g)\, d\mu(y)$$

for any $g \in G$.

Then we have

$$\Psi(c) = \int_{G/C} \tilde{F}(y; c)\, d\mu(y) = \varphi(c)$$

for all $c \in C$, i.e., Ψ also extends φ to G.

On the other hand, we have

$$\Psi(gg') = \int_{G/C} \tilde{F}(y; gg')\, d\mu(y) = \int_{G/C} (\tilde{F}(yp(g); g') + \tilde{F}(y, g))\, d\mu(y)$$
$$= \int_{G/C} \tilde{F}(y; g')\, d\mu(y) + \int_{G/C} \tilde{F}(y; g)\, d\mu(y) = \Psi(g) + \Psi(g');$$

i.e., $\Psi : G \longrightarrow \mathbb{R}$ is a homomorphism. By 2.2.11.2, Ψ is a Lie group homomorphism. □

Let G be a connected compact Lie group. Then, by 1.1.1, there exists an invariant inner product on $L(G)$, i.e., Ad is a Lie group homomorphism of G into $\mathrm{O}(L(G))$. This implies that ad is a Lie algebra morphism of $L(G)$ into the Lie algebra of $\mathrm{O}(L(G))$, i.e., all linear transformations $\mathrm{ad}(\xi)$, $\xi \in L(G)$, are antisymmetric.

Let $\mathfrak{h}$ be an ideal in $L(G)$. Then it is invariant under all $\mathrm{ad}(\xi)$, $\xi \in L(G)$. This implies that the orthogonal complement $\mathfrak{h}^\perp$ of $\mathfrak{h}$ is invariant for all $\mathrm{ad}(\xi)$, $\xi \in \mathfrak{g}$, i.e., $\mathfrak{h}^\perp$ is an ideal in $\mathfrak{g}$. It follows that $L(G) = \mathfrak{h} \oplus \mathfrak{h}^\perp$ as a linear space. On the other hand, for $\xi \in \mathfrak{h}$ and $\eta \in \mathfrak{h}^\perp$, we have $[\xi, \eta] \in \mathfrak{h} \cap \mathfrak{h}^\perp = \{0\}$, i.e., $L(G)$ is the product of $\mathfrak{h}$ and $\mathfrak{h}^\perp$ as a Lie algebra.

Therefore we established the following result.

1.6.8. LEMMA. *Let G be a connected compact Lie group. Let $\mathfrak{h}$ be an ideal in $L(G)$. Then $L(G)$ is the product of $\mathfrak{h}$ with the complementary ideal $\mathfrak{h}^\perp$.*

Let Z be the center of G. By 2.2.2.17, Z is a Lie subgroup of G and its Lie algebra $L(Z)$ is the center of $L(G)$. By 1.6.8, if we put $\mathfrak{k} = L(Z)^\perp$, $\mathfrak{k}$ is an ideal in $L(G)$ and $L(G) = \mathfrak{k} \oplus L(Z)$.

Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{a}$ and $\mathfrak{b}$ two ideals in $\mathfrak{g}$. Let $[\mathfrak{a}, \mathfrak{b}]$ be the span of all commutators $[\xi, \eta]$, $\xi \in \mathfrak{a}$, $\eta \in \mathfrak{b}$. Then $[\mathfrak{a}, \mathfrak{b}]$ is an ideal in $\mathfrak{g}$.

1.6.9. LEMMA. $\mathfrak{k} = [L(G), L(G)]$.

PROOF. Let $\xi, \eta \in L(G)$ and $\zeta \in L(Z)$. Then we have

$$([\xi, \eta] | \zeta) = (\mathrm{ad}(\xi)(\eta) | \zeta) = -(\eta | \mathrm{ad}(\xi)(\zeta)) = 0.$$

Therefore, we have $[L(G), L(G)] \subset L(Z)^\perp$. This implies that $L(Z) \subset [L(G), L(G)]^\perp$.

Conversely, let $\xi, \eta \in L(G)$ and $\zeta \in [L(G), L(G)]^\perp$. Then we have

$$0 = ([\xi, \eta] | \zeta) = (\mathrm{ad}(\xi)(\eta) | \zeta) = -(\eta | \mathrm{ad}(\xi)(\zeta)) = -(\eta | [\xi, \zeta]).$$

Since η is arbitrary, it follows that $[\xi, \zeta] = 0$ for any $\xi \in L(G)$. Therefore, ζ is in the center of $\mathfrak{g}$. It follows that $[L(G), L(G)]^\perp \subset L(Z)$. □

In particular, the decomposition $L(G) = \mathfrak{k} \oplus L(Z)$ does not depend on the choice of the invariant inner product on $L(G)$.

1.6.10. THEOREM. *Let G be a connected compact Lie group. Then the following statements are equivalent:*

(i) *The center Z of G is finite;*
(ii) *The universal covering group $\tilde{G}$ of G is compact.*

PROOF. Let $C \subset \tilde{G}$ be the kernel of the covering projection $p : \tilde{G} \longrightarrow G$. By 1.6.3, C is a finitely generated abelian subgroup of $\tilde{G}$. Assume that C is not finite. Then, by 1.7.7, we have $C = C_1 \times \mathbb{Z}$, for some finitely generated abelian group C_1. The projection to the second factor defines a homomorphism φ of C into $\mathbb{Z}$. By 1.6.6, this homomorphism extends to a Lie group homomorphism $\varphi : \tilde{G} \longrightarrow \mathbb{R}$. The kernel of $L(\varphi) : L(G) \longrightarrow \mathbb{R}$ is an ideal $\mathfrak{a}$ of codimension 1 in $L(G)$. Moreover, if $\xi, \eta \in L(G)$, we have

$$L(\varphi)([\xi, \eta]) = [L(\varphi)(\xi), L(\varphi)(\eta)] = 0.$$

Hence, $[L(G), L(G)] \subset \mathfrak{a}$. It follows that $[L(G), L(G)]$ is a nontrivial ideal in $L(G)$. By 1.6.9, this implies that $L(Z)$ is nonzero. Therefore, Z is not finite.

Therefore, we proved that if Z is finite, C must be finite too. Hence $\tilde{G}$ is a finite cover of G. By 1.6.2, this implies that $\tilde{G}$ is compact.

Conversely, assume that the center Z of G is infinite. Since Z is compact, it has finitely many components. Therefore, the identity component of Z has to be infinite. It follows that $L(Z)$ is nonzero and $q = \dim L(Z) > 0$. Let K be the integral subgroup of G corresponding to $[L(G), L(G)]$ and $\tilde{K}$ the universal covering group of K. Then $\tilde{K} \times \mathbb{R}^q$ is a simply connected, connected Lie group with Lie algebra isomorphic to $[L(G), L(G)] \times L(Z) \cong L(G)$. By 2.2.4.2, we conclude that $\tilde{G}$ is isomorphic to $\tilde{K} \times \mathbb{R}^q$. In particular, $\tilde{G}$ is not compact. □

1.7. Appendix: Finitely generated abelian groups. Let A be an abelian group. The group A is *finitely generated* if there exists elements $a_1, a_2, \dots, a_n$ such that the homomorphism

$$\mathbb{Z} \oplus \mathbb{Z} \oplus \cdots \oplus \mathbb{Z} \ni (m_1, m_2, \dots, m_n) \longmapsto m_1 a_1 + m_2 a_2 + \cdots + m_n a_n \in A$$

is surjective. The elements $a_1, a_2, \dots, a_n$ are *generators* of A.

A finitely generated abelian group A is *free*, if there is a family $a_1, a_2, \dots, a_n$ of generators of A such that the homomorphism $\mathbb{Z} \oplus \mathbb{Z} \oplus \cdots \oplus \mathbb{Z} \ni (m_1, m_2, \dots, m_n) \longmapsto m_1 a_1 + m_2 a_2 + \cdots + m_n a_n$ is an isomorphism. In this case, the family $a_1, a_2, \dots, a_n$ of generators is called a *basis* of A.

1.7.1. LEMMA. *All bases of a free finitely generated abelian group have same cardinality.*

PROOF. Let $a_1, a_2, \dots, a_n$ be a basis of A. Then $A/2A$ is a product of n copies of two-element group. Therefore, the number of elements of $A/2A$ is equal to 2^n. □

The cardinality of a basis of a free finitely generated abelian group is called the *rank* of A.

1.7.2. LEMMA. *Let A be a finitely generated abelian group and B a free finitely generated abelian group. Let $\varphi : A \longrightarrow B$ be a surjective group homomorphism. Let $C = \ker \varphi$. Then there exists a subgroup B' of A such that $A = C \oplus B'$ and the restriction of φ to B' is an isomorphism of B' onto B.*

PROOF. Let $b_1, b_2, \dots, b_n$ be a basis of B. We can pick $a_1, a_2, \dots, a_n$ such that $\varphi(a_i) = b_i$ for $1 \leq i \leq n$. Let B' be the subgroup generated by $a_1, a_2, \dots, a_n$. Then the homomorphism $\psi : \mathbb{Z} \oplus \mathbb{Z} \oplus \cdots \oplus \mathbb{Z} \ni (m_1, m_2, \dots, m_n) \longmapsto m_1 a_1 + m_2 a_2 + \cdots + m_n a_n$ is a surjection on B'. Moreover, since $b_1, b_2, \dots, b_n$ is a basis of B, $\varphi \circ \psi$ is an isomorphism. Therefore, ψ has to be injective. It follows that B' is a free finitely generated abelian group. Moreover, $B \cap C = \{0\}$.

Let $a \in A$. Then $\varphi(a) = m_1 b_1 + \cdots + m_n b_n$ for some integers $m_1, \dots, m_n$. This in turn implies that $a - (m_1 a_1 + \cdots + m_n a_n)$ is in the kernel of φ, i.e., it is in C. It follows that $a \in C \oplus B'$. □

1.7.3. LEMMA. *Let A be a free finitely generated abelian group. Let B be a subgroup of A. Then B is a free finitely generated abelian group and* $\operatorname{rank} B \leq \operatorname{rank} A$.

PROOF. We prove the statement by induction in the rank of A. If the rank is 1, A is isomorphic to $\mathbb{Z}$ and its subgroups are either isomorphic to $\mathbb{Z}$ or $\{0\}$.

Assume that the statement is true for free abelian groups of rank $\leq n-1$.

Assume that the rank of A is n. Let $a_1, a_2, \dots, a_n$ be a basis of A.

We can consider the homomorphism $\varphi : m_1 a_1 + m_2 a_2 + \cdots + m_n a_n \longmapsto m_n$ of A into $\mathbb{Z}$. Let A' be the kernel of φ. Then, A' is free abelian group with basis $a_1, a_2, \dots, a_{n-1}$. Moreover, $B' = B \cap A'$ is a free abelian group of rank $\leq n-1$ by the induction assumption.

Now, either B is a subset of A' or not. In the first case, $B = B'$ and B is a free abelian group of rank $\leq n-1$. In the second case, $\varphi(B)$ is a nontrivial subgroup of $\mathbb{Z}$. As we remarked above, this implies that $\varphi(B)$ is isomorphic to $\mathbb{Z}$. Therefore, by 1.7.2, $B = B' \oplus C$, where C is a subgroup isomorphic to $\mathbb{Z}$. It follows that B is a free abelian group of rank $\leq n$. □

1.7.4. LEMMA. *Let A be a finitely generated abelian group and B its subgroup. Then B is also finitely generated.*

PROOF. Since A is finitely generated, there exist a free finitely generated abelian group F and surjective group homomorphism $\varphi : F \longrightarrow A$. Let $B' = \varphi^{-1}(B)$. Then B' is a subgroup of F, and by 1.7.3, it is finitely generated. Let $b_1, b_2, \dots, b_p$ be a family of generators of B'. Then $\varphi(b_1), \varphi(b_2), \dots, \varphi(b_p)$ generate B. □

Let A be an abelian group. Let $a, b \in A$ be two cyclic elements in A, i.e., $pa = qb = 0$ for sufficiently large $p, q \in \mathbb{N}$. Then $pq(a+b) = 0$ and $a+b$ is also cyclic. This implies that all cyclic elements in A form a subgroup. This subgroup is called the *torsion subgroup* of A. We say that A is *torsion-free* if the torsion subgroup of A is trivial.

1.7.5. LEMMA. *Let A be a finitely generated abelian group. Then its torsion subgroup T is finite.*

PROOF. By 1.7.4, T is finitely generated. Let $t_1, t_2, \dots, t_n$ be a family of generators of T. Since $t_1, t_2, \dots, t_n$ are cyclic, there exists $p \in \mathbb{N}$ such that $pt_i = 0$ for all $1 \leq i \leq n$. This implies that any element $t \in T$ is of the form $t = m_1 t_1 + m_2 t_2 + \cdots + m_n t_n$ with $m_i \in \mathbb{Z}_+$ and $0 \leq m_i < p$. Therefore, T is finite. □

1.7.6. LEMMA. *Let A be a torsion-free finitely generated abelian group. Then A is free.*

PROOF. Assume that $A \neq \{0\}$. Let S be a finite set of generators of A. Then, it contains a nonzero element a. Hence, since A is torsion-free, $ma = 0$ implies $m = 0$.

Let $a_1, a_2, \dots, a_n$ be a maximal subset of S such that

$$m_1 a_1 + m_2 a_2 + \cdots + m_n a_n = 0$$

implies that $m_1 = m_2 = \cdots = m_n = 0$. Let B be the subgroup generated by $a_1, a_2, \dots, a_n$. Then B is a free finitely generated subgroup of A.

Let $a \in S$ different from $a_1, a_2, \dots, a_n$. By the maximality, there exist integers $m, m_1, m_2, \dots, m_n$, not all equal to zero, such that

$$ma + m_1 a_1 + m_2 a_2 + \cdots + m_n a_n = 0.$$

Again, by maximality, it follows that $m \neq 0$. Therefore, a multiple ma of a is in B. Since S is finite, there exists m such that $ma \in B$ for any $a \in S$. This implies that $mA \subset B$. Since A is torsion free the endomorphism $a \longmapsto ma$ of A is injective. Therefore, A is isomorphic to a subgroup mA of B. On the other hand, mA is a free finitely generated abelian group by 1.7.3. This implies that A is free. □

1.7.7. Theorem. *Let A be a finitely generated abelian group and T its torsion subgroup. Then there exists a subgroup B of A such that*

(i) *B is a free finitely generated abelian group;*
(ii) *$A = T \oplus B$.*

Proof. Let $\bar{a}$ be an element of A/T represented by $a \in A$. Assume that $m\bar{a} = 0$ for some $m \in \mathbb{N}$. Then $ma \in T$, and ma is cyclic. This in turn implies that a is cyclic, i.e., $a \in T$. It follows that $\bar{a} = 0$. Therefore, A/T is torsion-free. By 1.7.6, A/T is a free finitely generated abelian group. The statement follows from 1.7.2. □

1.8. Compact semisimple Lie groups. A Lie algebra $\mathfrak{g}$ is called *simple*, if it is not abelian and it doesn't contain any nontrivial ideals.

Clearly, all one-dimensional Lie algebras are abelian. The only nonabelian two-dimensional Lie algebra has an one-dimensional ideal. Therefore, there are no simple Lie algebras of dimension ≤ 2.

On the other hand, assume that $G = \mathrm{SU}(2)$. Then G is a connected, compact three-dimensional Lie group. Its Lie algebra $L(G)$ is the Lie algebra of all 2×2-skewadjoint matrices of trace zero, i.e., the Lie algebra of all matrices of the form

$$\begin{pmatrix} ia & b + ic \\ -b + ic & -ia \end{pmatrix} \text{ with } a, b, c \in \mathbb{R}.$$

Therefore, $L(G)$ is spanned by matrices

$$X = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \ Y = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \text{ and } Z = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} .$$

By a short computation we find that

$$[X, Y] = 2Z, \ [Y, Z] = 2X, \ [Z, X] = 2Y.$$

Assume that $\mathfrak{a}$ is a nonzero ideal in $L(G)$. Let $aX + bY + cZ \in \mathfrak{a}$. Since the above commutation relations are invariant under cyclic permutation of X, Y, Z, we can assume that $a \neq 0$. Then

$$\mathrm{ad}(Y)(aX + bY + cZ) = -2aZ + 2cX \in \mathfrak{a},$$

and finally

$$\mathrm{ad}(X)(aZ - cX) = -2aY \in \mathfrak{a}.$$

Therefore, $Y \in \mathfrak{a}$. From the commutation relations, we see that this immediately implies that X and Z are in $\mathfrak{a}$, and $\mathfrak{a} = L(G)$. Therefore, $L(G)$ is a simple Lie algebra.

A Lie algebra is called *semisimple* if it doesn't contain any nonzero abelian ideals. Clearly, a simple Lie algebra is semisimple.

Also, the center of a semisimple Lie algebra is always trivial.

A Lie group is called *semisimple* (resp. *simple*) if its Lie algebra is semisimple (resp. simple).

Consider now an arbitrary connected compact Lie group G. Let Z be the center of G. Since Z is compact and abelian, its identity component Z_0 is a torus in G. By 1.6.9, we have

$$L(G) = [L(G), L(G)] \oplus L(Z).$$

1.8.1. Lemma. *The ideal $[L(G), L(G)]$ in $L(G)$ is a semisimple Lie algebra.*

Proof. Let $\mathfrak{a}$ be an abelian ideal in $[L(G), L(G)]$. Then, by 1.6.8, $\mathfrak{a}^\perp$ is an ideal in $L(G)$ and $L(G)$ is the product of $\mathfrak{a}$ and $\mathfrak{a}^\perp$. This implies that $\mathfrak{a}$ is in the center $L(Z)$ of $L(G)$. By 1.6.9, it follows that that $\mathfrak{a} = \{0\}$. Therefore, $[L(G), L(G)]$ is semisimple. □

Let $H = G/Z_0$. Then H is a connected compact Lie group. Let $p : G \longrightarrow H$ be the natural projection. Then, $\ker L(p) = L(Z)$, by 2.2.2.9. Therefore, $L(p)$ induces a Lie algebra isomorphism of $[L(G), L(G)]$ onto $L(H)$. By 1.8.1, this implies that the center of $L(H)$ is trivial. Hence, by 2.2.2.17, it follows that the center of H is discrete. Since H is compact, the center of H is finite.

Let K be the integral subgroup of G corresponding to $[L(G), L(G)]$. Then we have the natural Lie group morphism $p : K \longrightarrow H$. As we remarked above, the differential of this morphism $L(p)$ is an isomorphism of $L(K) = [L(G), L(G)]$ onto $L(H)$. Hence, by 2.2.2.11, this map is a covering projection.

1.8.2. Lemma. *The integral subgroup K is a semisimple Lie subgroup of G.*

Proof. We remarked that K is a covering group of H. Moreover, H is a connected compact Lie group with finite center. Therefore, by 1.6.10, the universal covering group $\tilde{H}$ must be compact. This implies that K is a compact Lie group. Therefore, it must be closed in G. □

Consider the connected compact Lie group $K \times Z_0$ and the differentiable map $\varphi : K \times Z_0 \longrightarrow G$ given by $\varphi(k, z) = kz$ for $k \in K$ and $z \in Z_0$. Clearly, φ is a Lie group homomorphism and $L(\varphi)$ is an isomorphism of Lie algebras. Therefore, by 2.2.2.11, φ is a covering projection. The kernel of φ is a finite central subgroup of $K \times Z_0$. More precisely, we have

$$\ker \varphi = \{(k, z) \in K_0 \times Z_0 \mid kz = 1\} = \{(c, c^{-1}) \in K \times Z_0 \mid c \in K \cap Z_0\}.$$

Therefore we established the following result.

1.8.3. Proposition. *Let G be connected compact Lie group. Let $C = K \cap Z_0$ and $D = \{(c, c^{-1}) \in K \times Z_0 \mid c \in C\}$. Then φ induces an isomorphism of the Lie group $(K \times Z_0)/D$ with G.*

Therefore, any connected compact Lie group is a quotient by a finite central subgroup of a product of a connected compact semisimple Lie group with a torus.

This reduces the classification of connected compact Lie groups to the classification of connected compact semisimple Lie groups.

1.8.4. Example. Let $G = \mathrm{U}(2)$. Then $L(G)$ is the Lie algebra of all 2×2 skewadjoint matrices. The center of $L(G)$ consists of pure imaginary multiples of the identity matrix. Moreover, $[L(G), L(G)]$ is contained in the Lie subalgebra $\mathfrak{k}$ of 2×2 skewadjoint matrices of trace zero. Since the latter is the Lie algebra of the connected simple Lie subgroup $\mathrm{SU}(2)$, we conclude that $\mathfrak{k} = [L(G), L(G)]$. The center Z of G consists of matrices which are multiples of the identity matrix by a complex number α, $|\alpha| = 1$. Therefore, the center of G is connected. In addition, we have $\mathrm{SU}(2) \cap Z = \{\pm I\}$. Hence, $G = (\mathrm{SU}(2) \times Z)/\{\pm I\}$.

Moreover, we have the following result.

1.8.5. Corollary. *Let G be a connected compact Lie group. Then its universal covering group is a product of a simply connected, connected compact semisimple Lie group with $\mathbb{R}^p$ for some $p \in \mathbb{Z}_+$.*

Proof. As we have seen it the proof of 1.8.2 the universal covering group $\tilde{K}$ of K is a simply connected, connected compact semisimple Lie group. On the other hand, the universal cover of Z_0 is $\mathbb{R}^p$ for $p = \dim Z_0$. Therefore, by 1.8.3, $\tilde{K} \times \mathbb{R}^p$ is isomorphic to the universal covering group of G. □

Another byproduct of the above discussion is the following variation of 1.6.10.

1.8.6. Theorem. *Let G be a connected compact semisimple Lie group. Then its universal covering group $\tilde{G}$ is compact.*

This reduces the classification of connected compact semisimple Lie groups to the classification of simply connected, connected compact semisimple Lie groups.

In addition, we see that, for a connected semisimple Lie group G, its compactness depends only on the Lie algebra $L(G)$. To find an algebraic criterion for compactness, we need some preparation.

1.8.7. Proposition. *Let G be a compact Lie group. Let $\mathfrak{h}$ be a semisimple Lie subalgebra of $L(G)$. Then the integral subgroup H attached to $\mathfrak{h}$ is a compact Lie subgroup.*

Proof. Clearly, we can assume that G is connected. Assume first that $\mathfrak{h}$ is an ideal in $L(G)$. By 1.1.1, the Lie algebra $L(G)$ admits an G-invariant inner product. By 1.6.8, the Lie algebra $L(G)$ is a direct product of the ideal $\mathfrak{h}$ and its orthogonal complement ideal $\mathfrak{h}^\perp$.

Assume, in addition that G is also semisimple. In this case, by 1.8.6 the universal covering group $\tilde{G}$ of G is also compact. Let $\tilde{H}$ be the universal covering group of H. Also, let K be integral subgroup corresponding to $\mathfrak{h}^\perp$ and $\tilde{K}$ its universal covering group. Then $\tilde{H} \times \tilde{K}$ is a simply connected, connected Lie group with Lie algebra $\mathfrak{h} \times \mathfrak{h}^\perp \cong L(G)$. Hence, by 2.2.4.2, it follows that $\tilde{H} \times \tilde{K}$ is isomorphic to $\tilde{G}$. Since $\tilde{G}$ is compact, it follows that $\tilde{H}$ is also compact. This in turn implies that H is a compact subset of G and therefore closed in G. By 2.2.11.1, it follows that H is a Lie subgroup of G. This proves the assertion in this case.

Now we drop the assumption that G is semisimple. Let ζ be an element of the center $\mathfrak{z}$ of $L(G)$. Then $\zeta = \zeta' + \zeta''$ where $\zeta' \in \mathfrak{h}$ and $\zeta'' \in \mathfrak{h}^\perp$. Let $\xi \in \mathfrak{h}$. Then we have

$$[\xi, \zeta'] = [\xi, \zeta - \zeta''] = [\xi, \zeta] - [\xi, \zeta''] = 0.$$

It follows that ζ' is in the center of $\mathfrak{h}$. Since $\mathfrak{h}$ is semisimple, $\zeta' = 0$. It follows that $\zeta \in \mathfrak{h}^\perp$. Therefore, we have $\mathfrak{z} \subset \mathfrak{h}^\perp$. By 1.6.9, this implies that $\mathfrak{h} \subset \mathfrak{z}^\perp =$

$[L(G), L(G)]$. Let M be the integral subgroup of G corresponding to the ideal $[L(G), L(G)]$. By 1.8.2, M is a compact Lie subgroup of G. By 1.8.1, it is also semisimple. By 2.2.7.6, since $L(H) = \mathfrak{h} \subset L(M)$, we see that $H \subset M$ and we are in the situation above. Therefore, H is a compact Lie subgroup of M, and therefore a compact Lie subgroup of G.

Now we drop the assumption that $\mathfrak{h}$ is an ideal. Let $\bar{H}$ be the closure of H. Then $\bar{H}$ is a connected compact subgroup of G. By Cartan's theorem, 2.2.11.1, it is a Lie subgroup of G. Since $H \subset \bar{H}$, we see that $\mathfrak{h} \subset L(\bar{H})$. Clearly, $\mathrm{Ad}(h)(\mathfrak{h}) = \mathfrak{h}$ for any $h \in H$. Therefore, by continuity, $\mathrm{Ad}(h)(\mathfrak{h}) = \mathfrak{h}$ for any $h \in \bar{H}$. By differentiation, we see that $\mathrm{ad}(\xi)(\mathfrak{h}) \subset \mathfrak{h}$ for all $\xi \in L(\bar{H})$, i.e., $\mathfrak{h}$ is an ideal in $L(\bar{H})$. This reduces the situation to the second case. It follows that H is a compact Lie subgroup of $\bar{H}$. Therefore, it is also a compact subset of G. Hence, it is closed in G and equal to $\bar{H}$. □

Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over a filed k of characteristic 0. Define the bilinear form $B : \mathfrak{g} \times \mathfrak{g} \longrightarrow k$ by

$$B(\xi, \eta) = \mathrm{tr}(\mathrm{ad}(\xi)\,\mathrm{ad}(\eta))$$

for all $\xi, \eta \in \mathfrak{g}$. This form is called the *Killing form* of $\mathfrak{g}$.

Let A be an automorphism of $\mathfrak{g}$. Then for $\xi, \eta \in \mathfrak{g}$ we have

$$\mathrm{ad}(A\xi)(\eta) = [A\xi, \eta] = A[\xi, A^{-1}\eta] = (A\,\mathrm{ad}(\xi)A^{-1})(\eta),$$

i.e.,

$$\mathrm{ad}(A\xi) = A\,\mathrm{ad}(\xi)A^{-1}$$

for all $\xi \in \mathfrak{g}$. Therefore,

$$\begin{aligned} B(A\xi, A\eta) &= \mathrm{tr}(\mathrm{ad}(A\xi)\,\mathrm{ad}(A\eta)) = \mathrm{tr}(A\,\mathrm{ad}(\xi)A^{-1}A\,\mathrm{ad}(\eta)A^{-1}) \\ &= \mathrm{tr}(A\,\mathrm{ad}(\xi)\,\mathrm{ad}(\eta)A^{-1}) = \mathrm{tr}(\mathrm{ad}(\xi)\,\mathrm{ad}(\eta)) = B(\xi, \eta), \end{aligned}$$

for any $\xi, \eta \in \mathfrak{g}$.

Let $\mathrm{Aut}(\mathfrak{g})$ denote the automorphism group of $\mathfrak{g}$.

1.8.8. Lemma. *The Killing form on* $\mathfrak{g}$ *is* $\mathrm{Aut}(\mathfrak{g})$*-invariant.*

Let G be a Lie group. Then $\mathrm{Ad} : G \longrightarrow \mathrm{GL}(L(G))$ is a homomorphism of G into $\mathrm{Aut}(L(G))$. Therefore, the Killing form on $L(G)$ is $\mathrm{Ad}(G)$-invariant.

The following result gives a criterion for compactness of a connected semisimple Lie group in terms of its Lie algebra.

1.8.9. Theorem. *Let G be a connected Lie group. Then the following conditions are equivalent:*

(i) *G is compact and semisimple;*
(ii) *the Killing form on $L(G)$ is negative definite.*

Proof. (i)⇒(ii) Assume that G is compact and semisimple. Then, by 1.1.1, there exists an $\mathrm{Ad}(G)$-invariant inner product on $L(G)$. With respect to this inner product, Ad is a homomorphism of G into $\mathrm{O}(L(G))$. Therefore ad is a Lie algebra homomorphism of $L(G)$ into the Lie algebra of antisymmetric linear transformations on $L(G)$. Let $\xi \in L(G)$. Then $B(\xi, \xi) = \mathrm{tr}(\mathrm{ad}(\xi)^2)$ is the sum of squares of all (complex) eigenvalues of $\mathrm{ad}(\xi)$. Since $\mathrm{ad}(\xi)$ is antisymmetric, all its eigenvalues are pure imaginary. Hence their squares are negative. This implies that $B(\xi, \xi) \leq 0$ and $B(\xi, \xi) = 0$ implies that all eigenvalues of $\mathrm{ad}(\xi)$ are equal to 0. Since $\mathrm{ad}(\xi)$

is antisymmetric, it follows that $\operatorname{ad}(\xi) = 0$. Therefore, ξ is in the center of $L(G)$. Since $L(G)$ is semisimple, its center is equal to $\{0\}$, i.e., $\xi = 0$. Therefore, B is negative definite.

(ii)⇒(i) Assume that B is negative definite. Then, $(\xi|\eta) = -B(\xi, \eta)$ is an $\operatorname{Ad}(G)$-invariant inner product on $L(G)$. Therefore, Ad is a Lie group morphism of G into the compact Lie group $\mathrm{O}(L(G))$. Moreover, ad is a Lie algebra morphism of $L(G)$ into the Lie algebra of all antisymmetric linear transformations on $L(G)$.

Let $\mathfrak{a}$ be an abelian ideal in $L(G)$. As in the proof of 1.6.8 we see that $L(G) = \mathfrak{a} \oplus \mathfrak{a}^\perp$ and $\mathfrak{a}$ is in the center of $L(G)$. Let $\zeta \in \mathfrak{a}$. Then $\operatorname{ad}\zeta = 0$ since ζ is in the center of $L(G)$. It follows that $B(\zeta, \zeta) = 0$. Since B is negative definite, we see that $\zeta = 0$. Therefore, we see that $\mathfrak{a} = \{0\}$. Hence, G is semisimple.

By 2.2.2.17, the center Z of G is equal $\ker \operatorname{Ad}$ and its Lie algebra is equal to $\{0\}$. Hence Z is a discrete subgroup of G. Therefore, Ad induces an injective immersion of G/Z into $\mathrm{O}(L(G))$. Hence, the image $\operatorname{Ad}(G)$ is an integral subgroup of $\mathrm{O}(L(G))$ isomorphic to G/Z. Its Lie algebra is isomorphic to $L(G)$, hence it is semisimple. By 1.8.7, $\operatorname{Ad}(G)$ is a compact Lie subgroup of $\mathrm{O}(L(G))$. Hence, G is a covering group of a connected compact semisimple Lie group. By 1.8.6, G is a compact Lie group. □

Let G be a connected compact semisimple Lie group. Then, by 1.8.9, $(\xi, \eta) \longmapsto -B(\xi, \eta)$ is an $\operatorname{Ad}(G)$-invariant inner product on $L(G)$. Let $\mathfrak{a}$ be an ideal in $L(G)$. Then, by 1.6.8, $\mathfrak{a}^\perp$ is a complementary ideal in $L(G)$, i.e., $L(G) = \mathfrak{a} \oplus \mathfrak{a}^\perp$.

Assume that $\mathfrak{b}$ is another ideal in $L(G)$ such that $\mathfrak{a} \cap \mathfrak{b} = \{0\}$. Let $\xi \in \mathfrak{a}$ and $\eta \in \mathfrak{b}$. Then, $\operatorname{ad}(\eta)(L(G)) \subset \mathfrak{b}$ and

$$(\operatorname{ad}(\xi)\operatorname{ad}(\eta))(L(G)) = \operatorname{ad}(\xi)(\operatorname{ad}(\eta)(L(G)) \subset \operatorname{ad}(\xi)(\mathfrak{b}) \subset \mathfrak{a} \cap \mathfrak{b} = \{0\}.$$

Therefore, $\operatorname{ad}(\xi)\operatorname{ad}(\eta) = 0$ and $B(\xi, \eta) = \operatorname{tr}(\operatorname{ad}(\xi)\operatorname{ad}(\eta)) = 0$. It follows that $\mathfrak{b} \subset \mathfrak{a}^\perp$.

In particular, if $\mathfrak{b}$ is a direct complement of $\mathfrak{a}$, we must have $\mathfrak{b} = \mathfrak{a}^\perp$. Therefore, the complementary ideal is unique.

The set of all ideals in $L(G)$ is ordered by inclusion. Let $\mathfrak{m}$ be a minimal ideal in $L(G)$. Since $L(G)$ is semisimple, this ideal is not abelian.

Clearly, $L(G) = \mathfrak{m} \oplus \mathfrak{m}^\perp$. Let $\mathfrak{a} \subset \mathfrak{m}$ be an ideal in $\mathfrak{m}$. Then $[\mathfrak{a}, \mathfrak{m}^\perp] = \{0\}$ and $\mathfrak{a}$ is an ideal in $L(G)$. By the minimality of $\mathfrak{m}$, $\mathfrak{a}$ is either $\mathfrak{m}$ or $\{0\}$. It follows that $\mathfrak{m}$ is a simple Lie algebra.

Let $\mathfrak{a}$ be another ideal in $L(G)$. Then $\mathfrak{a} \cap \mathfrak{m}$ is an ideal in $L(G)$. By the minimality of $\mathfrak{m}$, we have either $\mathfrak{m} \subset \mathfrak{a}$ or $\mathfrak{a} \cap \mathfrak{m} = \{0\}$. By the above discussion, the latter implies that $\mathfrak{a} \subset \mathfrak{m}^\perp$, i.e., $\mathfrak{a}$ is perpendicular to $\mathfrak{m}$.

Let $\mathfrak{m}_1, \mathfrak{m}_2, \dots, \mathfrak{m}_p$ be a family of mutually different minimal ideals in $L(G)$. By the above discussion $\mathfrak{m}_i$ is perpendicular to $\mathfrak{m}_j$ for $i \neq j$, $1 \leq i, j \leq p$. Hence, p has to be smaller than $\dim L(G)$. Assume that p is maximal possible. Then $\mathfrak{a} = \mathfrak{m}_1 \oplus \mathfrak{m}_2 \oplus \dots \oplus \mathfrak{m}_p$ is an ideal in $L(G)$. Assume that $\mathfrak{a} \neq L(G)$. Then $L(G) = \mathfrak{a} \oplus \mathfrak{a}^\perp$. Let $\mathfrak{m}_{p+1}$ be a minimal ideal in $\mathfrak{a}^\perp$. Then $\mathfrak{m}_{p+1}$ is a minimal ideal in $L(G)$ different from $\mathfrak{m}_i$, $1 \leq i \leq p$, contradicting the maximality of p. It follows that $L(G) = \mathfrak{m}_1 \oplus \mathfrak{m}_2 \oplus \dots \oplus \mathfrak{m}_p$, i.e., we have the following result.

1.8.10. Lemma. *The semisimple Lie algebra $L(G)$ is the direct product of its minimal ideals. These ideals are simple Lie algebras.*

In particular, $L(G)$ is a product of simple Lie algebras.

Let $K_1, K_2, \ldots, K_p$ be the integral subgroups of G corresponding to Lie algebras $\mathfrak{m}_1, \mathfrak{m}_2, \ldots, \mathfrak{m}_p$. Let $\tilde{K}_1, \tilde{K}_2, \ldots, \tilde{K}_p$ be their universal covering groups. Then $\tilde{K}_1 \times \tilde{K}_2 \times \cdots \times \tilde{K}_p$ is a simply connected Lie group with Lie algebra isomorphic to $L(G) = \mathfrak{m}_1 \oplus \mathfrak{m}_2 \oplus \cdots \oplus \mathfrak{m}_p$. Hence, $\tilde{K}_1 \times \tilde{K}_2 \times \cdots \times \tilde{K}_p$ is isomorphic to $\tilde{G}$ by 2.4.2. Since $\tilde{G}$ is compact by 1.8.6, the subgroups $\tilde{K}_1, \tilde{K}_2, \ldots, \tilde{K}_p$ are also compact. This in turn implies that $K_1, K_2, \ldots, K_n$ are compact Lie subgroups of G. The map $\varphi : K_1 \times K_2 \times \cdots \times K_p \longrightarrow G$ given by $\varphi(k_1, k_2, \ldots, k_p) = k_1 k_2 \ldots k_p$ for any $k_1 \in K_1, k_2 \in K_2, \ldots, k_p \in K_p$ is a Lie group homomorphism. Clearly, it is a covering projection.

Therefore, we established the following result.

1.8.11. Theorem. *Connected compact semisimple Lie group G is a quotient by a finite central subgroup of a product $K_1 \times K_2 \times \cdots \times K_p$ of connected compact simple Lie groups.*

This reduces the study of connected compact Lie groups to the study connected compact simple Lie groups.

1.9. Fundamental group of a connected compact semisimple Lie group. Let G be a connected compact semisimple Lie group with Lie algebra $L(G)$. Let $\tilde{G}$ be the universal covering group of G and $p : \tilde{G} \longrightarrow G$ be the covering projection. By 1.8.6, $\tilde{G}$ is also compact. Hence, $\ker p$ is a finite central subgroup of $\tilde{G}$.

Then, as we remarked in 2.1.6, we have $\pi_1(G, 1) \cong \ker p$. In particular, $\pi(G, 1)$ is a finite abelian group.

Let T be a maximal torus in G and $L(T)$ its Lie algebra. By 1.2.1, its Lie algebra is a maximal abelian Lie subalgebra in $L(G)$. Let $\tilde{T}$ be the corresponding integral subgroup in $\tilde{G}$. Then, by 1.2.1, $\tilde{T}$ is a maximal torus in $\tilde{G}$. The map p induces a Lie group homomorphism q of $\tilde{T}$ onto T which is a covering map. Clearly, $\ker q \subset \ker p$.

Let Z be the center of $\tilde{G}$. By 1.4.2, Z is contained in $\tilde{T}$. In particular, $\ker p \subset Z \subset \tilde{T}$. This implies that $\ker q = \ker p$.

By 2.2.9.6, we have the commutative diagram

$$\begin{array}{ccc} & L(T) & \\ {\scriptstyle \exp_{\tilde{T}}} \swarrow & & \searrow {\scriptstyle \exp_T} \\ \tilde{T} & \xrightarrow{\quad q \quad} & T \end{array}$$

of Lie groups. Put $L = \ker \exp_T$ and $\tilde{L} = \ker \exp_{\tilde{T}}$. Then L and $\tilde{L}$ are discrete subgroups of $L(T)$ of rank $\dim L(T)$.

Clearly, $\tilde{L} \subset L$ and

$$\ker p = \ker q = L/\tilde{L}.$$

Therefore, any connected compact semisimple Lie group G with Lie algebra $L(G)$ determines a discrete subgroup L of $L(T)$ which contains $\tilde{L}$.

In the proof of 1.8.9 we proved that $G_0 = \mathrm{Ad}(\tilde{G})$ is a connected compact Lie group with Lie algebra $L(G)$. Moreover, by 2.2.2.17, the center Z of $\tilde{G}$ is equal to the kernel of $\mathrm{Ad} : \tilde{G} \longrightarrow G_0$. Let T_0 be the maximal torus in G_0 corresponding to $L(T)$. Then the above construction attaches to G_0 a discrete subgroup L_0 of $L(T)$ containing $\tilde{L}$.

In addition, we see that the following result holds.

1.9.1. LEMMA. *The center Z of $\tilde{G}$ is isomorphic to $L_0/\tilde{L}$.*

From the above discussion we see the following result.

1.9.2. THEOREM. *The map $G \longmapsto L$ defines a surjection from all connected compact semisimple Lie groups with Lie algebra $L(G)$ onto all discrete subgroups L in $L(T)$ such that $L_0 \supset L \supset \tilde{L}$.*

The center of G is isomorphic to L_0/L. The fundamental group $\pi_1(G, 1)$ is isomorphic to $L/\tilde{L}$.

1.9.3. EXAMPLE. Let $G = \mathrm{SU}(2)$. Then G is a connected compact simple Lie group. The subgroup

$$T = \left\{ \begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix} \middle| \varphi \in \mathbb{R} \right\}.$$

is a maximal torus in G. As we remarked in 2.1.8, the group G is simply connected and it is a two-fold covering of the group $\mathrm{SO}(3)$. The covering projection induces a Lie group morphism

$$\begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(2\varphi) & -\sin(2\varphi) \\ 0 & \sin(2\varphi) & \cos(2\varphi) \end{pmatrix}$$

of the torus T onto a torus T_0 in $\mathrm{SO}(3)$. Since the center of $\mathrm{SO}(3)$ is trivial, $\mathrm{SO}(3)$ is isomorphic to the adjoint group $\mathrm{Ad}(G)$. If we identify $L(T)$ with $\mathbb{R}$ and the exponential map with

$$\exp_T : \varphi \longmapsto \begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix}$$

the discrete subgroup L corresponds to $2\pi\mathbb{Z}$ and L_0 to $\pi\mathbb{Z}$.

CHAPTER 4

Basic Lie algebra theory

All Lie algebras in this chapter are finite dimensional Lie algebras over a field k of characteristic 0. All representations of Lie algebras are finite dimensional.

1. Solvable, nilpotent and semisimple Lie algebras

1.1. Derivations and characteristic ideals. Let $\mathfrak{g}$ be a Lie algebra over a field k. A *derivation* D of $\mathfrak{g}$ is a linear map on $\mathfrak{g}$ such that

$$D[x,y] = [Dx,y] + [x,Dy]$$

for all $x, y \in \mathfrak{g}$.

1.1.1. Lemma. *All derivations of* $\mathfrak{g}$ *for a Lie subalgebra* $\operatorname{Der}(\mathfrak{g})$ *of* $\mathcal{L}(\mathfrak{g})$.

Proof. Clearly, the set of all derivations of $\mathfrak{g}$ is a linear subspace of $\mathcal{L}(V)$. Let D, D' be two derivations of $\mathfrak{g}$. Then we have

$$(DD')[x,y] = D([D'x,y]+[x,D'y]) = [DD'x,y]+[D'x,Dy]+[Dx,D'y]+[x,DD'y]$$

and

$$\begin{aligned}[D,D']([x,y]) = (DD'-D'D)[x,y] = [DD'x,y]+[D'x,Dy]+[Dx,D'y]+[x,DD'y] \\ -[D'Dx,y]-[Dx,D'y]-[D'x,Dy]-[x,D'Dy] = [[D,D']x,y]+[x,[D,D']y]\end{aligned}$$

for all $x, y \in \mathfrak{g}$. Therefore, $[D, D']$ is a derivation of $\mathfrak{g}$. It follows that $\operatorname{Der}(\mathfrak{g})$ is a Lie subalgebra of $\mathcal{L}(\mathfrak{g})$. □

If $x \in \mathfrak{g}$, we have

$$\operatorname{ad} x([y,z]) = [x,[y,z]] = -[y,[z,x]] - [z,[x,y]] = [\operatorname{ad} x(y), z] + [y, \operatorname{ad} x(z)]$$

for all $y, z \in \mathfrak{g}$. Therefore, $\operatorname{ad} x$ is a derivation of $\mathfrak{g}$. The derivations $\operatorname{ad} x$, $x \in \mathfrak{g}$, are called the *inner derivations* of $\mathfrak{g}$.

Therefore, $\operatorname{ad} : \mathfrak{g} \longrightarrow \mathcal{L}(\mathfrak{g})$ is a Lie algebra homomorphism into $\operatorname{Der}(\mathfrak{g})$.

Let D be a derivation of $\mathfrak{g}$ and $x \in \mathfrak{g}$. Then

$$\operatorname{ad}(Dx)(y) = [Dx,y] = D[x,y] - [x,Dy] = [D, \operatorname{ad} x](y)$$

for any $y \in \mathfrak{g}$.

1.1.2. Lemma. *Let D be a derivation of $\mathfrak{g}$ and $x \in \mathfrak{g}$. Then*

$$\operatorname{ad}(Dx) = [D, \operatorname{ad} x].$$

The image of ad is the space of all inner derivations in $\operatorname{Der}(\mathfrak{g})$.

1.1.3. Lemma. *The linear space* $\operatorname{im} \operatorname{ad}$ *of all inner derivations is an ideal in* $\operatorname{Der}(\mathfrak{g})$.

Proof. This follows immediately from 1.1.2. □

If $\mathfrak{h}$ is an ideal in $\mathfrak{g}$, $\mathfrak{h}$ is an invariant subspace for $\operatorname{ad} x$ for any $x \in \mathfrak{g}$.

A linear subspace $\mathfrak{h}$ in $\mathfrak{g}$ is a *characteristic ideal* if $D(\mathfrak{h}) \subset \mathfrak{h}$ for all $D \in \operatorname{Der}(\mathfrak{g})$. Clearly, a characteristic ideal in $\mathfrak{g}$ is an ideal in $\mathfrak{g}$.

Let $\mathfrak{a}$ and $\mathfrak{b}$ be two characteristic ideals in $\mathfrak{g}$. Then $[\mathfrak{a}, \mathfrak{b}]$ is a characteristic ideal in $\mathfrak{g}$.

Let B be the Killing form on $\mathfrak{g}$, i.e.,

$$B(x, y) = \operatorname{tr}(\operatorname{ad}(x) \operatorname{ad}(y)) \text{ for } x, y \in \mathfrak{g}.$$

1.1.4. Lemma. *Let $D \in \operatorname{Der}(\mathfrak{g})$. Then*

$$B(Dx, y) + B(x, Dy) = 0$$

for any $x, y \in \mathfrak{g}$.

Proof. By 1.1.2, we have

$$\begin{aligned} B(Dx, y) + B(x, Dy) &= \operatorname{tr}(\operatorname{ad}(Dx) \operatorname{ad}(y)) + \operatorname{tr}(\operatorname{ad}(x) \operatorname{ad}(Dy)) \\ &= \operatorname{tr}([D, ad(x)] \operatorname{ad}(y)) + \operatorname{tr}(\operatorname{ad}(x)[D, \operatorname{ad}(y)]) \end{aligned}$$
$$= \operatorname{tr}(D \operatorname{ad}(x) \operatorname{ad}(y)) - \operatorname{tr}(\operatorname{ad}(x) D \operatorname{ad}(y)) + \operatorname{tr}(\operatorname{ad}(x) D \operatorname{ad}(y)) - \operatorname{tr}(\operatorname{ad}(x) \operatorname{ad}(y) D) = 0.$$

□

Let $\mathfrak{h}$ be a linear subspace in $\mathfrak{g}$. We denote by $\mathfrak{h}^\perp$ the linear space

$$\mathfrak{h}^\perp = \{x \in \mathfrak{g} \mid B(x, y) = 0 \text{ for all } y \in \mathfrak{h}\}.$$

1.1.5. Lemma. (i) *Let $\mathfrak{h}$ be an ideal in $\mathfrak{g}$. Then $\mathfrak{h}^\perp$ is an ideal in $\mathfrak{g}$.*
(ii) *Let $\mathfrak{h}$ be a characteristic ideal in $\mathfrak{g}$. Then $\mathfrak{h}^\perp$ is a characteristic ideal in $\mathfrak{g}$.*

Proof. (i) Let $x \in \mathfrak{h}^\perp$. Then

$$B(\operatorname{ad}(y)x, z) = -B(x, \operatorname{ad}(y)z) = 0$$

for any $y \in \mathfrak{g}$ and $z \in \mathfrak{h}$.

(ii) Let $x \in \mathfrak{h}^\perp$. Then

$$B(Dx, y) = -B(x, Dy) = 0$$

for any $y \in \mathfrak{h}$ and $D \in \operatorname{Der}(\mathfrak{g})$. □

1.2. Solvable Lie algebras. Let $\mathfrak{g}$ be a Lie algebra. We put

$$\mathcal{D}\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}].$$

This is the *derived ideal* of $\mathfrak{g}$. We put

$$\mathcal{D}^0\mathfrak{g} = \mathfrak{g},\ \mathcal{D}^1\mathfrak{g} = \mathcal{D}\mathfrak{g},\ \mathcal{D}^p\mathfrak{g} = [\mathcal{D}^{p-1}\mathfrak{g}, \mathcal{D}^{p-1}\mathfrak{g}] \text{ for } p \geq 2.$$

These are characteristic ideals in $\mathfrak{g}$. The decreasing sequence

$$\mathfrak{g} \supseteq \mathcal{D}\mathfrak{g} \supseteq \mathcal{D}^2\mathfrak{g} \supseteq \cdots \supseteq \mathcal{D}^p\mathfrak{g} \supseteq \ldots$$

is called the *derived series* of ideals in $\mathfrak{g}$.

Since $\mathfrak{g}$ is finite dimensional, the derived series has to stabilize, i.e., $\mathcal{D}^p\mathfrak{g} = \mathcal{D}^{p+1}\mathfrak{g} = \ldots$ for sufficiently large p.

We say that the Lie algebra $\mathfrak{g}$ is *solvable* if $\mathcal{D}^p\mathfrak{g} = \{0\}$ for some $p \in \mathbb{N}$.

Clearly, an abelian Lie algebra is solvable.

1.2.1. Lemma. (i) *Let $\mathfrak{g}$ be a solvable Lie algebra and $\mathfrak{h} \subset \mathfrak{g}$ a Lie subalgebra. Then $\mathfrak{h}$ is solvable.*

(ii) *Let $\mathfrak{g}$ be a solvable Lie algebra and $\mathfrak{h} \subset \mathfrak{g}$ an ideal in $\mathfrak{g}$. Then $\mathfrak{g}/\mathfrak{h}$ is solvable.*
(iii) *Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{h}$ an ideal in $\mathfrak{g}$. If $\mathfrak{h}$ and $\mathfrak{g}/\mathfrak{h}$ are solvable, $\mathfrak{g}$ is a solvable Lie algebra.*

PROOF. (i) We have
$$\mathcal{D}\mathfrak{h} = [\mathfrak{h}, \mathfrak{h}] \subseteq [\mathfrak{g}, \mathfrak{g}] = \mathcal{D}\mathfrak{g}.$$
Moreover, by induction in p, we get
$$\mathcal{D}^p\mathfrak{h} = [\mathcal{D}^{p-1}\mathfrak{h}, \mathcal{D}^{p-1}\mathfrak{h}] \subseteq [\mathcal{D}^{p-1}\mathfrak{g}, \mathcal{D}^{p-1}\mathfrak{g}] = \mathcal{D}^p\mathfrak{g}$$
for all $p \in \mathbb{N}$. Therefore, if $\mathfrak{g}$ is solvable, $\mathcal{D}^p\mathfrak{g} = \{0\}$ for some $p \in \mathbb{N}$. This in turn implies that $\mathcal{D}^p\mathfrak{h} = \{0\}$, i.e., $\mathfrak{h}$ is a solvable Lie algebra.

(ii) Let $\pi : \mathfrak{g} \longrightarrow \mathfrak{g}/\mathfrak{h}$ be the natural projection. Then $\mathcal{D}(\mathfrak{g}/\mathfrak{h}) = \pi(\mathcal{D}\mathfrak{g})$. By induction in p, we see that
$$\mathcal{D}^p(\mathfrak{g}/\mathfrak{h}) = [\mathcal{D}^{p-1}(\mathfrak{g}/\mathfrak{h}), \mathcal{D}^{p-1}(\mathfrak{g}/\mathfrak{h})] = \pi([\mathcal{D}^{p-1}\mathfrak{g}, \mathcal{D}^{p-1}\mathfrak{g}]) = \pi(\mathcal{D}^p\mathfrak{g})$$
for any $p \in \mathbb{N}$. If $\mathfrak{g}$ is solvable, $\mathcal{D}^p\mathfrak{g} = \{0\}$ for some $p \in \mathbb{N}$. This in turn implies that $\mathcal{D}^p(\mathfrak{g}/\mathfrak{h}) = \{0\}$, i.e., $\mathfrak{g}/\mathfrak{h}$ is a solvable Lie algebra.

(iii) Since $\mathfrak{g}/\mathfrak{h}$ is solvable, $\mathcal{D}^p(\mathfrak{g}/\mathfrak{h}) = \{0\}$ for some $p \in \mathbb{N}$. Therefore, $\mathcal{D}^p\mathfrak{g} \subset \mathfrak{h}$. Since $\mathfrak{h}$ is solvable, $\mathcal{D}^q\mathfrak{h} = \{0\}$ for some $q \in \mathbb{N}$. Therefore,
$$\mathcal{D}^{p+q}\mathfrak{g} = \mathcal{D}^q(\mathcal{D}^p\mathfrak{g}) \subseteq \mathcal{D}^q\mathfrak{h} = \{0\},$$
and $\mathfrak{g}$ is solvable. □

1.2.2. EXAMPLE. Let $\mathfrak{g}$ be the two-dimensional nonabelian Lie algebra discussed in 2.2.1. Then $\mathfrak{g}$ is spanned by e_1 and e_2 and $\mathcal{D}\mathfrak{g}$ is spanned by e_1. This implies that $\mathcal{D}^2\mathfrak{g} = \{0\}$, i.e., $\mathfrak{g}$ is a solvable Lie algebra.

Let $\mathfrak{a}$ and $\mathfrak{b}$ be two solvable ideals in the Lie algebra $\mathfrak{g}$. Then $\mathfrak{a} + \mathfrak{b}$ is an ideal in $\mathfrak{g}$. Moreover, by 1.2.1, $\mathfrak{a} \cap \mathfrak{b}$ is solvable. On the other hand, $(\mathfrak{a} + \mathfrak{b})/\mathfrak{b} = \mathfrak{a}/(\mathfrak{a} \cap \mathfrak{b})$. Therefore, by 1.2.1, $\mathfrak{a} + \mathfrak{b}$ is a solvable ideal.

Let $\mathcal{S}$ be the family of all solvable ideals in $\mathfrak{g}$. Since $\mathfrak{g}$ is finite dimensional, there exist maximal elements in $\mathcal{S}$. Let $\mathfrak{a}$ and $\mathfrak{b}$ be two maximal solvable ideals in $\mathfrak{g}$. Then $\mathfrak{a} + \mathfrak{b}$ is a solvable ideal containing $\mathfrak{a}$ and $\mathfrak{b}$. Therefore, we must have $\mathfrak{a} = \mathfrak{a} + \mathfrak{b} = \mathfrak{b}$. It follows that $\mathcal{S}$ contains the unique maximal element. This is the largest solvable ideal in $\mathfrak{g}$.

The largest solvable ideal of $\mathfrak{g}$ is called the *radical* of $\mathfrak{g}$.

1.3. Semisimple Lie algebras. A Lie algebra $\mathfrak{g}$ is *semisimple* if its radical is equal to $\{0\}$.

The next result shows that this definition is equivalent to the definition in 3.1.8.

1.3.1. LEMMA. *A Lie algebra is semisimple if and only if it has no nonzero abelian ideals.*

PROOF. If $\mathfrak{g}$ contains a nonzero abelian ideal $\mathfrak{a}$, the radical $\mathfrak{r}$ of $\mathfrak{g}$ contains $\mathfrak{a}$. Therefore, $\mathfrak{g}$ is not semisimple.

Let $x \in \mathfrak{g}$. Then $\operatorname{ad} x$ induces a derivation of $\mathfrak{r}$. Since $\mathcal{D}^p\mathfrak{r}$, $p \in \mathbb{N}$, are characteristic ideals in $\mathfrak{r}$, we see that $\operatorname{ad} x(\mathcal{D}^p\mathfrak{r}) \subset \mathcal{D}^p\mathfrak{r}$ for any $p \in \mathbb{N}$. Therefore, all $\mathcal{D}^p\mathfrak{r}$ are ideals in $\mathfrak{g}$. Let $q \in \mathbb{Z}_+$ be such that $\mathcal{D}^q\mathfrak{r} \neq \{0\}$ and $\mathcal{D}^{q+1}\mathfrak{r} = \{0\}$. Then, $\mathfrak{a} = \mathcal{D}^q\mathfrak{r}$ is a nonzero abelian ideal in $\mathfrak{g}$. □

In particular, the center of a semisimple Lie algebra is $\{0\}$. Since the center of $\mathfrak{g}$ is the kernel of ad we see that the following result holds.

1.3.2. LEMMA. *Let $\mathfrak{g}$ be a semisimple Lie algebra. Then* $\ker \operatorname{ad} = \{0\}$.

1.3.3. PROPOSITION. *Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{r}$ its radical. Then $\mathfrak{g}/\mathfrak{r}$ is semisimple.*

PROOF. Let $\pi : \mathfrak{g} \longrightarrow \mathfrak{g}/\mathfrak{r}$ be the natural projection. Assume that $\mathfrak{s}$ is a solvable ideal in $\mathfrak{g}/\mathfrak{r}$, and put $\pi^{-1}(\mathfrak{s}) = \mathfrak{s}'$. Then, $\mathfrak{s}'$ is an ideal in $\mathfrak{g}$ containing $\mathfrak{r}$. In addition, $\mathfrak{s}'/\mathfrak{r} = \mathfrak{s}$, and by 1.2.1, $\mathfrak{s}'$ is solvable. This in turn implies that $\mathfrak{s}' = \mathfrak{r}$ and $\mathfrak{s} = \{0\}$, i.e., the only solvable ideal in $\mathfrak{g}/\mathfrak{r}$ is $\{0\}$. Therefore, $\mathfrak{g}/\mathfrak{r}$ is semisimple. □

1.4. Nilpotent Lie algebras. Let $\mathfrak{g}$ be a Lie algebra. Define $\mathcal{C}\mathfrak{g} = \mathcal{D}\mathfrak{g}$, and

$$\mathcal{C}^0\mathfrak{g} = \mathfrak{g},\ \mathcal{C}^1\mathfrak{g} = \mathcal{C}\mathfrak{g},\ \mathcal{C}^p\mathfrak{g} = [\mathfrak{g}, \mathcal{C}^{p-1}\mathfrak{g}] \text{ for } p \geq 2.$$

Moreover,

$$\mathfrak{g} \supseteq \mathcal{C}\mathfrak{g} \supseteq \mathcal{C}^2\mathfrak{g} \supseteq \cdots \supseteq \mathcal{C}^p\mathfrak{g} \supseteq \ldots$$

is a decreasing sequence of characteristic ideals which is called the *descending central series.*

Since $\mathfrak{g}$ is finite dimensional, the descending central series has to stabilize, i.e., $\mathcal{C}^p\mathfrak{g} = \mathcal{C}^{p+1}\mathfrak{g} = \ldots$ for sufficiently large p.

Clearly, since $\mathcal{C}\mathfrak{g} = \mathcal{D}\mathfrak{g}$, by induction we have

$$\mathcal{C}^p\mathfrak{g} = [\mathfrak{g}, \mathcal{C}^{p-1}\mathfrak{g}] \supseteq [\mathcal{D}^{p-1}\mathfrak{g}, \mathcal{D}^{p-1}\mathfrak{g}] = \mathcal{D}^p\mathfrak{g}$$

for all $p \in \mathbb{N}$.

A Lie algebra $\mathfrak{g}$ is *nilpotent* if $\mathcal{C}^p\mathfrak{g} = \{0\}$ for some $p \in \mathbb{N}$.

Clearly, abelian Lie algebras are nilpotent. Also, nilpotent Lie algebras are solvable.

On the other hand, the two-dimensional solvable Lie algebra we considered in 1.2.2 is not nilpotent. As we remarked, $\mathcal{C}\mathfrak{g} = \mathcal{D}\mathfrak{g}$ is spanned by the vector e_1. This in turn implies that $\mathcal{C}^2\mathfrak{g} = [\mathfrak{g}, \mathcal{C}\mathfrak{g}] = \mathcal{C}\mathfrak{g}$ and inductively $\mathcal{C}^p\mathfrak{g} = \mathcal{C}\mathfrak{g}$ for all $p \in \mathbb{N}$.

1.4.1. LEMMA. (i) *Let $\mathfrak{g}$ be a nilpotent Lie algebra and $\mathfrak{h} \subset \mathfrak{g}$ a Lie subalgebra. Then $\mathfrak{h}$ is nilpotent.*
(ii) *Let $\mathfrak{g}$ be a nilpotent Lie algebra and $\mathfrak{h} \subset \mathfrak{g}$ an ideal in $\mathfrak{g}$. Then $\mathfrak{g}/\mathfrak{h}$ is nilpotent.*

PROOF. (i) We have $\mathcal{C}\mathfrak{h} \subseteq \mathcal{C}\mathfrak{g}$. Moreover, by induction in p, we get

$$\mathcal{C}^p\mathfrak{h} = [\mathfrak{h}, \mathcal{C}^{p-1}\mathfrak{h}] \subseteq [\mathfrak{g}, \mathcal{C}^{p-1}\mathfrak{g}] = \mathcal{C}^p\mathfrak{g}$$

for all $p \in \mathbb{N}$. Therefore, if $\mathfrak{g}$ is nilpotent, $\mathcal{C}^p\mathfrak{g} = \{0\}$ for some $p \in \mathbb{N}$. This in turn implies that $\mathcal{C}^p\mathfrak{h} = \{0\}$, i.e., $\mathfrak{h}$ is a nilpotent Lie algebra.

(ii) Let $\pi : \mathfrak{g} \longrightarrow \mathfrak{g}/\mathfrak{h}$ be the natural projection. Then $\mathcal{C}(\mathfrak{g}/\mathfrak{h}) = \pi(\mathcal{C}\mathfrak{g})$. By induction in p, we see that

$$\mathcal{C}^p(\mathfrak{g}/\mathfrak{h}) = [\mathfrak{g}/\mathfrak{h}, \mathcal{C}^{p-1}(\mathfrak{g}/\mathfrak{h})] = \pi([\mathfrak{g}, \mathcal{C}^{p-1}\mathfrak{g}]) = \pi(\mathcal{C}^p\mathfrak{g})$$

for any $p \in \mathbb{N}$. If $\mathfrak{g}$ is nilpotent, $\mathcal{C}^p\mathfrak{g} = \{0\}$ for some $p \in \mathbb{N}$. This in turn implies that $\mathcal{C}^p(\mathfrak{g}/\mathfrak{h}) = \{0\}$, i.e., $\mathfrak{g}/\mathfrak{h}$ is a nilpotent Lie algebra. □

On the other hand, the extensions of nilpotent Lie algebras do not have to be nilpotent. For example, the nonabelian two-dimensional solvable Lie algebra $\mathfrak{g}$ has a one-dimensional abelian ideal $\mathcal{D}\mathfrak{g}$ and the quotient $\mathfrak{g}/\mathcal{D}\mathfrak{g}$ is a one-dimensional abelian Lie algebra.

1.5. Engel's theorem. Let V be a finite-dimensional linear space and $\mathcal{L}(V)$ the Lie algebra of all linear transformations on V.

1.5.1. LEMMA. *Let $T \in \mathcal{L}(V)$. Then*

$$(\operatorname{ad} T)^p S = \sum_{i=0}^{p} (-1)^i \binom{p}{i} T^{p-i} S T^i$$

for any $p \in \mathbb{Z}_+$.

PROOF. We prove this statement by induction in p. It is obvious for $p = 0$. Therefore, we have

$$\begin{aligned}(\operatorname{ad} T)^{p+1} S = T(\operatorname{ad} T)^p S - (\operatorname{ad} T)^p S T &= \sum_{i=0}^{p} (-1)^i \binom{p}{i} \left(T^{p-i+1} S T^i - T^{p-i} S T^{i+1}\right) \\ &= \sum_{i=0}^{p} (-1)^i \binom{p}{i} T^{p-i+1} S T^i + \sum_{i=1}^{p+1} (-1)^i \binom{p}{i-1} T^{p-i+1} S T^i \\ &= \sum_{i=0}^{p+1} (-1)^i \left(\binom{p}{i} + \binom{p}{i-1}\right) T^{p-i+1} S T^i = \sum_{i=0}^{p+1} \binom{p+1}{i} T^{p+1-i} S T^i.\end{aligned}$$

□

1.5.2. COROLLARY. *Let T is a nilpotent linear transformation on V. Then $\operatorname{ad} T$ is a nilpotent linear transformation on $\mathcal{L}(V)$.*

PROOF. We have $T^p = 0$ for some $p \in \mathbb{Z}_+$. By 1.5.1, it follows that $(\operatorname{ad} T)^{2p} = 0$, and $\operatorname{ad} T$ is nilpotent. □

1.5.3. THEOREM (Engel). *Let V be a finite dimensional linear space and $\mathfrak{g}$ a Lie subalgebra of $\mathcal{L}(V)$ consisting of nilpotent linear transformations. Then there exists a vector $v \in V$, $v \neq 0$, such that $Tv = 0$ for all $T \in \mathfrak{g}$.*

PROOF. We prove the theorem by induction in dimension of $\mathfrak{g}$. The statement is obvious if $\dim \mathfrak{g} = 1$.

Now we want to show that $\mathfrak{g}$ contains an ideal $\mathfrak{a}$ of codimension 1. Let $\mathfrak{h}$ be an arbitrary Lie subalgebra of $\mathfrak{g}$ such that $\dim \mathfrak{h} < \dim \mathfrak{g}$. Let $T \in \mathfrak{h}$. Then T is a nilpotent linear transformation, and by 1.5.2, $\operatorname{ad} T$ is a nilpotent linear transformation on $\mathcal{L}(V)$. Since $\mathfrak{g}$ and $\mathfrak{h}$ invariant subspaces for $\operatorname{ad} T$, it induces a nilpotent linear transformation $\sigma(T)$ on $\mathfrak{g}/\mathfrak{h}$. Clearly, $\sigma : \mathfrak{h} \longrightarrow \mathcal{L}(\mathfrak{g}/\mathfrak{h})$ is a representation of $\mathfrak{h}$. By the induction assumption, there exists a linear transformation $R \in \mathfrak{g}$, $R \notin \mathfrak{h}$, such that $\sigma(T)(R + \mathfrak{h}) = 0$, i.e., $[T, R] = \operatorname{ad} T(R) \in \mathfrak{h}$ for all $T \in \mathfrak{h}$. Let $\mathfrak{h}'$ be the linear span of R and $\mathfrak{h}$. Then $\mathfrak{h}'$ is a Lie subalgebra of $\mathfrak{g}$, $\dim \mathfrak{h}' = \dim \mathfrak{h} + 1$, and $\mathfrak{h}$ is an ideal in $\mathfrak{h}'$ of codimension 1. By induction in dimension of $\mathfrak{h}$, starting with $\mathfrak{h} = \{0\}$, we show that $\mathfrak{g}$ contains an ideal $\mathfrak{a}$ of codimension 1.

Let $T \in \mathfrak{g}$, $T \notin \mathfrak{a}$. By the induction assumption, there exists $w \in V$, $w \neq 0$, such that $Sw = 0$ for any $S \in \mathfrak{a}$. Consider the linear subspace $U = \{u \in V \mid Su = 0 \text{ for all } S \in \mathfrak{a}\}$. Clearly, U is nonzero. If $u \in U$, we have

$$S(Tu) = STu - TSu = [S, T]u = 0,$$

for all $S \in \mathfrak{a}$, since $[S, T] \in \mathfrak{a}$. Therefore, $Tu \in U$. It follows that U is invariant for T. Since T is nilpotent, there exists $v \in U$, $v \neq 0$, such that $Tv = 0$. Hence, v is annihilated by all elements of $\mathfrak{g}$. □

The following result characterizes nilpotent Lie algebras in terms of their adjoint representations.

1.5.4. PROPOSITION. *Let $\mathfrak{g}$ be a Lie algebra. Then the following conditions are equivalent:*

(i) *$\mathfrak{g}$ is nilpotent;*
(ii) *all* $\operatorname{ad} x$*, $x \in \mathfrak{g}$, are nilpotent.*

PROOF. (i)⇒(ii) Assume that $\mathfrak{g}$ is nilpotent. Let $\mathcal{C}^p\mathfrak{g} = \{0\}$. By induction, we can find a basis of $\mathfrak{g}$ by completing the basis of $\mathcal{C}^s\mathfrak{g}$ to a basis of $\mathcal{C}^{s-1}\mathfrak{g}$ for all $s \le p$. In this basis, all $\operatorname{ad} x$, $x \in \mathfrak{g}$, are upper triangular matrices with zeros on diagonal.

(ii)⇒(i) If all $\operatorname{ad} x$ are nilpotent, by 1.5.3, there exists $y \in \mathfrak{g}$, $y \neq 0$, such that $[x, y] = \operatorname{ad}(x)y = 0$ for all $x \in \mathfrak{g}$. Therefore, the center $\mathfrak{z}$ of $\mathfrak{g}$ is different from $\{0\}$.

We proceed by induction in dimension of $\mathfrak{g}$. If $\mathfrak{g}$ is abelian, the statement is obvious. Assume that $\mathfrak{g}$ is not abelian, and consider $\mathfrak{g}/\mathfrak{z}$. Clearly, $\dim(\mathfrak{g}/\mathfrak{z}) < \dim \mathfrak{g}$. Moreover, for any $x \in \mathfrak{g}/\mathfrak{z}$, $\operatorname{ad} x$ is nilpotent. Therefore, by the induction assumption, $\mathfrak{g}/\mathfrak{z}$ is nilpotent. This implies that $\mathcal{C}^p(\mathfrak{g}/\mathfrak{z}) = \{0\}$ for some $p \in \mathbb{N}$. It follows that $\mathcal{C}^p\mathfrak{g} \subset \mathfrak{z}$. Hence, $\mathcal{C}^{p+1}\mathfrak{g} = \{0\}$, and $\mathfrak{g}$ is nilpotent. □

The next result implies that all Lie algebras which satisfy the conditions of 1.5.3 are nilpotent.

1.5.5. COROLLARY. *Let V be a finite-dimensional linear space. Let $\mathfrak{g}$ be a Lie subalgebra of $\mathcal{L}(V)$ consisting of nilpotent linear transformations. Then $\mathfrak{g}$ is nilpotent.*

PROOF. Let $T \in \mathfrak{g}$. Then, by 1.5.2, $\operatorname{ad} T$ is nilpotent linear transformation on $\mathcal{L}(V)$. Therefore, it is a nilpotent linear transformation on $\mathfrak{g}$. By 1.5.4, $\mathfrak{g}$ is nilpotent. □

1.5.6. EXAMPLE. Let $M_n(k)$ be the Lie algebra of $n \times n$ matrices with entries in k. Let $\mathfrak{n}(n, k)$ be the Lie subalgebra of all upper triangular matrices in $M_n(k)$ with zeros on the diagonal. Then $\mathfrak{n}(n, k)$ is a nilpotent Lie algebra.

Let $\mathfrak{g}$ be a nilpotent Lie algebra. In the proof of 1.5.4, we proved that there exists a basis of $\mathfrak{g}$ such that the matrices of $\operatorname{ad} x$, $x \in \mathfrak{g}$, in this basis are upper triangular and nilpotent. Therefore, for any $x, y \in \mathfrak{g}$, the matrix of $\operatorname{ad}(x)\operatorname{ad}(y)$ is upper triangular and nilpotent. In particular $B(x, y) = \operatorname{tr}(\operatorname{ad}(x)\operatorname{ad}(y)) = 0$. This proves the following result.

1.5.7. LEMMA. *Let $\mathfrak{g}$ be a nilpotent Lie algebra. Then its Killing form B is trivial.*

1.6. Lie's theorem. In this section we prove some basic properties of solvable Lie algebras over an algebraically closed field k.

1.6.1. LEMMA. *Let $\mathfrak{g}$ be a Lie algebra over an algebraically closed field k and $\mathfrak{r}$ its radical. Let π be an irreducible representation of $\mathfrak{g}$ on a linear space V over k. Then there exists a linear form λ on $\mathfrak{r}$ such that $\pi(x) = \lambda(x)1_V$ for all $x \in \mathfrak{r}$.*

PROOF. Let $\mathfrak{a} = \pi(\mathfrak{g})$ and $\mathfrak{p} = \pi(\mathfrak{r})$. Then $\mathfrak{a}$ is a Lie subalgebra of $\mathcal{L}(V)$ and $\mathfrak{p}$ is a solvable ideal in $\mathfrak{a}$.

Fix $p \in \mathbb{Z}_+$ such that $\mathfrak{b} = \mathcal{D}^p\mathfrak{p} \neq 0$, $\mathcal{D}^{p+1}\mathfrak{p} = \{0\}$. Clearly, $\mathfrak{b}$ is an abelian characteristic ideal in $\mathfrak{p}$. Therefore, it is an ideal in $\mathfrak{a}$. Since the field k is algebraically closed, $T \in \mathfrak{b}$ have a common eigenvector $v \in V$, $v \neq 0$. Therefore,

$$Tv = \lambda(T)v \text{ for all } T \in \mathfrak{b}.$$

Clearly, λ is a linear form on $\mathfrak{b}$.

Let $S \in \mathfrak{a}$. Since $\mathfrak{b}$ is an ideal in $\mathfrak{a}$, we have $[S, T] \in \mathfrak{b}$ for all $T \in \mathfrak{b}$. We claim that $\lambda([S, T]) = 0$.

Let V_n be the subspace of V spanned by $v, Sv, \dots, S^n v$. Clearly, we have

$$V_0 \subseteq V_1 \subseteq \dots \subseteq V_n \subseteq \dots$$

and this increasing sequence stabilizes since V is finite dimensional. Assume that $V_{m-1} \neq V_m = V_{m+1} = \dots$. Then V_m is invariant for S. Moreover, $v, Sv, \dots, S^m v$ form a basis of of V_m, and $\dim V_m = m + 1$.

We claim that

$$TS^n v - \lambda(T)S^n v \in V_{n-1}$$

for all $T \in \mathfrak{b}$ and $n \in \mathbb{Z}_+$. This is obvious for $n = 0$. We prove the statement by induction in n. Since

$$\begin{aligned} TS^{n+1}v - \lambda(T)S^{n+1}v &= [T, S]S^n v + STS^n v - \lambda(T)S^{n+1}v \\ &= [T, S]S^n v + S(TS^n v - \lambda(T)S^n v), \end{aligned}$$

and $[T, S] \in \mathfrak{b}$, by the induction assumption we have

$$[T, S]S^n v - \lambda([T, S])S^n v \in V_{n-1} \text{ and } TS^n v - \lambda(T)S^n v \in V_{n-1}.$$

Therefore,

$$TS^{n+1}v - \lambda(T)S^{n+1}v \in \lambda([T, S])S^n v + V_{n-1} + S(V_{n-1}) \subset V_n.$$

This proves the above statement.

It follows that V_m is invariant for $T \in \mathfrak{b}$. Moreover, in the basis $v, Sv, \dots, S^m v$ of V_m, $T \in \mathfrak{b}$ act by upper triangular matrices with $\lambda(T)$ on the diagonal. Therefore, we have

$$\operatorname{tr}(T|_{V_m}) = (m + 1)\lambda(T)$$

for any $T \in \mathfrak{b}$. In particular, this holds for $[T, S]$. Therefore,

$$(m + 1)\lambda([T, S]) = \operatorname{tr}([T, S]|_{V_m}) = \operatorname{tr}([T|_{V_m}, S|_{V_m}]) = 0.$$

Hence, we have $\lambda([T, S]) = 0$ as we claimed above.

Consider now the linear subspace

$$W = \{v \in V \mid Tv = \lambda(T)v, T \in \mathfrak{b}\}.$$

Then $v \in W$ and $W \neq \{0\}$. For any $w \in W$ and $S \in \mathfrak{a}$, we have

$$TSw = [T, S]w + STw = \lambda([T, S])w + \lambda(T)Sw = \lambda(T)Sw$$

for all $T \in \mathfrak{b}$. Hence, $Sw \in W$. Therefore, W is $\mathfrak{a}$-invariant. Since π is irreducible, we must have $W = V$. It follows that $T = \lambda(T)1_V$ for any $T \in \mathfrak{b}$. If $p > 0$, $\mathfrak{b}$ is spanned by commutators and the trace of any element of $\mathfrak{b}$ must be zero. This implies that $\lambda = 0$ and $\mathfrak{b} = \{0\}$, contradicting the choice of p. Therefore, $p = 0$, and $\mathfrak{b} = \mathfrak{p}$. □

The following result is an immediate consequence of 1.6.1.

1.6.2. Corollary. *Let $\mathfrak{g}$ be a solvable Lie algebra over an algebraically closed field k. Then any irreducible representation of $\mathfrak{g}$ on a linear space V over k is one-dimensional.*

1.6.3. Theorem (Lie). *Let $\mathfrak{g}$ be a solvable Lie algebra over an algebraically closed field k. Let π be a representation of $\mathfrak{g}$ on a linear space V over k. Then there exist a basis of V such that all matrices of $\pi(x)$, $x \in \mathfrak{g}$, are upper triangular.*

Proof. We prove this statement by induction in $\dim V$. If $\dim V = 1$, the statement is obvious.

Assume that $\dim V = n > 1$. Let W be a minimal invariant subspace in V for π. Then the representation of $\mathfrak{g}$ on W is irreducible. By 1.6.2, $\dim W = 1$. Let e_1 be a nonzero vector in W. Clearly, π defines a representation σ of $\mathfrak{g}$ on V/W and $\dim V/W = \dim V - 1$. Therefore, by the induction assumption, there exist vectors $e_2, \dots, e_n$ in V such that $e_2 + W, \dots, e_n + W$, form a basis of V/W such that $\sigma(x)$ are upper triangular in that basis, i.e, the subspaces spanned by $e_2 + W, \dots, e_k + W$, $2 \leq k \leq n$, are σ-invariant. This in turn implies that the subspaces spanned by $e_1, \dots, e_k$, $2 \leq k \leq n$, are π-invariant, i.e., the matrices of $\pi(x)$, $x \in \mathfrak{g}$, are upper triangular. □

2. Lie algebras and field extensions

2.1. k-structures on linear spaces. Let k be a field of characteristic 0. Let K be an algebraically closed field containing k.

Let U be a linear space over k. Then $K \otimes_k U$ has a natural structure of a linear space over K.

A *k-structure* on a linear space V over K is a k-linear subspace $V_k \subset V$ such that the natural map

$$K \otimes_k V_k \longrightarrow V$$

is an isomorphism. This means that V_k spans V over K and the elements of V_k linearly independent over k are also linearly independent over K.

We say that the elements of V_k are *rational* over k.

Let V be a linear space over K and V_k its k-structure. Let U be a linear subspace of V. We put $U_k = U \cap V_k$. We say that U is *defined* (or *rational*) *over* k if U_k is a k-structure on U. This is equivalent to U_k spanning U.

If $W = V/U$, we write W_k for the projection of V_k into W. We say that W is *defined over* k if W_k is a k-structure on W.

2.1.1. Lemma. *Let U be a linear subspace of V and $W = V/U$. Then the following conditions are equivalent:*

(i) *U is defined over k;*
(ii) *W is defined over k.*

Proof. Consider the k-linear map $V_k \longrightarrow W$ induced by the natural projection $V \longrightarrow W$. Then, its kernel is $V_k \cap U = U_k$. Therefore, we have the short exact sequence

$$0 \longrightarrow U_k \longrightarrow V_k \longrightarrow W_k \longrightarrow 0.$$

By tensoring it with K, we get the short exact sequence

$$0 \longrightarrow K \otimes_k U_k \longrightarrow K \otimes_k V_k \longrightarrow K \otimes_k W_k \longrightarrow 0.$$

This leads to the commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & K \otimes_k U_k & \xrightarrow{a} & K \otimes_k V_k & \xrightarrow{b} & K \otimes_k W_k & \longrightarrow & 0 \\ & & \alpha \downarrow & & \beta \downarrow & & \gamma \downarrow & & \\ 0 & \longrightarrow & U & \xrightarrow{A} & V & \xrightarrow{B} & W & \longrightarrow & 0 \end{array}$$

of linear spaces over K. The rows in this diagram are exact and the middle vertical arrow is an isomorphism. From the diagram it is evident that the first vertical arrow must be an injection and the last vertical arrow must be a surjection.

We claim that the first arrow is surjective if and only if the last one is injective.

Assume first that α is an isomorphism. Let $w \in \ker \gamma$. Then, $w = b(v)$ for some $v \in K \otimes_k V_k$ and

$$B(\beta(v)) = \gamma(b(v)) = \gamma(w) = 0.$$

Therefore, $\beta(v) \in \ker B$. It follows that $\beta(v) = A(u)$ for some $u \in U$. Since α is an isomorphism, $u = \alpha(u')$ for some $u' \in K \otimes_k U$. Hence, we have

$$\beta(v) = A(u) = A(\alpha(u')) = \beta(a(u')).$$

Since β is an isomorphism, this implies that $v = a(u')$. Hence, $w = b(v) = b(a(u')) = 0$. It follows that γ is injective.

Consider now that γ is an isomorphism. Let $u \in U$. Then $A(u) = \beta(v)$ for some $v \in K \otimes_k V_k$. It follows that

$$\gamma(b(v)) = B(\beta(v)) = B(A(u)) = 0.$$

By our assumption, this implies that $b(v) = 0$, and $v = a(u')$ for some $u' \in K \otimes_k U_k$. Hence, we have

$$A(u) = \beta(v) = \beta(a(u')) = A(\alpha(u')).$$

Since A is injective, it follows that $u = \alpha(u')$, i.e., α is surjective. □

Let $\mathrm{Aut}_k(K)$ be the group of all k-linear automorphisms of K. Let $\sigma \in \mathrm{Aut}_k(K)$. Then the k-linear map $\sigma : K \longrightarrow K$ defines a k-bilinear map $K \times V_k \longrightarrow V$ by $(\lambda, v) \longmapsto \sigma(\lambda)v$. This map defines a k-linear map of $K \otimes_k V_k$ into V by

$$\sigma_V(\lambda \otimes v) = \sigma(\lambda)v.$$

Since V_k is a k-structure of V, we can view σ_V as a k-linear automorphism of V. Therefore, we get a homomorphism of $\mathrm{Aut}_k(K)$ into the group of all k-linear automorphisms of V. We say that this action of $\mathrm{Aut}_k(K)$ *corresponds* to the k-structure V_k.

2.1.2. Lemma. *The k-structure V_k of V is the fixed point set of the action of $\mathrm{Aut}_k(K)$ on V.*

Proof. Let $v \in V$. Then $v = \sum_i \lambda_i v_i$ for some finite independent set of vectors $v_1, v_2, \ldots, v_n$ in V_k and $\lambda_i \in K$. Therefore, $\sigma_V(v) = \sum_i \sigma(\lambda_i) v_i$ for any $\sigma \in \mathrm{Aut}_k(K)$. It follows that v is fixed by σ_V if and only if λ_i are fixed by σ. By Galois theory, $\lambda \in K$ is fixed by all $\sigma \in \mathrm{Aut}_k(K)$ if and only if $\lambda \in k$. Therefore, v is fixed by $\mathrm{Aut}_k(K)$ if and only if $v \in V_k$. □

Let $v, w \in V$ and $\lambda, \mu \in K$. We can represent v and w as

$$v = \sum_i \alpha_i v_i \text{ and } w = \sum_j \beta_j w_j$$

where $\alpha_i, \beta_j \in K$ and $v_i, w_j \in V_k$. Therefore, we have

$$\begin{aligned}\sigma_V(\lambda v + \mu w) = \sigma_V\left(\lambda \sum_i \alpha_i v_i + \mu \sum_j \beta_j w_j\right) = \sigma_V\left(\sum_i \lambda\alpha_i v_i + \sum_j \mu\beta_j w_j\right)\\ = \sum_i \sigma(\lambda\alpha_i) v_i + \sum_j \sigma(\mu\beta_j) w_j = \sigma(\lambda)\sum_i \sigma(\alpha_i) v_i + \sigma(\mu)\sum_j \sigma(\beta_j) w_j\\ = \sigma(\lambda)\sigma_V(v) + \sigma(\mu)\sigma_V(w).\end{aligned}$$

It follows that for any K-linear subspace U of V, $\sigma_V(U)$ is also a K-linear subspace of V. Therefore, $\mathrm{Aut}_k(K)$ permutes K-linear subspaces of V.

Assume that the K-linear subspace U is defined over k. Then, any $u \in U$ can be written as $u = \sum_i \lambda_i u_i$ for $\lambda_i \in K$ and $u_i \in U_k$. Hence, $\sigma_V(u) = \sum_i \sigma(\lambda_i) u_i \in U$. Therefore, $\sigma_V(U) = U$ and U is invariant for the action of $\mathrm{Aut}_k(K)$.

2.1.3. Lemma. *If U is a K-linear subspace of V defined over k, U is stable for the action of* $\mathrm{Aut}_k(K)$.

Moreover, $\mathrm{Aut}_k(K)$ induces the action on U which corresponds to the k-structure U_k.

Let $W = V/U$. Then the action of $\mathrm{Aut}_k(K)$ induces an action on W. if $p : V \longrightarrow W$ is the canonical projection,

$$\sigma_W(p(v)) = \sigma_W\left(\sum_i \lambda_i p(v_i)\right) = \sum_i \sigma(\lambda_i) p(v_i)$$

for $\lambda_i \in K$ and $v_i \in V_k$. Therefore, the action of $\mathrm{Aut}_k(K)$ on W is corresponds to the k-structure W_k.

Now we want to prove the converse of the above lemma.

2.1.4. Proposition. *Let U be a K-linear subspace of V stable for the action of* $\mathrm{Aut}_k(K)$. *Then U is defined over k.*

Proof. Let $U_k = U \cap V_k$. Also, put $U' = K \otimes_k U_k \subset V$. Then U' is defined over k, and $U' \subset U$.

Let $\bar{V} = V/U'$ with the induced k-structure. The image $\bar{U}$ of U in $\bar{V}$ is $\mathrm{Aut}_k(K)$-invariant. Let $\bar{u} \in \bar{U} \cap \bar{V}_k$. Then $\bar{u} = p(u) = p(v)$ for some $u \in U$ and $v \in V_k$. Hence, $p(u - v) = 0$, and $u - v \in U'$. In particular, $u - v \in U$. Hence, $v \in U$, and $v \in U_k$. It follows that $v \in U'$ and $p(v) = 0$. Therefore, $\bar{u} = 0$. It follows that $\bar{U}_k = \bar{U} \cap \bar{V}_k = \{0\}$. To prove the claim, we have to show that $\bar{U} = \{0\}$. This immediately implies that $U = U'$, i.e., U is defined over k.

Therefore, we can assume from the beginning that $U_k = \{0\}$.

Assume that $U \neq \{0\}$. Let $u \in U$, $u \neq 0$, be such that $u = \sum_{i=1}^n \lambda_i v_i$ where $\lambda_i \in K$ and $v_i \in V_k$ and n is smallest possible. Then, $v_1, v_2, \dots, v_n$ must be linearly independent over K and all λ_i different from 0.

Then, multiplying by $\frac{1}{\lambda_1}$ we can assume that $u \in U$ has the form $u = v_1$ if $n = 1$ or

$$u = v_1 + \sum_{i=2}^n \lambda_i v_i.$$

In the first case, $u \in V_k$ and $u \in U \cap V_k = U_k$ contradicting the assumption that $U_k = \{0\}$. Hence, we must have $n > 1$. For any $\sigma \in \mathrm{Aut}_k(K)$, we have

$$\sigma_V(u) = v_1 + \sum_{i=2}^{n} \sigma(\lambda_i) v_i$$

and

$$\sigma_V(u) - u = \sum_{i=2}^{n} (\sigma(\lambda_i) - \lambda_i) v_i.$$

Since $\sigma_V(u) - u \in U$, and the sum on the right has $n-1$ terms, we must have $\sigma_V(u) - u = 0$ and $\sigma(\lambda_i) = \lambda_i$ for $2 \leq i \leq n$.

Let $\lambda \in K$. By Galois theory, if $\sigma(\lambda) = \lambda$ for all $\sigma \in \mathrm{Aut}_k(K)$, λ is in subfield k. Therefore, we conclude that all $\lambda_i \in k$. This implies that $u \in V_k$ and again $u \in U \cap V_k = \{0\}$, contradicting our assumption. Therefore, $U = \{0\}$. □

2.2. k-structures on Lie algebras. Let k be a field of characteristic 0 and K its algebraically closed extension. Let $\mathfrak{g}$ be a Lie algebra over k. Then we can define the commutator on $\mathfrak{g}_K = K \otimes_k \mathfrak{g}$ by

$$[\lambda \otimes x, \mu \otimes y] = \lambda\mu \otimes [x, y]$$

for any $x, y \in \mathfrak{g}$ and $\lambda, \mu \in K$. One can check that $\mathfrak{g}_K$ is a Lie algebra over K.

If $\varphi : \mathfrak{g} \longrightarrow \mathfrak{h}$ is a morphism of Lie algebras over k, by linearity it extends to the morphism $\varphi_K : \mathfrak{g}_K \longrightarrow \mathfrak{h}_K$.

In this way we construct an exact functor from the category of Lie algebras over k in to the category of Lie algebras over K. This functor is called the functor of *extension of scalars*.

If V is a linear space over k and $V_K = K \otimes_k V$. One checks that $\mathcal{L}(V)_K = \mathcal{L}(V_K)$.

Conversely, if $\mathfrak{g}$ is a Lie algebra over K, a k-linear subspace $\mathfrak{g}_k$ of $\mathfrak{g}$ is a *k-structure on Lie algebra* $\mathfrak{g}$ if

(i) $\mathfrak{g}_k$ is a k-structure on the linear space $\mathfrak{g}$;
(ii) $\mathfrak{g}_k$ is a Lie subalgebra of $\mathfrak{g}$ considered as a Lie algebra over k.

Let $\mathfrak{g}_k$ be a k-structure on the Lie algebra $\mathfrak{g}$ over K. Let $x, y \in \mathfrak{g}$. Then $x = \sum_i \lambda_i x_i$, $y = \sum_j \mu_j y_j$ for some $x_i, y_j \in \mathfrak{g}_k$ and $\lambda_i, \mu_j \in K$. Therefore,

$$\sigma_{\mathfrak{g}}([x,y]) = \sigma_{\mathfrak{g}}\left(\sum_{i,j} \lambda_i \mu_j [x_i, y_j]\right) = \sum_{i,j} \sigma(\lambda_i)\sigma(\mu_j)[x_i, y_j] = [\sigma_{\mathfrak{g}}(x), \sigma_{\mathfrak{g}}(y)]$$

for any $\sigma \in \mathrm{Aut}_k(K)$, i.e., $\mathrm{Aut}_k(K)$ acts on $\mathfrak{g}$ by k-linear automorphisms. This implies that the action of $\mathrm{Aut}_k(K)$ on $\mathfrak{g}$ permutes Lie subalgebras, resp. ideals, in the Lie algebra $\mathfrak{g}$.

A Lie subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is *defined over* k, if $\mathfrak{h}_k = \mathfrak{h} \cap \mathfrak{g}_k$ is a k-structure on $\mathfrak{h}$ as a linear space.

Since $\mathfrak{h}$ and $\mathfrak{g}_k$ are Lie subalgebras of $\mathfrak{g}$ considered as a Lie algebra over k, $\mathfrak{h}_k$ is a Lie algebra over k. Therefore, $\mathfrak{h}_k$ is k-structure on the Lie algebra $\mathfrak{h}$.

Let $\mathfrak{h}$ be an ideal in $\mathfrak{g}$. Then $\mathfrak{h}_k = \mathfrak{h} \cap \mathfrak{g}_k$ is an ideal in $\mathfrak{g}_k$.

2.2.1. LEMMA. *Let $\mathfrak{g}$ be a Lie algebra over K with k-structure $\mathfrak{g}_k$. If $\mathfrak{a}$ and $\mathfrak{b}$ are two ideals of $\mathfrak{g}$ defined over k, the ideal $[\mathfrak{a}, \mathfrak{b}]$ is defined over k.*

PROOF. Since $\mathfrak{a}$ and $\mathfrak{b}$ are defined over k, they are invariant under the action of $\mathrm{Aut}_k(K)$. By 2.1.4, this in turn implies that $[\mathfrak{a}, \mathfrak{b}]$ is defined over k. □

This immediately implies the following result.

2.2.2. COROLLARY. *Let $\mathfrak{g}$ be a Lie algebra over k, and $\mathfrak{g}_K$ the Lie algebra obtained by extension of scalars.*

(i) $\mathcal{D}^p\mathfrak{g}_K = (\mathcal{D}^p\mathfrak{g})_K$ *for all* $p \in \mathbb{Z}_+$*;*
(ii) $\mathcal{C}^p\mathfrak{g}_K = (\mathcal{C}^p\mathfrak{g})_K$ *for all* $p \in \mathbb{Z}_+$*.*

Therefore, we have the following result.

2.2.3. THEOREM. *Let $\mathfrak{g}$ be a Lie algebra over k, and $\mathfrak{g}_K$ the Lie algebra obtained by extension of scalars.*

(i) $\mathfrak{g}$ *is solvable if and only if* $\mathfrak{g}_K$ *is solvable.*
(ii) $\mathfrak{g}$ *is nilpotent if and only if* $\mathfrak{g}_K$ *is nilpotent.*

Let $\mathfrak{g}$ be a Lie algebra over K and $\mathfrak{g}_k$ its k-structure. As we remarked $\mathrm{Aut}_k(K)$ permutes ideals in $\mathfrak{g}$. Clearly, if $\mathfrak{a}$ is a solvable ideal, $\sigma_{\mathfrak{g}}(\mathfrak{a})$ is also a solvable ideal. Therefore, $\mathrm{Aut}_k(K)$ permutes solvable ideals. Since this action clearly preserves the partial ordering given by inclusion, we conclude that the radical $\mathfrak{r}$ of $\mathfrak{g}$ is fixed by the action of $\mathrm{Aut}_k(K)$. Hence, by 2.1.4, $\mathfrak{r}$ is defined over k. This implies the following result.

2.2.4. LEMMA. *Let $\mathfrak{g}$ be a Lie algebra over k, and $\mathfrak{g}_K$ the Lie algebra obtained by extension of scalars. Let $\mathfrak{r}$ be the radical of $\mathfrak{g}$. Then $\mathfrak{r}_K$ is the radical of $\mathfrak{g}_K$.*

This has the following immediate consequence.

2.2.5. THEOREM. *Let $\mathfrak{g}$ be a Lie algebra over k, and $\mathfrak{g}_K$ the Lie algebra obtained by extension of scalars. Then $\mathfrak{g}$ is semisimple if and only if $\mathfrak{g}_K$ is semisimple.*

The following observation follows immediately from the definitions.

2.2.6. LEMMA. *Let $\mathfrak{g}$ be a Lie algebra over k and $\mathfrak{g}_K$ the Lie algebra obtained by extension of scalars. Then the Killing form $B_{\mathfrak{g}_K}$ of $\mathfrak{g}_K$ is the linear extension of the Killing form $B_{\mathfrak{g}}$ on $\mathfrak{g}$.*

Also, we can prove the following characterization of solvable Lie algebras.

2.2.7. PROPOSITION. *Let $\mathfrak{g}$ be a Lie algebra. Then the following conditions are equivalent:*

(i) $\mathfrak{g}$ *is solvable;*
(ii) $\mathcal{D}\mathfrak{g}$ *is nilpotent.*

PROOF. (ii)⇒(i) The Lie algebra $\mathfrak{g}$ is an extension of $\mathfrak{g}/\mathcal{D}\mathfrak{g}$ by the ideal $\mathcal{D}\mathfrak{g}$. Clearly, $\mathfrak{g}/\mathcal{D}\mathfrak{g}$ is abelian. Hence, if $\mathcal{D}\mathfrak{g}$ is nilpotent, $\mathfrak{g}$ has to be solvable by 1.2.1.

(i)⇒(ii) Let K be the algebraic closure of k. Then $\mathfrak{g}_K$ is solvable by 2.2.3. By 1.6.3, there exists a basis of $\mathfrak{g}_K$ such that the matrices of $\operatorname{ad} x$, $x \in \mathfrak{g}_K$, are upper triangular. Hence, the matrices of $\operatorname{ad}[x, y] = [\operatorname{ad} x, \operatorname{ad} y]$ are upper triangular with zeros on the diagonal. It follows that $\operatorname{ad} x$, $x \in \mathcal{D}\mathfrak{g}_k$, are nilpotent. By 1.5.4, it follows that $\mathcal{D}\mathfrak{g}_K$ is nilpotent. By 2.2.2 and 2.2.3, $\mathcal{D}\mathfrak{g}$ is nilpotent. □

2.2.8. EXAMPLE. Let $M_n(k)$ be the Lie algebra of all $n \times n$ matrices with entries in k. Denote by $\mathfrak{s}(n, k)$ the Lie subalgebra of all upper triangular matrices in $M_n(k)$. Then, as we remarked in 1.5.6, $\mathcal{D}\mathfrak{s}(n, k) = \mathfrak{n}(n, k)$ is a nilpotent Lie algebra. Therefore, $\mathfrak{s}(n, k)$ is a solvable Lie algebra.

3. Cartan's criterion

3.1. Jordan decomposition. Let k be a field of characteristic 0 and K its algebraic closure.

3.1.1. Lemma. *Let $P \in k[X]$ be a polynomial with simple zeros in K. Then P and P' are relatively prime, i.e, there exist $S, T \in k[X]$ such that $SP + TP' = 1$.*

Proof. By our assumption, $P(X) = \prod_{i=1}^{p}(X - \lambda_i)$ for $\lambda_i \in K$, and λ_i, $1 \leq i \leq n$, are mutually different. Therefore, for any $1 \leq i \leq n$, $P(X) = (X - \lambda_i)Q(X)$ and $Q(\lambda_i) \neq 0$. It follows that $P'(X) = Q(X) + (X - \lambda_i)Q'(X)$ and $P'(\lambda_i) = Q(\lambda_i) \neq 0$.

Let I be the ideal generated by P and P' in $k[X]$. Assume that $I \neq k[X]$. Since $k[X]$ is a principal ideal domain, in this case $I = (R)$ for some polynomial $R \in k[X]$. Therefore, a zero of R in K must be a common zero of P and P' which is impossible. It follows that $I = k[X]$. □

3.1.2. Lemma. *Let $P \in k[X]$ be a polynomial with simple zeros in K. Let $n \in \mathbb{N}$. Assume that $Q \in k[X]$ is a polynomial such that $P \circ Q$ is in the ideal in $k[X]$ generated by P^n. Then there exists a polynomial $A_n \in k[X]$ such that $P \circ (Q - A_n P^n)$ is in the ideal generated by P^{n+1}.*

Proof. By Taylor's formula

$$P(X + Y) = P(X) + P'(X)Y + Y^2 R(X, Y)$$

for some $R \in k[X, Y]$. Therefore, for any polynomial A_n we have

$$P \circ (Q - A_n P^n) = P \circ Q - (P' \circ Q)A_n P^n + SP^{n+1}$$

where $S \in k[X]$. By our assumption, $P \circ Q = TP^n$ for some polynomial $T \in k[X]$. Hence, we have

$$P \circ (Q - A_n P^n) = TP^n - (P' \circ Q)A_n P^n + SP^{n+1} = (T - (P' \circ Q)A_n)P^n + SP^{n+1}.$$

By 3.1.1, there exists $A, B \in k[X]$ such that $1 = AP' + BP$. Therefore,

$$1 = (A \circ Q)(P' \circ Q) + (B \circ Q)(P \circ Q).$$

If we put $A_n = T(A \circ Q)$, we get

$$\begin{aligned} P \circ (Q - A_n P^n) &= (T - (P' \circ Q)T(A \circ Q))P^n + SP^{n+1} \\ &= T(1 - (P' \circ Q)(A \circ Q))P^n + SP^{n+1} = T(B \circ Q)(P \circ Q)P^n + SP^{n+1} \\ &= T^2(B \circ Q)P^{2n} + SP^{n+1}. \end{aligned}$$

□

By induction, from this lemma we deduce the following result.

3.1.3. Lemma. *Let $P \in k[X]$ be a polynomial with simple zeros in K. Let $n \in \mathbb{Z}_+$. Then there exist polynomials $A_0 = 0, A_1, \ldots, A_n$ such that the polynomial*

$$P\left(X - \sum_{i=0}^{n} A_i(X)P(X)^i\right)$$

is in the ideal generated by P^{n+1}.

Proof. If $n = 0$, the statement is evident for $A_0 = 0$.

Assume that the statement holds for $n-1$. then there exist polynomials $A_0 = 0, A_1, \dots, A_{n-1}$ such that

$$P\left(X - \sum_{i=0}^{n-1} A_i(X)P(X)^i\right)$$

is in the ideal generated by P^n. Put

$$Q = X - \sum_{i=0}^{n-1} A_i(X)P(X)^i.$$

Then the existence of A_n follows from 3.1.2. □

Let V be a linear space over a field k. Let K be the algebraic closure of k. A linear transformation S on V is semisimple if its minimal polynomial has simple zeros in K.

3.1.4. Theorem. *Let T be a linear transformation on a linear space V over k. Then there exist unique linear transformations S and N on V such that*

(i) *S is semisimple and N is nilpotent;*
(ii) *S and N commute;*
(iii) *$T = S + N$.*

Also, $S = P(T)$ and $N = Q(T)$ where $P, Q \in k[X]$ without constant term.

Proof. Let K be the algebraic closure of k. Let λ_i, $1 \leq i \leq n$, be the mutually different eigenvalues of T in K. Since the characteristic polynomial of T is in $k[X]$, if λ is an eigenvalue of T in K and $\sigma \in \mathrm{Aut}_k(K)$, $\sigma(\lambda)$ is also an eigenvalue of T. It follows that $\{\lambda_i, 1 \leq i \leq n\}$, is a union of $\mathrm{Aut}_k(K)$-orbits. Let $P(X) = \prod_{i=1}^n (X - \lambda_i)$. Clearly, P is in $k[X]$. Then, for some $p \in \mathbb{N}$, the characteristic polynomial of T divides P^p and $P(T)^p = 0$. By 3.1.3, for $n = p-1$, we know that there exist polynomials $A_0 = 0, A_1, \dots, A_{p-1}$ such that

$$P\left(T - \sum_{i=0}^{p-1} A_i(T)P(T)^i\right) = 0.$$

If we put

$$N = \sum_{i=0}^{p-1} A_i(T)P(T)^i$$

and

$$S = T - \sum_{i=0}^{p-1} A_i(T)P(T)^i,$$

we immediately see that S is semisimple. On the other hand, since $A_0 = 0$, we see that $N = P(T)Q(T)$ for some $Q \in k[X]$. Therefore, $N^p = P(T)^p Q(T)^p = 0$ and N is nilpotent. This proves the existence of S and N.

It remains to establish the uniqueness. Assume that S', N' is another pair of linear transformations satisfying the above conditions. Since S' and N' commute, they commute with $T = S' + N'$. On the other hand, S and N are polynomials in T, and we conclude that S' and N' commute with S and N. This implies that $S - S'$ is a semisimple linear transformation and $N - N'$ is a nilpotent linear transformation.

On the other hand, $S + N = T = S' + N'$ implies $S - S' = N' - N$. Therefore, $S - S' = N' - N = 0$. This proves the uniqueness of S and N. □

The linear transformation S is called the *semisimple part* of T and the linear transformation N is called the *nilpotent part* of T. The decomposition $T = S + N$ is called the *Jordan decomposition* of T.

Let $e_1, e_2, \ldots, e_n$ be a basis of V. Denote by E_{ij} the linear transformations on V such that $E_{ij}e_k = 0$ if $j \neq k$, and $E_{ij}e_j = e_i$, for all $1 \leq i, j, k \leq n$. Then E_{ij}, $1 \leq i, j \leq n$, form a basis of $\mathcal{L}(V)$. Let S be such that

$$Se_i = \lambda_i e_i \text{ for all } 1 \leq i \leq n.$$

Then we have

$$\begin{aligned}(\operatorname{ad} S(E_{ij}))(e_k) &= SE_{ij}e_k - E_{ij}Se_k = \delta_{jk}Se_i - \lambda_k E_{ij}e_k \\ &= \lambda_i \delta_{jk} e_i - \lambda_k \delta_{jk} e_i = (\lambda_i - \lambda_j)\delta_{jk} e_i = (\lambda_i - \lambda_j) E_{ij} e_k\end{aligned}$$

for all $1 \leq k \leq n$, i.e.,

$$\operatorname{ad} S(E_{ij}) = (\lambda_i - \lambda_j) E_{ij}$$

for all $1 \leq i, j \leq n$.

3.1.5. Lemma. *Let V be a linear space over an algebraically closed field k. Let T be a linear transformation on V and $T = S + N$ its Jordan decomposition. Then* $\operatorname{ad} T = \operatorname{ad} S + \operatorname{ad} N$ *is the Jordan decomposition of* $\operatorname{ad} T$.

Proof. Clearly, $T = S+N$ implies $\operatorname{ad} T = \operatorname{ad} S + \operatorname{ad} N$. Moreover, $[\operatorname{ad} S, \operatorname{ad} N] = \operatorname{ad}[S, N] = 0$. By 1.5.2, $\operatorname{ad} N$ is a nilpotent linear transformation. Hence, it remains to show that $\operatorname{ad} S$ is semisimple. Let $e_1, e_2, \ldots, e_n$ be a basis of V such that

$$Se_i = \lambda_i e_i \text{ for all } 1 \leq i \leq n.$$

Then

$$\operatorname{ad} S(E_{ij}) = SE_{ij} - E_{ij}S = (\lambda_i - \lambda_j)E_{ij} \text{ for all } 1 \leq i, j \leq n.$$

Hence, $\operatorname{ad} S$ is semisimple. □

Finally, we prove a result which will play the critical role in the next section.

3.1.6. Lemma. *Let V be a linear space over an algebraically closed field k. Let $U \subset W$ be two linear subspaces of $\mathcal{L}(V)$ and*

$$\mathcal{S} = \{T \in \mathcal{L}(V) \mid \operatorname{ad}(T)(W) \subset U\}.$$

If $A \in \mathcal{S}$ and $\operatorname{tr}(AB) = 0$ for every $B \in \mathcal{S}$, then A is nilpotent.

Proof. If $B \in \mathcal{S}$, clearly U and W are invariant subspaces for $\operatorname{ad}(B)$.

Let $A \in \mathcal{S}$ and $\operatorname{tr}(AB) = 0$ for all $B \in \mathcal{S}$. Let $A = S + N$ be the Jordan decomposition of A. Fix a basis $e_1, e_2, \ldots, e_n$ of V such that $Se_i = \lambda_i e_i$ for $1 \leq i \leq n$.

Let L be the linear subspace of k over the rational numbers $\mathbb{Q}$ spanned by $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let $f : L \longrightarrow \mathbb{Q}$ be a $\mathbb{Q}$-linear form on L. Let T be a linear transformation on V given by

$$Te_i = f(\lambda_i)e_i \text{ for } 1 \leq i \leq n.$$

Then

$$\operatorname{ad}(T)(E_{ij}) = (f(\lambda_i) - f(\lambda_j))E_{ij} \text{ for all } 1 \leq i, j \leq n.$$

The numbers $\lambda_i - \lambda_j$, $1 \leq i, j \leq n$, are in L. Moreover, $\lambda_i - \lambda_j = \lambda_p - \lambda_q$ for some $1 \leq i, j, p, q \leq n$, implies that

$$f(\lambda_i) - f(\lambda_j) = f(\lambda_i - \lambda_j) = f(\lambda_p - \lambda_q) = f(\lambda_p) - f(\lambda_q).$$

In addition, if $\lambda_i - \lambda_j = 0$ for some $1 \leq i, j \leq n$, i.e., we have $\lambda_i = \lambda_j$ and $f(\lambda_i) - f(\lambda_j) = 0$. Therefore, there exists a polynomial $P \in k[X]$ such that $P(\lambda_i - \lambda_j) = f(\lambda_i) - f(\lambda_j)$ for all $1 \leq i, j \leq n$, and P has no constant term. It follows that $P(\operatorname{ad} S) = \operatorname{ad} T$. On the other hand, $\operatorname{ad} S = Q(\operatorname{ad} A)$ for some polynomial Q with no constant term. Hence, $\operatorname{ad} T = (P \circ Q)(\operatorname{ad} A)$. Since $P \circ Q$ has no constant term, $\operatorname{ad}(T)(W) \subset U$, i.e, $T \in \mathcal{S}$.

This implies that $\operatorname{tr}(AT) = 0$. On the eigenspace of S for eigenvalue λ_i, T acts as multiplication by $f(\lambda_i)$. The eigenspace is also invariant for N. Therefore, it is invariant for AT and the restriction of AT to this subspace has only one eigenvalue $\lambda_i f(\lambda_i)$. Therefore, we have

$$0 = \operatorname{tr}(AT) = \sum_{i=1}^{n} \lambda_i f(\lambda_i).$$

Hence, since $f(\lambda_i) \in \mathbb{Q}$, we have

$$0 = f\left(\sum_{i=1}^{n} f(\lambda_i)\lambda_i\right) = \sum_{i=1}^{n} f(\lambda_i)^2.$$

Moreover, $f(\lambda_i)^2 \geq 0$ for all $1 \leq i \leq n$. Therefore, we conclude that $f(\lambda_i) = 0$ for all $1 \leq i \leq n$. It follows that $f = 0$. Since f is an arbitrary linear form on L, it follows that $L = \{0\}$. Therefore, $\lambda_i = 0$ for all $1 \leq i \leq n$. It follows that $S = 0$ and $A = N$, i.e., A is nilpotent. □

3.2. Cartan's criterion. In this section we prove the following solvability criterion.

3.2.1. Theorem (Cartan). *Let V be a linear space over k. Let $\mathfrak{g}$ be a Lie subalgebra of $\mathcal{L}(V)$. Define*

$$\beta(T, S) = \operatorname{tr}(TS), \text{ for } T, S \in \mathfrak{g}.$$

Then the following conditions are equivalent:

(i) *$\mathfrak{g}$ is solvable;*
(ii) *$\mathcal{D}\mathfrak{g}$ is orthogonal to $\mathfrak{g}$ with respect to β.*

Proof. Let K be the algebraic closure of k. By 2.2.3, the Lie algebra $\mathfrak{g}_K$ obtained by extension of scalars is solvable if and only if $\mathfrak{g}$ is solvable. On the other hand, $\mathfrak{g}_K$ is a Lie subalgebra of $\mathcal{L}(V)_K = \mathcal{L}(V_K)$. The bilinear form $\beta_K : (T, S) \longmapsto \operatorname{tr}(TS)$ for $T, S \in \mathcal{L}(V_K)$ is obtained from β by linear extension. Therefore, it is enough to prove the statements for Lie algebras over K.

(i)⇒(ii) By 1.6.3 we can find a basis of V such that the matrices of all $T \in \mathfrak{g}$ are upper triangular. Then the matrices of $\mathcal{D}\mathfrak{g}$ are upper triangular with zeros on the diagonal. Therefore, it follows immediately that $\beta(T, S) = 0$ for $T \in \mathfrak{g}$ and $S \in \mathcal{D}\mathfrak{g}$.

(ii)⇒(i) To prove this implication, by 2.2.7, it is enough to show that $\mathcal{D}\mathfrak{g}$ is nilpotent. By 1.5.5, $\mathcal{D}\mathfrak{g}$ is nilpotent if all $R \in \mathcal{D}\mathfrak{g}$ are nilpotent. To prove this we consider 3.1.6 for $U = \mathcal{D}\mathfrak{g}$ and $W = \mathfrak{g}$. In this case we have

$$\mathcal{S} = \{T \in \mathcal{L}(V) \mid \operatorname{ad}(T)(\mathfrak{g}) \subset \mathcal{D}\mathfrak{g}\}.$$

Clearly, $\mathfrak{g} \subset \mathcal{S}$.

Let $T \in \mathcal{S}$ and $A, B \in \mathfrak{g}$. Then $[T, A] \in \mathcal{D}\mathfrak{g}$ and

$$\begin{aligned}\operatorname{tr}(T[A,B]) &= \operatorname{tr}(TAB - TBA) = \operatorname{tr}(TAB) - \operatorname{tr}(TBA)\\ &= \operatorname{tr}(TAB) - \operatorname{tr}(ATB) = \operatorname{tr}([T,A]B) = \beta([T,A],B) = 0\end{aligned}$$

by the assumption. Hence, $\operatorname{tr}(TR) = 0$ for all $R \in \mathcal{D}\mathfrak{g}$.

It follows that $\operatorname{tr}(RT) = 0$ for all $R \in \mathcal{D}\mathfrak{g}$ and $T \in \mathcal{S}$. On the other hand, as we remarked above, $R \in \mathcal{S}$. By 3.1.6, we see that R is nilpotent. □

3.3. Radical is a characteristic ideal. The main goal of this section is to prove the following result.

3.3.1. THEOREM. *Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{r}$ its radical. Then $\mathfrak{r}$ is the orthogonal to $\mathcal{D}\mathfrak{g}$ with respect to the Killing form of $\mathfrak{g}$.*

3.3.2. COROLLARY. *The radical $\mathfrak{r}$ of a Lie algebra $\mathfrak{g}$ is a characteristic ideal.*

PROOF. Clearly, $\mathcal{D}\mathfrak{g}$ is a characteristic ideal. Therefore, by 1.1.5, $\mathcal{D}\mathfrak{g}^{\perp}$ is a characteristic ideal. By 3.3.1, $\mathfrak{r} = (\mathcal{D}\mathfrak{g})^{\perp}$. □

We first want to prove that the radical $\mathfrak{r}$ is contained in the characteristic ideal $\mathfrak{r}' = (\mathcal{D}\mathfrak{g})^{\perp}$.

We first need a technical result.

3.3.3. LEMMA. *Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{r}$ the radical of $\mathfrak{g}$. Let π be a representation of $\mathfrak{g}$ on linear space V. Then $\operatorname{tr}(\pi(x)\pi(y)) = 0$ for all $x \in \mathcal{D}\mathfrak{g}$ and $y \in \mathfrak{r}$.*

PROOF. Let K be the algebraic closure of k. Let $V_K = K \otimes_k V$. Let π_K be the representation obtained by extension of scalars from π. Then π_K is the representation of $\mathfrak{g}_K$ on V_K. By 2.2.4, $\mathfrak{r}_K$ is the radical of $\mathfrak{g}_K$. The bilinear form $\beta(x,y) = \operatorname{tr}(\pi(x)\pi(y))$, $x, y \in \mathfrak{g}$, extends by linearity to $\beta_K(x,y) = \operatorname{tr}(\pi_K(x)\pi_K(y))$ for $x, y \in \mathfrak{g}_K$. Therefore, it is enough to prove that $\beta_K(x,y) = 0$ for $x \in (\mathcal{D}\mathfrak{g})_K$ and $y \in \mathfrak{r}_K$.

Hence, we can assume from the beginning that k is algebraically closed. Assume first that π is irreducible. By 1.6.1, there exists a linear form λ on $\mathfrak{r}$ such that $\pi(y) = \lambda(y)1_V$ for all $y \in \mathfrak{r}$. Therefore, $\beta(x,y) = \operatorname{tr}(\pi(x)\pi(y)) = \lambda(y)\operatorname{tr}(\pi(x))$ for all $x \in \mathcal{D}\mathfrak{g}$ and $y \in \mathfrak{r}$. On the other hand, $\mathcal{D}\mathfrak{g}$ is spanned by commutators, hence the linear form $x \longmapsto \operatorname{tr}\pi(x)$ vanishes on $\mathcal{D}\mathfrak{g}$. It follows that $\beta(x,y) = 0$ for $x \in \mathcal{D}\mathfrak{g}$ and $y \in \mathfrak{r}$.

Assume that π is reducible. Then we prove the statement by induction in length of π. Let W be a minimal invariant subspace of V. then the representation π' of $\mathfrak{g}$ induced on W is irreducible. Let π'' be the representation of $\mathfrak{g}$ induced on V/W. Then

$$\beta(x,y) = \operatorname{tr}(\pi(x)\pi(y)) = \operatorname{tr}(\pi'(x)\pi'(y)) + \operatorname{tr}(\pi''(x)\pi''(y)) = \operatorname{tr}(\pi''(x)\pi''(y))$$

for any $x \in \mathcal{D}\mathfrak{g}$ and $y \in \mathfrak{r}$. Clearly the length of π'' is less than the length of π. Hence, by the induction assumption, $\beta(x,y) = 0$ for $x \in \mathcal{D}\mathfrak{g}$ and $y \in \mathfrak{r}$. □

Applying the lemma to the adjoint representation of $\mathfrak{g}$ we see that $\mathfrak{r} \subset \mathfrak{r}'$.

It remains to show that $\mathfrak{r}'$ is a solvable ideal in $\mathfrak{g}$. We first need a result about the Killing form.

3.3.4. LEMMA. *Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{h}$ an ideal in $\mathfrak{g}$. Then the Killing form $B_{\mathfrak{h}}$ of $\mathfrak{h}$ is the restriction of the Killing form $B_{\mathfrak{g}}$ of $\mathfrak{g}$ to $\mathfrak{h} \times \mathfrak{h}$.*

PROOF. Since $\mathfrak{h}$ is an ideal in $\mathfrak{g}$, for any $x \in \mathfrak{h}$ we have $\operatorname{ad} x(\mathfrak{g}) \subset \mathfrak{h}$. Moreover, $\operatorname{ad}(x)\operatorname{ad}(y)(\mathfrak{g}) \subset \mathfrak{h}$ for $x, y \in \mathfrak{h}$. Therefore,

$$B_{\mathfrak{g}}(x,y) = \operatorname{tr}(\operatorname{ad}(x)\operatorname{ad}(y)) = \operatorname{tr}(\operatorname{ad}_{\mathfrak{h}}(x)\operatorname{ad}_{\mathfrak{h}}(y)) = B_{\mathfrak{h}}(x,y)$$

for all $x, y \in \mathfrak{h}$. □

By 3.3.4, we have $B_{\mathfrak{r}'} = B_{\mathfrak{g}}|_{\mathfrak{r}' \times \mathfrak{r}'}$. Therefore, since $\mathfrak{r}'$ is orthogonal to $\mathcal{D}\mathfrak{g}$ with respect to $B_{\mathfrak{g}}$, we see that $\mathfrak{r}'$ is orthogonal to $\mathcal{D}\mathfrak{r}'$ for $B_{\mathfrak{r}'}$. By 3.2.1, this implies that $\operatorname{ad}\mathfrak{r}'$ is a solvable Lie subalgebra of $\mathcal{L}(\mathfrak{r}')$. On the other hand, if $\mathfrak{z}$ is the center of $\mathfrak{r}'$, we have the exact sequence

$$0 \longrightarrow \mathfrak{z} \longrightarrow \mathfrak{r}' \longrightarrow \operatorname{ad}\mathfrak{r}' \longrightarrow 0,$$

where $\mathfrak{z}$ is abelian. Therefore, by 1.2.1, $\mathfrak{r}'$ is a solvable Lie algebra. Therefore, $\mathfrak{r}'$ is a solvable ideal in $\mathfrak{g}$. It follows that $\mathfrak{r}' \subset \mathfrak{r}$. Hence, it follows that $\mathfrak{r} = \mathfrak{r}'$ which completes the proof of 3.3.1.

4. Semisimple Lie algebras

In this section we generalize some results about Lie algebras of compact semisimple Lie groups from 3.1.8.

4.1. Killing form and semisimple Lie algebras. The following result gives a new characterization of semisimple Lie algebras.

4.1.1. THEOREM. *Let $\mathfrak{g}$ be a Lie algebra. Then the following conditions are equivalent:*

(i) $\mathfrak{g}$ *is semisimple;*
(ii) *the Killing form B of $\mathfrak{g}$ is nondegenerate.*

If these conditions hold, $\mathfrak{g} = \mathcal{D}\mathfrak{g}$.

PROOF. (i)⇒(ii) Assume that $\mathfrak{g}$ is semisimple. Then the radical $\mathfrak{r}$ of $\mathfrak{g}$ is equal to $\{0\}$. By 3.3.1, $(\mathcal{D}\mathfrak{g})^{\perp} = \{0\}$. Moreover, $\mathcal{D}\mathfrak{g} \subseteq \mathfrak{g}$ implies $\mathfrak{g}^{\perp} \subseteq (\mathcal{D}\mathfrak{g})^{\perp} = \{0\}$. It follows that $\mathfrak{g}^{\perp} = \{0\}$, i.e., the Killing form B is nondegenerate.

In addition, in this situation $\mathfrak{g}^{\perp} = \{0\} = (\mathcal{D}\mathfrak{g})^{\perp}$ implies $\mathfrak{g} = \mathcal{D}\mathfrak{g}$.

(ii)⇒(i) Let $\mathfrak{a}$ be an abelian ideal in $\mathfrak{g}$. Let $x \in \mathfrak{g}$ and $y \in \mathfrak{a}$. Then $\operatorname{ad}(y)(\mathfrak{g}) \subset \mathfrak{a}$, and $\operatorname{ad}(x)\operatorname{ad}(y)(\mathfrak{g}) \subset \mathfrak{a}$. Therefore, $(\operatorname{ad}(x)\operatorname{ad}(y))^2(\mathfrak{g}) \subset \operatorname{ad}(x)\operatorname{ad}(y)(\mathfrak{a}) = \{0\}$, and $\operatorname{ad}(x)\operatorname{ad}(y)$ is a nilpotent linear transformation on $\mathfrak{g}$. This implies that $B(x,y) = \operatorname{tr}(\operatorname{ad}(x)\operatorname{ad}(y)) = 0$. Hence, $\mathfrak{a} \subset \mathfrak{g}^{\perp}$. Since B is nondegenerate, we see that $\mathfrak{a} = \{0\}$. By 1.3.1, $\mathfrak{g}$ is a semisimple Lie algebra. □

Let $\mathfrak{g}$ be a semisimple Lie algebra. Let $\mathfrak{a}$ be an ideal in $\mathfrak{g}$. Let $\mathfrak{a}^{\perp}$ be the orthogonal to $\mathfrak{a}$ with respect to the Killing form B of $\mathfrak{g}$. Then, by 1.1.5, $\mathfrak{a}^{\perp}$ is an ideal in $\mathfrak{g}$. This implies that $\mathfrak{b} = \mathfrak{a} \cap \mathfrak{a}^{\perp}$ is an ideal in $\mathfrak{g}$. If $x, y \in \mathfrak{b}$, we have $B(x,y) = 0$. By 3.2.1, we see that $\operatorname{ad}\mathfrak{b}$ is a solvable Lie algebra. Since ad is injective by 1.3.2, we conclude that $\mathfrak{b}$ is solvable. Therefore, $\mathfrak{b} = \{0\}$ and $\mathfrak{a} \cap \mathfrak{a}^{\perp} = \{0\}$. Since B is nondegenerate, $\dim \mathfrak{a}^{\perp} = \dim \mathfrak{g} - \dim \mathfrak{a}$, i.e., $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^{\perp}$ as a linear space. This in turn implies the following result.

4.1.2. LEMMA. *Let $\mathfrak{g}$ be a semisimple Lie algebra and $\mathfrak{a}$ an ideal in $\mathfrak{g}$. Then* $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$.

Moreover, $\mathfrak{a}$ and $\mathfrak{a}^\perp$ are semisimple ideals in $\mathfrak{g}$.

PROOF. It is enough to prove that $\mathfrak{a}$ is semisimple. Let $x \in \mathfrak{a}$ be such that $B_\mathfrak{a}(x, y) = 0$ for all $y \in \mathfrak{a}$. Then, by 3.3.4, $B_\mathfrak{g}(x, y) = 0$ for all $y \in \mathfrak{a}$. This in turn implies that $B_\mathfrak{g}(x, y) = 0$ for all $y \in \mathfrak{g}$. Since $B_\mathfrak{g}$ is nondegenerate, $x = 0$. This implies that $B_\mathfrak{a}$ is nondegenerate, and $\mathfrak{a}$ is semisimple by 4.1.1. □

A Lie algebra is *simple* if it is not abelian and it has no nontrivial ideals. By 1.3.1, a simple Lie algebra is semisimple.

A minimal ideal $\mathfrak{a}$ in a semisimple Lie algebra cannot be abelian by 1.3.1. On the other hand, by 4.1.2, any ideal in $\mathfrak{a}$ is an ideal in $\mathfrak{g}$. Hence, by minimality, $\mathfrak{a}$ has to be simple. Therefore any semisimple Lie algebra contains a simple ideal.

4.1.3. LEMMA. *Let $\mathfrak{g}$ be a semisimple Lie algebra and π a representation of $\mathfrak{g}$ on a linear space V. Then $\pi(\mathfrak{g})$ is contained in the Lie algebra $\mathfrak{sl}(V)$ of all traceless linear transformations on V.*

PROOF. Let $x, y \in \mathfrak{g}$. Then

$$\operatorname{tr}(\pi([x,y])) = \operatorname{tr}([\pi(x), \pi(y)]) = \operatorname{tr}(\pi(x)\pi(y)) - \operatorname{tr}(\pi(y)\pi(x)) = 0.$$

Therefore, the linear form $x \longmapsto \operatorname{tr}\pi(x)$ vanishes on $\mathcal{D}\mathfrak{g}$. By 4.1.1, it vanishes on $\mathfrak{g}$. □

4.2. Derivations are inner. The next result is a generalization of the result about the nondegeneracy of the Killing form.

4.2.1. LEMMA. *Let $\mathfrak{g}$ be a semisimple Lie algebra and π a faithful representation of $\mathfrak{g}$. Then the bilinear form $(x, y) \longmapsto \beta(x, y) = \operatorname{tr}(\pi(x)\pi(y))$ is nondegenerate on $\mathfrak{g}$.*

PROOF. Let

$$\mathfrak{s} = \{x \in \mathfrak{g} \mid \beta(x, y) = 0 \text{ for all } y \in \mathfrak{g}\}.$$

Let $x \in \mathfrak{s}$ and $y \in \mathfrak{g}$. Then

$$\begin{aligned}\beta([y,x], z) = \operatorname{tr}(\pi([y,x])\pi(z)) = \operatorname{tr}(\pi(y)\pi(x)\pi(z)) - \operatorname{tr}(\pi(x)\pi(y)\pi(z))\\ = \operatorname{tr}(\pi(x)\pi([y,z])) = 0,\end{aligned}$$

for all $z \in \mathfrak{g}$. Therefore, $[y, x] \in \mathfrak{s}$ for all $y \in \mathfrak{g}$. Hence, $\mathfrak{s}$ is an ideal in $\mathfrak{g}$. Moreover, $\mathfrak{s}$ is orthogonal onto itself with respect to β. By 3.2.1, this implies that $\pi(\mathfrak{s})$ is solvable. Since the π is faithful, this implies that $\mathfrak{s}$ is solvable. Hence, $\mathfrak{s} = \{0\}$. This implies that β is nondegenerate. □

This has the following consequence which generalizes 4.1.2.

4.2.2. LEMMA. *Let $\mathfrak{g}$ be a Lie algebra and B its Killing form. Let $\mathfrak{a}$ be a semisimple Lie subalgebra of $\mathfrak{g}$. Then the orthogonal $\mathfrak{h} = \mathfrak{a}^\perp$ of $\mathfrak{a}$ is a direct complement to $\mathfrak{a}$ in $\mathfrak{g}$ and $\operatorname{ad}(x)(\mathfrak{h}) \subset \mathfrak{h}$ for all $x \in \mathfrak{a}$.*

If $\mathfrak{a}$ is an ideal in $\mathfrak{g}$, $\mathfrak{h}$ is an ideal in $\mathfrak{g}$ and

$$\mathfrak{h} = \{x \in \mathfrak{g} \mid \operatorname{ad}(x)(\mathfrak{a}) = \{0\}\}.$$

In particular, $\mathfrak{g} = \mathfrak{a} \times \mathfrak{h}$.

PROOF. Since $\mathfrak{a}$ is semisimple, its center is equal to $\{0\}$ and $\operatorname{ad}_{\mathfrak{g}} : \mathfrak{a} \longrightarrow \mathcal{L}(\mathfrak{g})$ is faithful. Hence, by 4.2.1, $B_{\mathfrak{g}}|_{\mathfrak{a}\times\mathfrak{a}}$ is nondegenerate. Therefore, $\mathfrak{a} \cap \mathfrak{h} = \{0\}$. Moreover, we have $\dim \mathfrak{h} = \dim \mathfrak{g} - \dim \mathfrak{a}$, and $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{h}$ as a linear space. In addition, for $x \in \mathfrak{a}$ and $y \in \mathfrak{h}$, we have $B([x, y], z) = -B(y, [x, z]) = 0$ for all $z \in \mathfrak{a}$. Hence $[x, y] \in \mathfrak{h}$. It follows that $\mathfrak{h}$ is invariant for all $\operatorname{ad} x$, $x \in \mathfrak{a}$.

If $\mathfrak{a}$ is an ideal, $\mathfrak{h}$ is an ideal by 1.1.5. Hence, $\mathfrak{g} = \mathfrak{a} \times \mathfrak{h}$, and since $\mathfrak{a}$ has trivial center, the rest of the statement follows. □

The next result says that all derivations of a semisimple Lie algebra are inner.

4.2.3. PROPOSITION. *Let $\mathfrak{g}$ be a semisimple Lie algebra. Then* $\operatorname{ad} : \mathfrak{g} \longrightarrow \operatorname{Der}(\mathfrak{g})$ *is an isomorphism of Lie algebras.*

PROOF. Since the center of $\mathfrak{g}$ is $\{0\}$, ad is injective and $\mathfrak{g}$ is isomorphic to the ideal $\operatorname{ad} \mathfrak{g}$ of inner derivations in $\operatorname{Der}(\mathfrak{g})$ by 1.1.3.

By 4.2.2, $\operatorname{Der}(\mathfrak{g}) = \operatorname{ad} \mathfrak{g} \times \mathfrak{h}$, where

$$\mathfrak{h} = \{D \in \operatorname{Der}(\mathfrak{g}) \mid [D, \operatorname{ad} x] = 0 \text{ for all } x \in \mathfrak{g}\}.$$

Let $D \in \mathfrak{h}$. Then, by 1.1.2, we have $\operatorname{ad}(Dx) = [D, \operatorname{ad} x] = 0$ for all $x \in \mathfrak{g}$. Since ad is injective, $Dx = 0$ for all $x \in \mathfrak{g}$, and $D = 0$. Hence, $\mathfrak{h} = \{0\}$. □

4.3. Decomposition into product of simple ideals. Let $\mathfrak{g}$ be a semisimple Lie algebra and $\mathfrak{a}$ an ideal in $\mathfrak{g}$. Then $\mathfrak{a}$ is semisimple by 4.1.2. Assume that $\mathfrak{b}$ is another ideal in $\mathfrak{g}$ such that $\mathfrak{a} \cap \mathfrak{b} = \{0\}$. Let $x \in \mathfrak{a}$ and $y \in \mathfrak{b}$. Then, $\operatorname{ad}(y)(\mathfrak{g}) \subset \mathfrak{b}$ and

$$(\operatorname{ad}(x) \operatorname{ad}(y))(\mathfrak{g}) = \operatorname{ad}(x)(\operatorname{ad}(y)(\mathfrak{g})) \subset \operatorname{ad}(x)(\mathfrak{b}) \subset \mathfrak{a} \cap \mathfrak{b} = \{0\}.$$

Therefore, $\operatorname{ad}(x) \operatorname{ad}(y) = 0$ and $B(x, y) = \operatorname{tr}(\operatorname{ad}(x) \operatorname{ad}(y)) = 0$. It follows that $\mathfrak{b} \subset \mathfrak{a}^\perp$.

In particular, if $\mathfrak{b}$ is a direct complement of $\mathfrak{a}$, we must have $\mathfrak{b} = \mathfrak{a}^\perp$. Therefore, the complementary ideal is unique.

The set of all ideals in $\mathfrak{g}$ is ordered by inclusion. Let $\mathfrak{m}$ be a minimal ideal in $\mathfrak{g}$. As we remarked in the preceding section $\mathfrak{m}$ is a simple ideal.

Let $\mathfrak{a}$ be another ideal in $\mathfrak{g}$. Then $\mathfrak{a} \cap \mathfrak{m}$ is an ideal in $\mathfrak{g}$. By the minimality of $\mathfrak{m}$, we have either $\mathfrak{m} \subset \mathfrak{a}$ or $\mathfrak{a} \cap \mathfrak{m} = \{0\}$. By the above discussion, the latter implies that $\mathfrak{a} \subset \mathfrak{m}^\perp$, i.e., $\mathfrak{a}$ is perpendicular to $\mathfrak{m}$.

Let $\mathfrak{m}_1, \mathfrak{m}_2, \ldots, \mathfrak{m}_p$ be a family of mutually different minimal ideals in $\mathfrak{g}$. By the above discussion $\mathfrak{m}_i$ is perpendicular to $\mathfrak{m}_j$ for $i \neq j$, $1 \leq i, j \leq p$. Hence, p has to be smaller than $\dim \mathfrak{g}$. Assume that p is maximal possible. Then $\mathfrak{a} = \mathfrak{m}_1 \oplus \mathfrak{m}_2 \oplus \cdots \oplus \mathfrak{m}_p$ is an ideal in $\mathfrak{g}$. Assume that $\mathfrak{a} \neq \mathfrak{g}$. Then $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$. Let $\mathfrak{m}_{p+1}$ be a minimal ideal in $\mathfrak{a}^\perp$. Then $\mathfrak{m}_{p+1}$ is a minimal ideal in $\mathfrak{g}$ different from $\mathfrak{m}_i$, $1 \leq i \leq p$, contradicting the maximality of p. It follows that $\mathfrak{g} = \mathfrak{m}_1 \oplus \mathfrak{m}_2 \oplus \cdots \oplus \mathfrak{m}_p$, i.e., we have the following result.

4.3.1. THEOREM. *Any semisimple Lie algebra $\mathfrak{g}$ is the direct product of its minimal ideals. These ideals are simple Lie algebras.*

In particular, a semisimple Lie algebra is a product of simple Lie algebras.

4.4. Jordan decomposition in semisimple Lie algebras. In this section we prove a version of Jordan decomposition for semisimple Lie algebras.

Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{a}$ a Lie subalgebra of $\mathfrak{g}$. Let

$$\mathfrak{n} = \{x \in \mathfrak{g} \mid \operatorname{ad}(x)(\mathfrak{a}) \subset \mathfrak{a}\}.$$

Clearly, $\mathfrak{n}$ is a Lie subalgebra of $\mathfrak{g}$ and $\mathfrak{a}$ is an ideal in $\mathfrak{n}$. This Lie algebra is called the *normalizer* of $\mathfrak{a}$ in $\mathfrak{g}$. Clearly, the normalizer is the largest Lie subalgebra of $\mathfrak{g}$ which contains $\mathfrak{a}$ as an ideal.

4.4.1. THEOREM. *Let $\mathfrak{g}$ be a semisimple Lie algebra. Let $x \in \mathfrak{g}$. Then there exist unique elements $s, n \in \mathfrak{g}$ such that*

(i) $\operatorname{ad} s$ *is semisimple and* $\operatorname{ad} n$ *is nilpotent;*
(ii) $[s, n] = 0$*;*
(iii) $x = s + n$*.*

The element s is the *semisimple part* of x and n is the *nilpotent part* of x. The decomposition $x = s + n$ is called the *Jordan decomposition* of x in $\mathfrak{g}$.

PROOF. Assume first that k is algebraically closed. Since the center of $\mathfrak{g}$ is equal to zero, the adjoint representation $\operatorname{ad} : \mathfrak{g} \longrightarrow \mathcal{L}(\mathfrak{g})$ is injective. Therefore, we can view $\mathfrak{g}$ as a Lie subalgebra $\operatorname{ad} \mathfrak{g}$ in $\mathcal{L}(\mathfrak{g})$.

Let $\mathfrak{n}$ be the normalizer of $\mathfrak{g}$ in $\mathcal{L}(\mathfrak{g})$, i.e.,

$$\mathfrak{n} = \{T \in \mathcal{L}(\mathfrak{g}) \mid \operatorname{ad}(T)(\operatorname{ad} \mathfrak{g}) \subset \operatorname{ad} \mathfrak{g}\}.$$

Clearly, $\mathfrak{n}$ is a Lie subalgebra of $\mathcal{L}(\mathfrak{g})$, and $\operatorname{ad} \mathfrak{g}$ is an ideal in $\mathfrak{n}$. Let $T \in \mathfrak{n}$. Let $T = S + N$ be the Jordan decomposition of the linear transformation T. Then, by 3.1.5, $\operatorname{ad} T = \operatorname{ad} S + \operatorname{ad} N$. Moreover, by 3.1.4, $\operatorname{ad} S$ and $\operatorname{ad} N$ are polynomials in $\operatorname{ad} T$. Hence, $S, N \in \mathfrak{n}$.

By 4.2.2, we have $\mathfrak{n} = \mathfrak{a} \times \mathfrak{g}$, where

$$\mathfrak{a} = \{T \in \mathfrak{n} \mid \operatorname{ad}(T)(\operatorname{ad} \mathfrak{g}) = \{0\}\}.$$

Let

$$\mathfrak{n}' = \{T \in \mathfrak{n} \mid T(\mathfrak{h}) \subset \mathfrak{h} \text{ for any ideal } \mathfrak{h} \text{ of } \mathfrak{g}\}.$$

Since $\operatorname{ad} \mathfrak{g} \subset \mathfrak{n}'$, we have

$$\mathfrak{n}' = (\mathfrak{a} \cap \mathfrak{n}') \times \mathfrak{g},$$

i.e., $\mathfrak{n}' = \mathfrak{a}' \times \mathfrak{g}$ where $\mathfrak{a}' = \mathfrak{a} \cap \mathfrak{n}'$.

Let $T \in \mathfrak{n}'$ and let $T = S + N$ be its Jordan decomposition. As we already remarked, $S, N \in \mathfrak{n}$. Moreover, by 3.1.4, S and N are polynomials in T, so $S(\mathfrak{h}) \subset \mathfrak{h}$ and $N(\mathfrak{h}) \subset \mathfrak{h}$ for all ideals $\mathfrak{h} \subset \mathfrak{g}$. It follows that $S, N \in \mathfrak{n}'$.

Let $\mathfrak{g} = \mathfrak{m}_1 \times \mathfrak{m}_2 \times \cdots \times \mathfrak{m}_p$ be the decomposition of $\mathfrak{g}$ into a product of simple ideals. Then for $T \in \mathfrak{a}'$ we have $T(\mathfrak{m}_i) \subset \mathfrak{m}_i$ for $1 \le i \le p$. On the other hand, T commutes with $\operatorname{ad} x$ for any $x \in \mathfrak{g}$. Let $1 \le i \le p$ and λ_i be an eigenvalue of the restriction of T to $\mathfrak{m}_i$. Let $y \in \ker(T - \lambda_i I) \cap \mathfrak{m}_i$. Then

$$(T - \lambda_i I)([x, y]) = (T - \lambda_i I) \operatorname{ad}(x) y = \operatorname{ad}(x)(T - \lambda_i I) y = 0$$

for any $x \in \mathfrak{g}$. Hence, $[x, y] \in \ker(T - \lambda_i I) \cap \mathfrak{m}_i$ for any $x \in \mathfrak{g}$, i.e., $\ker(T - \lambda_i I) \cap \mathfrak{m}_i$ is an ideal in $\mathfrak{m}_i$. Since $\mathfrak{m}_i$ is minimal, $\ker(T - \lambda_i I) \supset \mathfrak{m}_i$, i.e., $T|_{\mathfrak{m}_i}$ is multiplication by λ_i. This implies that T is semisimple.

Let N be a nilpotent linear transformation in $\mathfrak{n}'$. Then $N = P + Q$ for $P \in \mathfrak{a}'$ and $Q \in \operatorname{ad} \mathfrak{g}$. By the above argument, P is semisimple. Since $\mathfrak{m}_i$ are invariant for N, P, Q for all $1 \le i \le p$, we have

$$N|_{\mathfrak{m}_i} = P|_{\mathfrak{m}_i} + Q|_{\mathfrak{m}_i}$$

for all $1 \le i \le p$. Since N is nilpotent, $N|_{\mathfrak{m}_i}$ is nilpotent and

$$0 = \operatorname{tr}(N|_{\mathfrak{m}_i}) = \operatorname{tr}(P|_{\mathfrak{m}_i}) + \operatorname{tr}(Q|_{\mathfrak{m}_i}).$$

The ideal $\mathfrak{m}_i$ is invariant for the adjoint representation of $\mathfrak{g}$. Hence, by 4.1.3, we have $\operatorname{tr}(Q|_{\mathfrak{m}_i}) = 0$. This in turn implies that $\operatorname{tr}(P|_{\mathfrak{m}_i}) = 0$. On the other hand, the above argument shows that $P|_{\mathfrak{m}_i}$ is a multiple of the identity. Hence, $P|_{\mathfrak{m}_i} = 0$ for all $1 \le i \le p$. It follows that $P = 0$, and $N \in \operatorname{ad}\mathfrak{g}$.

Let $x \in \mathfrak{g}$ and let $\operatorname{ad} x = S + N$ be the Jordan decomposition of $\operatorname{ad} x$ in $\mathcal{L}(\mathfrak{g})$. By the above remarks, S and N are in $\mathfrak{n}'$. By the above argument, $N = \operatorname{ad} n$ for some $n \in \mathfrak{g}$. This implies that $s = x - n \in \mathfrak{g}$ and

$$\operatorname{ad} s = \operatorname{ad} x - \operatorname{ad} n = \operatorname{ad} x - N = S$$

is semisimple. Finally,

$$\operatorname{ad}[s, n] = [\operatorname{ad} s, \operatorname{ad} n] = [S, N] = 0$$

and $[s, n] = 0$. This proves the existence of the decomposition.

Assume that $s', n' \in \mathfrak{g}$ satisfy $x = s' + n'$, $[s', n'] = 0$ and $\operatorname{ad} s'$ is semisimple and $\operatorname{ad} n'$ is nilpotent. Then $\operatorname{ad} x = \operatorname{ad} s' + \operatorname{ad} n'$ is the Jordan decomposition of $\operatorname{ad} x$ in $\mathcal{L}(\mathfrak{g})$. Therefore, by the uniqueness of that decomposition, $\operatorname{ad} s' = S = \operatorname{ad} s$ and $\operatorname{ad} n' = N = \operatorname{ad} n$. This in turn implies that $s' = s$ and $n' = n$.

Finally, assume that k is not algebraically closed. Let K be the algebraic closure of k. Let $\mathfrak{g}_K$ be the algebra obtained from $\mathfrak{g}$ by the extension of the field of scalars. Then $\mathfrak{g}_K$ is semisimple by 2.2.5. Let $x \in \mathfrak{g} \subset \mathfrak{g}_K$. Let $x = s + n$ be the Jordan decomposition of x in $\mathfrak{g}_K$. Since x is stable under the action of $\operatorname{Aut}_k(K)$, for any $\sigma \in \operatorname{Aut}_k(K)$, we have $x = \sigma_{\mathfrak{g}_K}(s) + \sigma_{\mathfrak{g}_K}(n)$, $[\sigma_{\mathfrak{g}_K}(s), \sigma_{\mathfrak{g}_K}(n)] = \sigma_{\mathfrak{g}_K}([s, n]) = 0$, $\operatorname{ad}\sigma_{\mathfrak{g}_K}(s)$ is semisimple and $\operatorname{ad}\sigma_{\mathfrak{g}_K}(n)$ is nilpotent. Therefore, by the uniqueness of the Jordan decomposition, we have $\sigma_{\mathfrak{g}_K}(s) = s$ and $\sigma_{\mathfrak{g}_K}(n) = n$ for any $\sigma \in \operatorname{Aut}_k(K)$. Therefore, s and n are in $\mathfrak{g}$. □

4.5. Lie algebra $\mathfrak{sl}(n, k)$. Let V be a linear space over the field k and V^* its linear dual. We can define a bilinear map $\varphi : V \times V^* \longrightarrow \mathcal{L}(V)$ by $\varphi(v, f)(w) = f(w)v$ for any $v, w \in V$ and $f \in V^*$. This map defines the linear map $\Phi : V \otimes_k V^* \longrightarrow \mathcal{L}(V)$ such that $\Phi(v \otimes f)(w) = f(w)v$ for any $v, w \in V$ and $f \in V^*$.

4.5.1. Lemma. *The linear map $\Phi : V \otimes_k V^* \longrightarrow \mathcal{L}(V)$ is a linear isomorphism.*

Proof. Clearly, we have

$$\dim(V \otimes V^*) = \dim V \dim V^* = (\dim V)^2 = \dim \mathcal{L}(V).$$

Therefore, it is enough to show that Φ is injective. Let $v_1, v_2, \dots, v_n$ be a basis of V and $f_1, f_2, \dots, v_n$ the dual basis of V^*. Then $v_i \otimes f_j$, $1 \le i, j \le n$, is a basis of $V \otimes_k V^*$. Let $\Phi(z) = 0$ for some $z = \sum_{i,j=1}^n \alpha_{ij} v_i \otimes f_j \in V \otimes_k V^*$. Then

$$0 = \Phi(z)(v_k) = \Phi\left(\sum_{i,j=1}^n \alpha_{ij} v_i \otimes f_j\right)(v_k) = \sum_{i,j=1}^n \alpha_{ij} f_j(v_k) v_i = \sum_{i=1}^n \alpha_{ik} v_i$$

for any $1 \le k \le n$. Therefore, $\alpha_{ij} = 0$ for all $1 \le i, j \le n$. □

Let V and W be two linear spaces over k. Let $S \in \mathcal{L}(V)$ and $T \in \mathcal{L}(W)$. Then they define a bilinear map $(v, w) \longmapsto Sv \otimes Tw$ from $V \times W$ into $V \otimes_k W$. This bilinear map induces a linear endomorphism $S \otimes T$ of $V \otimes_k W$ given by

$$(S \otimes T)(v \otimes w) = Sv \otimes Tw$$

for any $v \in V$ and $w \in W$.

Let T be a linear transformation on V. Then T acts on $\mathcal{L}(V)$ by left (resp. right) multiplication. Then we have the following commutative diagram

$$\begin{array}{ccc} V \otimes V^* & \xrightarrow{T \otimes I} & V \otimes V^* \\ {\scriptstyle\Phi}\big\downarrow & & \big\downarrow{\scriptstyle\Phi} \\ \mathcal{L}(V) & \xrightarrow[T\circ -]{} & \mathcal{L}(V) \end{array}$$

since

$$T\Phi(v \otimes f)(w) = T(f(w)v) = f(w)Tv = \Phi(Tv \otimes f)(w) = \Phi((T \otimes I)(v \otimes f))(w)$$

for all $v, w \in V$ and $f \in V^*$. Also, we have the following commutative diagram

$$\begin{array}{ccc} V \otimes V^* & \xrightarrow{I \otimes T^*} & V \otimes V^* \\ {\scriptstyle\Phi}\big\downarrow & & \big\downarrow{\scriptstyle\Phi} \\ \mathcal{L}(V) & \xrightarrow[-\circ T]{} & \mathcal{L}(V) \end{array}$$

since

$$\Phi(v \otimes f)T(w) = f(Tw)v = (T^*f)(w)v = \Phi(v \otimes T^*f)(w) = \Phi((I \otimes T^*)(v \otimes f))(w)$$

for all $v, w \in V$ and $f \in V^*$.

Therefore, we finally conclude that the we have the commutative diagram

$$\begin{array}{ccc} V \otimes V^* & \xrightarrow{T \otimes I - I \otimes T^*} & V \otimes V^* \\ {\scriptstyle\Phi}\big\downarrow & & \big\downarrow{\scriptstyle\Phi} \\ \mathcal{L}(V) & \xrightarrow[\operatorname{ad} T]{} & \mathcal{L}(V) \end{array} .$$

Therefore, the adjoint representation of $\mathcal{L}(V)$ is equivalent to the representation $T \longmapsto T \otimes I - I \otimes T^*$ on $V \otimes_k V^*$.

4.5.2. Lemma. *Let V and W be two linear spaces over k. Let $S \in \mathcal{L}(V)$ and $T \in \mathcal{L}(W)$. Then*

$$\operatorname{tr}(T \otimes S) = \operatorname{tr}(T)\operatorname{tr}(S).$$

Proof. Let $v_1, v_2, \dots v_n$ be a basis of V and $w_1, w_2, \dots, w_m$ a basis of W. Then $v_i \otimes w_j$, $1 \le i \le n$, $1 \le j \le m$, is a basis of $V \otimes_k W$. If α_{ij} and β_{pq} are the matrix entries of these transformations in these bases, we have

$$(S \otimes T)(v_i \otimes w_j) = Sv_i \otimes Tw_j = \sum_{p=1}^{n}\sum_{q=1}^{m} \alpha_{pi}\beta_{qj} v_p \otimes w_q.$$

Therefore,

$$\operatorname{tr}(S \otimes T) = \sum_{i=1}^{n}\sum_{j=1}^{q} \alpha_{ii}\beta_{jj} = \operatorname{tr}(S)\operatorname{tr}(T).$$

□

If $n = \dim V$, this implies that the Killing form on $\mathcal{L}(V)$ is given by

$$\begin{aligned} B(S,T) &= \operatorname{tr}(\operatorname{ad}(T)\operatorname{ad}(S)) = \operatorname{tr}((S \otimes I - I \otimes S^*)(T \otimes I - I \otimes T^*)) \\ &= \operatorname{tr}(ST \otimes I) - \operatorname{tr}(S \otimes T^*) - \operatorname{tr}(T \otimes S^*) + \operatorname{tr}(I \otimes S^*T^*) \\ &= n\operatorname{tr}(ST) - \operatorname{tr}(S)\operatorname{tr}(T^*) - \operatorname{tr}(T)\operatorname{tr}(S^*) + n\operatorname{tr}(S^*T^*) = 2n\operatorname{tr}(ST) - 2\operatorname{tr}(S)\operatorname{tr}(T) \end{aligned}$$

for $S, T \in \mathcal{L}(V)$.

On the other hand, let $\mathfrak{sl}(V)$ be the ideal of $\mathcal{L}(V)$ consisting of all traceless linear transformations on V. Then, by 3.3.4, we have the following result.

4.5.3. LEMMA. *Let V be a n-dimensional linear space over k. The Killing form on the algebra $\mathfrak{sl}(V)$ is given by $B(S,T) = 2n\ \operatorname{tr}(ST)$ for $S, T \in \mathfrak{sl}(V)$.*

This has the following direct consequence.

4.5.4. PROPOSITION. *Let $n \geq 2$. The Lie algebra $\mathfrak{sl}(n,k)$ of all $n \times n$ traceless matrices is semisimple.*

PROOF. By 4.1.1, it is enough to show that the Killing form is nondegenerate on $\mathfrak{sl}(n,k)$.

Let $T \in \mathfrak{sl}(n,k)$ be such that $B(T,S) = 0$ for all $S \in \mathfrak{sl}(n,k)$. Let E_{ij} be the matrix with all entries equal to zero except the entry in i^{th} row and j^{th} column. Then E_{ij}, $i \neq j$, are in $\mathfrak{sl}(n,k)$. Moreover, if we denote by t_{ij} the matrix entries of T, we have $0 = B(T, E_{ij}) = 2nt_{ji}$. Hence, $T_{ji} = 0$. Hence, T must be diagonal. On the other hand $E_{ii} - E_{jj} \in \mathfrak{sl}(n,k)$ for $1 \leq i,j \leq n$. Also, $0 = B(T, E_{ii} - E_{jj}) = 2n(t_{ii} - t_{jj})$ for all $1 \leq i,j \leq n$. Hence T is a multiple of the identity matrix. Since $\operatorname{tr}(T) = 0$, we must have $T = 0$. It follows that B is nondegenerate. □

4.6. Three-dimensional Lie algebras. In this section we want to classify all three dimensional Lie algebras over an algebraically closed field k. We start with the following observation.

4.6.1. LEMMA. *Let $\mathfrak{g}$ be a three-dimensional Lie algebra. Then $\mathfrak{g}$ is either solvable or simple Lie algebra.*

PROOF. Assume that $\mathfrak{g}$ is not solvable. Let $\mathfrak{r}$ be the radical of $\mathfrak{g}$. Then $\mathfrak{r} \neq \mathfrak{g}$. Therefore, $\mathfrak{g}/\mathfrak{r}$ is a Lie algebra of dimension 1, 2 or 3. By 1.3.3, $\mathfrak{g}/\mathfrak{r}$ is semisimple. Since all Lie algebras of dimension 1 and 2 are solvable, $\mathfrak{g}/\mathfrak{r}$ must be three-dimensional, i.e., $\mathfrak{r} = \{0\}$. Let $\mathfrak{h}$ be a nonzero ideal in $\mathfrak{g}$. Then its dimension is either 1, 2 or 3. Since the ideals of dimension 1 or 2 have to be solvable, this contradicts the fact that $\mathfrak{g}$ is semisimple. Therefore, $\mathfrak{h} = \mathfrak{g}$, i.e., $\mathfrak{g}$ is simple. □

We are going to classify the three-dimensional Lie algebras by discussing the possible cases of $\dim \mathcal{D}\mathfrak{g}$.

First, if $\mathcal{D}\mathfrak{g} = \{0\}$, $\mathfrak{g}$ is abelian.

Consider now the case $\dim \mathcal{D}\mathfrak{g} = 1$. Let $\mathfrak{z}$ be the center of $\mathfrak{g}$. Then there are two subcases.

First, assume that $\mathcal{D}\mathfrak{g} \subset \mathfrak{z}$. Then, we can pick $e \in \mathcal{D}\mathfrak{g}$, $e \neq 0$, which spans $\mathcal{D}\mathfrak{g}$. Since $e \in \mathfrak{z}$, there are $f, g \in \mathfrak{g}$ such that (e, f, g) is a basis of $\mathfrak{g}$ and $[e,f] = [e,g] = 0$. Finally, $[f,g] = \lambda e$ with $\lambda \in k$. The number λ must be different from 0 since $\mathfrak{g}$ is not abelian. By replacing f with $\frac{1}{\lambda}f$, we get that $[f,g] = e$. Therefore, there exists at most one three-dimensional Lie algebra with the above properties. On the other hand, Let $\mathfrak{g} = \mathfrak{n}(3, k)$ be the Lie algebra upper triangular nilpotent matrices in $M_3(k)$. Then $\mathfrak{g}$ is three-dimensional, and its basis

$$e = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad g = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

satisfies the above commutation relations. This shows the existence of the above Lie algebra. By 1.5.6, this is a nilpotent Lie algebra.

In the second subcase, we assume that $\mathcal{D}\mathfrak{g} \cap \mathfrak{z} = \{0\}$. Let $e \in \mathcal{D}\mathfrak{g}$, $e \neq 0$. Since e is not in the center, there exists $f \in \mathfrak{g}$ such that $[e, f] = \lambda e$ with $\lambda \neq 0$. By replacing f with $\frac{1}{\lambda} f$, we can assume that $[e, f] = e$. Therefore, the Lie algebra $\mathfrak{h}$ spanned by e, f is the two-dimensional nonabelian Lie algebra from 1.2.2. Since $\mathcal{D}\mathfrak{g} \subset \mathfrak{h}$, $\mathfrak{h}$ is an ideal in $\mathfrak{g}$.

Let $g \in \mathfrak{g}$ be a vector outside $\mathfrak{h}$. Then $[g, e] = ae$ and $[g, f] = be$. This in turn implies that

$$[g + \lambda e + \mu f, e] = ae - \mu e = (a - \mu)e \text{ and } [g + \lambda e + \mu f, f] = be + \lambda e = (b + \lambda)e.$$

Hence, if we put $\lambda = -b$ and $\mu = a$, and replace g with $g + \lambda e + \mu f$, we see that $[e, g] = [f, g] = 0$. Therefore, g spans an abelian ideal complementary to $\mathfrak{h}$. Therefore, $\mathfrak{g}$ is the product of $\mathfrak{h}$ and a one-dimensional abelian Lie algebra. Therefore, this is a solvable Lie algebra, which is not nilpotent.

Now we want to study the case of $\dim \mathcal{D}\mathfrak{g} = 2$. There are two subcases, the two-dimensional Lie subalgebra $\mathfrak{h} = \mathcal{D}\mathfrak{g}$ can be isomorphic to the two-dimensional nonabelian Lie algebra from 1.2.2, or it can be abelian.

Assume first that $\mathcal{D}\mathfrak{g}$ is not abelian. We need a simple observation.

4.6.2. Lemma. *All derivations of the two-dimensional nonabelian solvable Lie algebra are inner.*

Proof. Let (e, f) be a basis of $\mathfrak{h}$ such that $[e, f] = e$. Since $\mathcal{D}\mathfrak{g}$ is a characteristic ideal spanned by e, $De = \lambda e$ with $\lambda \in k$ for any derivation $D \in \operatorname{Der}(\mathfrak{h})$. By replacing D with $D + \lambda \operatorname{ad} f$ we can assume that $De = 0$. Let $Df = ae + bf$ for some $a, b \in k$. Then we have

$$0 = De = D([e, f]) = [De, f] + [e, Df] = b[e, f] = be$$

and $b = 0$. It follows that $Df = ae$. Now, $(a \operatorname{ad} e)(f) = ae$ and $(a \operatorname{ad} e)(e) = 0$. Hence, $D = a \operatorname{ad} e$. □

Now we return to the study of the Lie algebra $\mathfrak{g}$. Let $g \in \mathfrak{g}$, $g \notin \mathfrak{h}$. Then $\operatorname{ad} g|_{\mathfrak{h}}$ is a derivation of $\mathfrak{g}$. By 4.6.2, there exists $x \in \mathfrak{h}$ such that $\operatorname{ad} g|_{\mathfrak{h}} = \operatorname{ad} x|_{\mathfrak{h}}$. Hence, by replacing g by $g - x$, we can assume that $[e, g] = [f, g] = 0$. This is impossible, since this would imply that $\mathcal{D}\mathfrak{g} = \mathcal{D}\mathfrak{h}$ is one-dimensional contrary to our assumption.

Therefore, in this case $\mathfrak{h}$ has to be abelian. Let (e, f) be a basis of $\mathfrak{h}$, and g a vector outside $\mathfrak{h}$. Then $\mathfrak{h}$ is spanned by $\operatorname{ad} g(e)$ and $\operatorname{ad} g(f)$, i.e., $A = \operatorname{ad} g|_{\mathfrak{h}}$ is a linear automorphism of $\mathfrak{h}$. If we replace g with $ag + be + cf$, the linear transformation A is replaced by aA. Therefore, the quotient of the eigenvalues of A is unchanged and independent of the choice of g. There are two options:

(1) the matrix A is semisimple;
(2) the matrix A is not semisimple.

In the first case, we can pick e and f to be the eigenvectors of A. Also, we can assume that the eigenvalue of A corresponding to e is equal to 1. We denote by α the other eigenvalue of A. Clearly $\alpha \in k^*$. In this case, we have

$$[e,f]=0,\ [g,e]=e,\ [g,f]=\alpha f.$$

Let

$$e=\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix},\ f=\begin{pmatrix}0&0&0\\0&0&1\\0&0&0\end{pmatrix},\ g=\begin{pmatrix}1&0&0\\0&\alpha&0\\0&0&0\end{pmatrix}.$$

Then these three matrices span the Lie algebra isomorphic to $\mathfrak{g}$. This proves the existence of $\mathfrak{g}$.

If we switch the order of eigenvalues the quotient $\frac{1}{\alpha}$ is replaced by α. In this case, switching e and f and replacing g by αg establishes the isomorphism of the corresponding Lie algebras. Therefore, the Lie algebras parametrized by $\alpha, \alpha' \in k^*$ are isomorphic if and only if $\alpha = \alpha'$ or $\alpha = \frac{1}{\alpha'}$. This gives an infinite family of solvable Lie algebras. They are not nilpotent, since $\mathcal{C}^p\mathfrak{g} = \mathfrak{h}$ for $p \geq 1$.

If A is not semisimple, its eigenvalues are equal, by changing g we can assume that they are equal to 1. Therefore, we can assume that A is given by the matrix

$$A=\begin{pmatrix}1&1\\0&1\end{pmatrix}$$

in the basis (e,f). Hence, we have

$$[e,f]=0,\ [g,e]=e,\ [g,f]=e+f.$$

Let

$$e=\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix},\ f=\begin{pmatrix}0&0&0\\0&0&1\\0&0&0\end{pmatrix},\ g=\begin{pmatrix}1&1&0\\0&1&0\\0&0&0\end{pmatrix}.$$

Then these three matrices span the Lie algebra isomorphic to $\mathfrak{g}$. This proves the existence of $\mathfrak{g}$. This Lie algebra is solvable, but not nilpotent.

Finally, consider the case of $\dim \mathcal{D}\mathfrak{g} = 3$. In this case, $\mathfrak{g} = \mathcal{D}\mathfrak{g}$, and $\mathfrak{g}$ cannot be solvable. By 4.6.1, $\mathfrak{g}$ is simple.

By 1.5.4, there exists $x \in \mathfrak{g}$ which is not nilpotent. Hence, it has a nonzero eigenvalue $\lambda \in k$. By multiplying it with $\frac{2}{\lambda}$, we get an element $h \in \mathfrak{g}$ such that $\operatorname{ad} h$ has eigenvalue 2. Since $\operatorname{ad} h(h) = 0$, 0 is also an eigenvalue of $\operatorname{ad} h$. Since $\mathfrak{g}$ is three-dimensional, $\operatorname{ad} h$ has at most three eigenvalues. Moreover, by 4.1.3, $\operatorname{tr} \operatorname{ad} h = 0$. Therefore, -2 is also an eigenvalue of $\operatorname{ad} h$. This in turn implies that the corresponding eigenspaces must be one-dimensional. Therefore, we can find $e, f \in \mathfrak{g}$ such that (e,f,h) is a basis of $\mathfrak{g}$ and

$$[h,e]=2e,\ [h,f]=-2f.$$

In addition, we have

$$\operatorname{ad} h([e,f]) = [\operatorname{ad} h(e), f] + [e, \operatorname{ad} h(f)] = 2[e,f] - 2[e,f] = 0$$

and $[e,f]$ is proportional to h. Clearly, $\mathcal{D}\mathfrak{g}$ is spanned by $[h,e]$, $[h,f]$ and $[e,f]$. Hence, $[e,f] \neq 0$. It follows that $[e,f] = \lambda h$ with $\lambda \neq 0$. By replacing e by $\frac{1}{\lambda}e$ we see that there exists a basis (e,f,h) such that

$$[h,e]=2e,\ [h,f]=-2f,\ [e,f]=h.$$

Therefore, there exists at most one three-dimensional simple Lie algebra over k. If $\mathfrak{g} = \mathfrak{sl}(2,k)$, and we put

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \ h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

we easily check that the above relations hold. Therefore, in the only three-dimensional simple Lie algebra is $\mathfrak{sl}(2,k)$.

4.6.3. Remark. Let $\mathfrak{g}$ be the Lie algebra $\mathfrak{sl}(2,\mathbb{C})$ with the basis (e,f,h) we described above. It has the obvious $\mathbb{R}$-structure $\mathfrak{sl}(2,\mathbb{R})$ which is spanned over $\mathbb{R}$ by (e,f,h). This real Lie algebra is the Lie algebra of the Lie group $\mathrm{SL}(2,\mathbb{R})$. On the other hand, in 1.8, we considered the Lie algebra of the group $\mathrm{SU}(2)$ which is spanned by another three linearly independent traceless 2×2 matrices. Therefore, the complexification of this Lie algebra is again $\mathfrak{g}$. In other words, the Lie algebra of $\mathrm{SU}(2)$ is another $\mathbb{R}$-structure of $\mathfrak{g}$. This shows that a complex Lie algebra can have several different $\mathbb{R}$-structures which correspond to quite different Lie groups.

4.7. Irreducible finite-dimensional representations of $\mathfrak{sl}(2,k)$. Let k be an algebraically closed field and $\mathfrak{g} = \mathfrak{sl}(2,k)$. As before, we chose the basis of $\mathfrak{g}$ given by the matrices

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \ h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Then we have, as we already remarked,

$$[h,e] = 2e, \ [h,f] = -2f, \ [e,f] = h.$$

Let V be a finite-dimensional linear space over k and $\pi : \mathfrak{g} \longrightarrow \mathcal{L}(V)$ a representation of $\mathfrak{g}$. Let $v \in V$, $v \neq 0$, be an eigenvector of $\pi(h)$ for an eigenvalue $\lambda \in k$, i.e. $\pi(h)v = \lambda v$. Then

$$\pi(h)\pi(e)v = \pi([h,e])v + \pi(e)\pi(h)v = (\lambda+2)\pi(e)v.$$

Hence, either $\pi(e)v = 0$ or $\pi(e)v$ is an eigenvector of $\pi(h)$ with the eigenvalue $\lambda + 2$. By induction, either $\pi(e)^k v$, $k \in \mathbb{Z}_+$, are nonzero eigenvectors of $\pi(h)$ with eigenvalues $\lambda + 2k$, or $\pi(e)^k v \neq 0$ and $\pi(e)^{k+1} v = 0$ for some $k \in \mathbb{Z}_+$. In the first case, since $\pi(e)^k v$ correspond to different eigenvalues of $\pi(h)$, these vectors must be linearly independent. This leads to a contradiction with $\dim V < \infty$. Hence, there exists $k \in \mathbb{Z}_+$ such that $u = \pi(e)^k v \neq 0$ is an eigenvector of $\pi(h)$ for the eigenvalue $\lambda + 2k$ and $\pi(e)u = \pi(e)^{k+1}v = 0$. Therefore, we proved the following result.

4.7.1. Lemma. *Let (π, V) be a finite-dimensional representation of $\mathfrak{g}$. Then there exists a vector $v \in V$, $v \neq 0$, such that $\pi(e)v = 0$ and $\pi(h)v = \lambda v$ for some $\lambda \in k$.*

The vector v is called the *primitive vector* of *weight* λ.

Let $v_0 \in V$ be a primitive vector of the representation π of weight λ. We put $v_n = \pi(f)^n v_0$ for $n \in \mathbb{Z}_+$. We claim that

$$\pi(h)v_n = (\lambda - 2n)v_n, \ n \in \mathbb{Z}_+.$$

This is true for $n = 0$. Assume that it holds for $m \in \mathbb{Z}_+$. Then, by the induction assumption, we have

$$\pi(h)v_{m+1} = \pi(h)\pi(f)v_m = \pi([h,f])v_m + \pi(f)\pi(h)v_m$$
$$= -2\pi(f)v_m + (\lambda - 2m)\pi(f)v_m = (\lambda - 2m - 2)\pi(f)v_m = (\lambda - 2(m+1))v_{m+1}.$$

Therefore, the assertion holds by induction in m.

We also claim that

$$\pi(e)v_n = n(\lambda - n + 1)v_{n-1}$$

for all $n \in \mathbb{Z}_+$. This is true for $n = 0$. Assume that it holds for $n = m$. Then we have

$$\pi(e)v_{m+1} = \pi(e)\pi(f)v_m = \pi([e,f])v_m + \pi(f)\pi(e)v_m$$
$$= \pi(h)v_n + m(\lambda - m + 1)\pi(f)v_{m-1} = (\lambda - 2m + m(\lambda - m + 1))v_m = (m+1)(\lambda - m)v_m,$$

and the above statement follows by induction in m.

Now $v_n \neq 0$ for all $n \in \mathbb{Z}_+$ would contradict the finitedimensionality of V, hence there exists $m \in \mathbb{Z}_+$ such that $v_m \neq 0$ and $v_{m+1} = 0$. This in turn implies that $\pi(e)v_{m+1} = (m+1)(\lambda - m)v_m = 0$. Therefore, we must have $\lambda = m$, i.e., the weight λ must be a nonnegative integer.

Therefore, we established the following addition to 4.7.1.

4.7.2. Lemma. *Let (π, V) be a finite-dimensional representation of $\mathfrak{g}$ and v a primitive vector in V. Then then the weight of v is a nonnegative integer.*

Let $\lambda = m$ be the weight of v_0. Then

$$\pi(h)v_n = (m - 2n)v_n, \ \pi(e)v_n = n(m - n + 1)v_{n-1}, \ \pi(f)v_n = v_{n+1}$$

for all $n = 0, 1, \ldots, m$. Also, $\pi(f)v_m = 0$. Therefore, the linear span of $v_0, v_1, \ldots, v_m$ is a $(m+1)$-dimensional linear subspace invariant for π.

If π is irreducible, this invariant subspace must be equal to V. This proves the exhaustion part of the following result.

4.7.3. Theorem. *Let $n \in \mathbb{Z}_+$, and V_n be the $(n+1)$-dimensional linear space with the basis $e_0, e_1, \ldots, e_n$. Define*

$$\pi_n(h)e_k = (n - 2k)e_k,$$
$$\pi_n(e)e_k = (n - k + 1)e_{k-1}$$
$$\pi_n(f)e_k = (k+1)e_{k+1}$$

for $0 \leq k \leq n$. Then (π_n, V_n) is an irreducible representation of $\mathfrak{g}$.

All irreducible finite-dimensional representations of $\mathfrak{g}$ are isomorphic to one of these representations.

Proof. It remains to check that π_n are representations. We have

$$[\pi_n(h), \pi_n(e)]e_k = \pi_n(h)\pi_n(e)e_k - \pi_n(e)\pi_n(h)e_k$$
$$= (n - k + 1)\pi_n(h)e_{k-1} - (n - 2k)\pi_n(e)e_k$$
$$= ((n - k + 1)(n - 2k + 2) - (n - 2k)(n - k + 1))e_{k-1}$$
$$= 2(n - k + 1)e_{k-1} = 2\pi_n(e)e_k$$

for all $0 \leq k \leq n$, i.e., $[\pi_n(h), \pi_n(e)] = 2\pi_n(e)$. Also, we have

$$[\pi_n(h), \pi_n(f)]e_k = \pi_n(h)\pi_n(f)e_k - \pi_n(f)\pi_n(h)e_k$$
$$= (k+1)\pi_n(h)e_{k+1} - (n - 2k)\pi_n(f)e_k$$
$$= ((k+1)(n - 2k - 2) - (n - 2k)(k+1))e_{k+1}$$
$$= -2(k+1)e_{k+1} = -2\pi_n(f)e_k$$

for all $0 \le k \le n$, i.e., $[\pi_n(h), \pi_n(f)] = -2\pi_n(f)$. Finally, we have

$$\begin{aligned}[\pi_n(e), \pi_n(f)]e_k &= \pi_n(e)\pi_n(f)e_k - \pi_n(f)\pi_n(e)e_k \\ &= (k+1)\pi_n(e)e_{k+1} - (n-k+1)\pi_n(f)e_{k-1} \\ &= ((k+1)(n-k) - (n-k+1)k)e_k \\ &= (n-2k)e_k = \pi_n(h)e_k\end{aligned}$$

for all $0 \le k \le n$, i.e., $[\pi_n(e), \pi_n(f)] = \pi_n(h)$. Therefore, π_n is a representation. □

Fix $n \in \mathbb{Z}_+$. From above formulas we see that the kernel of $\pi_n(e)$ is spanned by e_0 and the kernel of $\pi_n(f)$ is spanned by e_n. Moreover, for any $0 \le k \le \left[\frac{n}{2}\right]$, e_k is an eigenvector of $\pi_n(h)$ with eigenvalue $n - 2k > 0$. By induction in p, this implies that $\pi(f)^p e_k$ is a vector proportional to e_{k+p} and nonzero for $0 \le p \le n - 2k$. Therefore, we have

$$\pi_n(f)^{n-2k} e_k \ne 0$$

for any $0 \le k \le \left[\frac{n}{2}\right]$.

Analogously, for any $\left[\frac{n}{2}\right] \le k \le n$, e_k is an eigenvector of $\pi_n(h)$ with eigenvalue $n - 2k < 0$. By induction in p, this implies that $\pi(e)^p e_k$ is a vector proportional to e_{k-p} and nonzero for $0 \le p \le -(n-2k)$. Therefore, we have

$$\pi_n(e)^{-(n-2k)} e_k \ne 0$$

for any $\left[\frac{n}{2}\right] \le k \le n$.

This implies the following result.

4.7.4. Corollary. *Let (π, V) be a finite-dimensional representation of $\mathfrak{g}$. Then all eigenvalues of $\pi(h)$ are integral.*

Let $v \in V$ be a nonzero vector such that $\pi(h)v = pv$ for some $p \in k$. Then $p \in \mathbb{Z}$, and

(i) *if $p > 0$, $\pi(f)^p v \ne 0$;*
(ii) *if $p < 0$, $\pi(e)^{-p} v \ne 0$.*

Proof. Let $v \ne 0$ be an eigenvector of $\pi(h)$. Let

$$\{0\} = V_0 \subset V_1 \subset \cdots \subset V_{n-1} \subset V_m = V$$

be a maximal flag of subspaces in V invariant under π. Then there exists $1 \le i \le n$ such that $v \in V_i$, $v \notin V_{i-1}$. The quotient representation ρ on $W = V_i/V_{i-1}$ must be an irreducible representation of $\mathfrak{g}$. Therefore, it must be isomorphic to one of π_n, $n \in \mathbb{Z}_+$. The image w of v in W is an eigenvector of $\rho(h)$ for the eigenvalue p. By 4.7.3, p must be an integer. This proves the first statement.

Moreover, if $p = n - 2k$, w corresponds to a vector proportional to e_k. From the above discussion, we see that, if p is positive, $\rho(f)^p w \ne 0$ and, if p is negative, $\rho(e)^{-p} w \ne 0$. This immediately implies our assertion. □

5. Cartan subalgebras

5.1. Regular elements. Let $\mathfrak{g}$ be a Lie algebra over an algebraically closed field k. For $h \in \mathfrak{g}$ and $\lambda \in k$ we put

$$\mathfrak{g}(h, \lambda) = \{x \in \mathfrak{g} \mid (\operatorname{ad} h - \lambda I)^p x = 0 \text{ for some } p \in \mathbb{N}\}.$$

Then $\mathfrak{g}(h,\lambda) \neq \{0\}$ if and only if λ is an eigenvalue of $\operatorname{ad} h$. Also, since $\operatorname{ad} h(h) = 0$, we see that $\mathfrak{g}(h,0) \neq \{0\}$. Moreover, by the Jordan decomposition of $\operatorname{ad} h$ we know that

$$\mathfrak{g} = \bigoplus_{i=0}^{p} \mathfrak{g}(h,\lambda_i)$$

where $\lambda_0 = 0, \lambda_1, \dots, \lambda_p$ are distinct eigenvalues of $\operatorname{ad} h$.

For two linear subspaces $\mathfrak{a}$ and $\mathfrak{b}$ in $\mathfrak{g}$, we denote by $[\mathfrak{a},\mathfrak{b}]$ the linear span of $[x,y]$, $x \in \mathfrak{a}$, $y \in \mathfrak{b}$.

5.1.1. Lemma. *Let $h \in \mathfrak{g}$. Then*

$$[\mathfrak{g}(h,\lambda),(h,\mu)] \subset \mathfrak{g}(h,\lambda+\mu)$$

for any $\lambda,\mu \in k$.

Proof. Let $x \in \mathfrak{g}(h,\lambda)$, $y \in \mathfrak{g}(h,\mu)$. Then we have

$$\begin{aligned}(\operatorname{ad} h - (\lambda+\mu)I)[x,y] &= [\operatorname{ad} h(x),y] + [x,\operatorname{ad} h(y)] - (\lambda+\mu)[x,y] \\ &= [(\operatorname{ad} h - \lambda I)x,y] + [x,(\operatorname{ad} h - \mu I)y],\end{aligned}$$

and by induction in m, we get

$$(\operatorname{ad} h - (\lambda+\mu)I)^m[x,y] = \sum_{j=0}^{m} \binom{m}{j} [(\operatorname{ad} h - \lambda I)^j x, (\operatorname{ad} - \mu I)^{m-j} y]$$

for any $m \in \mathbb{N}$. Therefore, if $(\operatorname{ad} h - \lambda I)^p x = 0$ and $(\operatorname{ad} h - \mu I)^q y = 0$, we have $(\operatorname{ad} h - (\lambda+\mu)I)^{p+q}[x,y] = 0$. □

In particular, we have the following result.

5.1.2. Corollary. *The linear subspace $\mathfrak{g}(h,0)$ is a nonzero Lie subalgebra of $\mathfrak{g}$.*

Let

$$P_h(\lambda) = \det(\lambda I - \operatorname{ad} h)$$

be the characteristic polynomial of $\operatorname{ad} h$. Then, if $n = \dim \mathfrak{g}$, we have

$$P_h(\lambda) = \sum_{i=0}^{n} a_i(h)\lambda^i$$

where $a_1, a_2, \dots, a_n$ are polynomial functions on $\mathfrak{g}$. Since 0 is an eigenvalue of $\operatorname{ad} h$, 0 is a zero of P_h and $a_0(h) = 0$. In addition, $a_n = 1$. Let

$$\ell = \min\{i \in \mathbb{Z}_+ \mid a_i \neq 0\}.$$

The number ℓ is called the *rank* of $\mathfrak{g}$. Clearly, $0 < \ell \leq n$, i.e.,

$$0 < \operatorname{rank} \mathfrak{g} \leq \dim \mathfrak{g}.$$

Moreover, $\operatorname{rank} \mathfrak{g} = \dim \mathfrak{g}$ if and only if all $\operatorname{ad} x$, $x \in \mathfrak{g}$, are nilpotent. Therefore, by 1.5.4, we have the following result.

5.1.3. Lemma. *A Lie algebra $\mathfrak{g}$ is nilpotent if and only if* $\operatorname{rank} \mathfrak{g} = \dim \mathfrak{g}$.

An element $h \in \mathfrak{g}$ is called *regular* if $a_\ell(h) \neq 0$. Regular elements form a nonempty Zariski open set in $\mathfrak{g}$.

Let φ be an automorphism of $\mathfrak{g}$. Then we have

$$\operatorname{ad}(\varphi(h)) = \varphi \operatorname{ad}(h) \varphi^{-1}.$$

Therefore, it follows that

$$P_{\varphi(h)}(\lambda) = \det(\lambda I - \operatorname{ad}(\varphi(h))) = \det(\lambda I - \varphi \operatorname{ad} h \, \varphi^{-1}) \\ = \det(\varphi\,(\lambda I - \operatorname{ad} h)\,\varphi^{-1}) = P_h(\lambda).$$

Hence, $a_\ell(\varphi(h)) = a_\ell(h)$ for all $h \in \mathfrak{g}$. It follows that the set of all regular elements is invariant under the action of $\operatorname{Aut}(\mathfrak{g})$.

5.1.4. Lemma. *The set of regular elements in $\mathfrak{g}$ is a dense Zariski open set in $\mathfrak{g}$, stable under the action of the group $\operatorname{Aut}(\mathfrak{g})$ of automorphisms of $\mathfrak{g}$.*

Since the multiplicity of 0 as a zero of P_h is equal to $\dim \mathfrak{g}(h,0)$, we see that

$$\dim \mathfrak{g}(h,0) \geq \operatorname{rank} \mathfrak{g}$$

and the equality is attained for regular $h \in \mathfrak{g}$.

5.1.5. Example. Let $\mathfrak{g} = \mathfrak{sl}(2,k)$. Fix the standard basis e, f, h with commutation relations

$$[h,e] = 2e, \ [h,f] = -2f, \ [e,f] = h.$$

Then

$$\operatorname{ad} e = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \ \operatorname{ad} f = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ -1 & 0 & 0 \end{pmatrix}, \ \operatorname{ad} h = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Therefore, for

$$x = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$$

the characteristic polynomial of $\operatorname{ad} x$ is equal to

$$P_x(\lambda) = \begin{vmatrix} \lambda - 2a & 0 & 2b \\ 0 & \lambda + 2a & -2c \\ c & -b & \lambda \end{vmatrix} = \lambda(\lambda^2 - 4a^2) - 4cb\lambda = \lambda(\lambda^2 + 4\det x).$$

Therefore, $a_1(x) = 4\det(x)$ for all $x \in \mathfrak{g}$. It follows that $\operatorname{rank} \mathfrak{g} = 1$. Moreover, x is regular if and only if $\det(x) \neq 0$. Since $\operatorname{tr}(x) = 0$, this implies that x is regular if and only if it is not nilpotent. A regular x has two different nonzero eigenvalues μ and $-\mu$, and therefore is a semisimple matrix.

Let h_0 be a regular element in $\mathfrak{g}$. Put

$$\mathfrak{h} = \mathfrak{g}(h_0, 0).$$

5.1.6. Lemma. *The Lie algebra $\mathfrak{h}$ is nilpotent.*

Proof. Let $\lambda_0 = 0, \lambda_1, \ldots, \lambda_p$ be the distinct eigenvalues of $\operatorname{ad}(h_0)$. Put

$$\mathfrak{g}_1 = \bigoplus_{i=1}^{p} \mathfrak{g}(h_0, \lambda_i).$$

Then, by 5.1.1, we have $[\mathfrak{h}, \mathfrak{g}_1] \subset \mathfrak{g}_1$. Hence the restriction of the adjoint representation of $\mathfrak{g}$ to $\mathfrak{h}$ induces a representation ρ of $\mathfrak{h}$ on $\mathfrak{g}_1$. Consider the function

$h \longmapsto d(h) = \det \rho(h)$ on $\mathfrak{h}$. This is clearly a polynomial function on $\mathfrak{h}$. Also, if $q_i = \dim \mathfrak{g}(h_0, \lambda_i)$, we have $d(h_0) = \lambda_1^{q_1} \lambda_2^{q_2} \dots \lambda_p^{q_p} \neq 0$. Hence, $d \neq 0$. It follows that there exist a dense Zariski open set in $\mathfrak{h}$ on which d is nonzero.

Let $h \in \mathfrak{h}$ be such that $d(h) \neq 0$. The the eigenvalues of $\rho(h)$ are all nonzero. Hence $\mathfrak{g}(h, 0) \subset \mathfrak{h}$. Since h_0 is regular, the dimension of $\dim \mathfrak{h} = \operatorname{rank} \mathfrak{g}$, and $\dim \mathfrak{g}(h, 0) \geq \operatorname{rank} \mathfrak{g}$. Hence, we see that $\mathfrak{g}(h, 0) = \mathfrak{h}$. This implies that $\operatorname{ad}_{\mathfrak{h}} h$ is nilpotent. Therefore, $(\operatorname{ad}_{\mathfrak{h}} h)^q = 0$ for $q \geq \operatorname{rank} \mathfrak{g}$. Clearly, the matrix entries of $(\operatorname{ad}_{\mathfrak{h}} h)^q$ are polynomial functions on $\mathfrak{h}$. Therefore, by Zariski continuity, we must have $(\operatorname{ad}_{\mathfrak{h}} h)^q = 0$ for all $h \in \mathfrak{h}$. This implies that all $\operatorname{ad}_{\mathfrak{h}} h$, $h \in \mathfrak{h}$, are nilpotent. By 1.5.4, $\mathfrak{h}$ is a nilpotent Lie algebra. □

5.1.7. LEMMA. *The Lie algebra $\mathfrak{h}$ is equal to its normalizer.*

PROOF. Let $\mathfrak{n}$ be the normalizer of $\mathfrak{h}$ and $x \in \mathfrak{n}$. Then $[h_0, x] \in \mathfrak{h}$. Since $\mathfrak{h} = \mathfrak{g}(h_0, 0)$, we see that there exists $p \in \mathbb{Z}_+$ such that $\operatorname{ad}(h_0)^p([h_0, x]) = \operatorname{ad}(h_0)^{p+1} x = 0$. This in turn implies that $x \in \mathfrak{h}$. Therefore, $\mathfrak{n} = \mathfrak{h}$. □

5.2. Cartan subalgebras. Let $\mathfrak{g}$ be a Lie algebra. A Lie subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is a *Cartan subalgebra* of $\mathfrak{g}$ if

(i) $\mathfrak{h}$ is a nilpotent Lie algebra;
(ii) $\mathfrak{h}$ is equal to its own normalizer.

5.2.1. PROPOSITION. *Let $\mathfrak{h}$ be a Cartan subalgebra in $\mathfrak{g}$. Then $\mathfrak{h}$ is a maximal nilpotent Lie subalgebra of $\mathfrak{g}$.*

PROOF. Let $\mathfrak{n}$ be a nilpotent Lie algebra containing $\mathfrak{h}$. Assume that $\mathfrak{n} \neq \mathfrak{h}$. Then the adjoint representation of $\mathfrak{n}$ restricted to $\mathfrak{h}$ defines a representation σ of $\mathfrak{h}$ on $\mathfrak{n}/\mathfrak{h}$. By 1.5.4, this is a representation of $\mathfrak{h}$ by nilpotent linear transformations. By 1.5.3, there exists a nonzero vector $v \in \mathfrak{n}/\mathfrak{h}$ such that $\sigma(x)v = 0$ for all $x \in \mathfrak{h}$. Let $y \in \mathfrak{n}$ be a representative of the coset v. Then $[x, y] = \operatorname{ad}(x)y \in \mathfrak{h}$ for all $x \in \mathfrak{h}$. Therefore, y is in the normalizer of $\mathfrak{h}$. Since $\mathfrak{h}$ is a Cartan subalgebra, this implies that $y \in \mathfrak{h}$, i.e., $v = 0$ and we have a contradiction. Therefore, $\mathfrak{n} = \mathfrak{h}$, i.e., $\mathfrak{h}$ is a maximal nilpotent Lie subalgebra of $\mathfrak{g}$. □

5.2.2. EXAMPLE. There exist maximal nilpotent Lie subalgebras which are not Cartan subalgebras. For example, let $\mathfrak{g} = \mathfrak{sl}(2, k)$. Then the the abelian Lie subalgebra spanned by e is maximal nilpotent. To show this, assume that $\mathfrak{n}$ is a nilpotent Lie subalgebra containing e. Then $\dim \mathfrak{n}$ must be ≤ 2. Hence, it must be abelian. Let $g = \alpha e + \beta f + \gamma h$ be an element of $\mathfrak{n}$. Then

$$0 = [e, g] = \beta h - 2\gamma e.$$

Therefore $\beta = \gamma = 0$, and g is proportional to e. It follows that $\mathfrak{n}$ is spanned by e. On the other hand, the Lie subalgebra of all upper triangular matrices in $\mathfrak{g}$ normalizes $\mathfrak{n}$, so $\mathfrak{n}$ is not a Cartan subalgebra.

5.2.3. THEOREM. *Let $\mathfrak{g}$ be a Lie algebra over k. Then $\mathfrak{g}$ contains a Cartan subalgebra.*

Assume first that k is algebraically closed. Let $h \in \mathfrak{g}$ be a regular element. Then, by 5.1.6 and 5.1.7, $\mathfrak{g}(h, 0)$ is a Cartan subalgebra in $\mathfrak{g}$.

Assume now that k is not algebraically closed. Let K the algebraic closure of k.

5.2.4. Lemma. *Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{g}_K$ the Lie algebra obtained from $\mathfrak{g}$ by extension of the field of scalars.*

Let $\mathfrak{h}$ be a Lie subalgebra of $\mathfrak{g}$ and $\mathfrak{h}_K$ the Lie subalgebra of $\mathfrak{g}_K$ spanned by $\mathfrak{h}$ over K. Then the following conditions are equivalent:

(i) *$\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$;*
(ii) *$\mathfrak{h}_K$ is a Cartan subalgebra of $\mathfrak{g}_K$.*

Proof. By 2.2.3, $\mathfrak{h}$ is nilpotent if and only if $\mathfrak{h}_K$ is nilpotent.

Let $\mathfrak{n}$ be the normalizer of $\mathfrak{h}_K$. If $x \in \mathfrak{n}$ then $(\operatorname{ad} x)(\mathfrak{h}) \subset \mathfrak{h}$. Since $\mathfrak{h}$ is defined over k, it is invariant under the action of the Galois group $\operatorname{Aut}_k(K)$ on $\mathfrak{g}_K$. This implies that $(\operatorname{ad} \sigma_{\mathfrak{g}_K}(x))(\mathfrak{h}) \subset \mathfrak{h}$ for any $\sigma \in \operatorname{Aut}_k(K)$. Therefore, $\mathfrak{n}$ is stable for the action of $\operatorname{Aut}_k(K)$. By 2.1.4, $\mathfrak{n}$ is defined over k.

Assume that $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$. Then $\mathfrak{h}$ is nilpotent. By the above remark, $\mathfrak{h}_K$ is also nilpotent. Let $\mathfrak{n}$ be the normalizer of $\mathfrak{h}_K$. Then it is defined over k. Hence, it is spanned by elements in $\mathfrak{g}$ which normalize $\mathfrak{h}$. Since $\mathfrak{h}$ is equal to its normalizer in $\mathfrak{g}$, it follows that $\mathfrak{n} = \mathfrak{h}_K$. Therefore, $\mathfrak{h}_K$ is a Cartan subalgebra in $\mathfrak{g}_K$.

Assume that $\mathfrak{h}_K$ is a Cartan subalgebra in $\mathfrak{g}_K$. Then $\mathfrak{h}$ is nilpotent. Moreover, for any $x \in \mathfrak{g}$ such that $(\operatorname{ad} x)(\mathfrak{h}) \subset \mathfrak{h}$, by linearity we have $(\operatorname{ad} x)(\mathfrak{h}_K) \subset \mathfrak{h}_K$. Since $\mathfrak{h}_K$ is equal to its normalizer, this implies that $x \in \mathfrak{h}_K$ and finally $x \in \mathfrak{h}$. Therefore, the normalizer of $\mathfrak{h}$ is equal to $\mathfrak{h}$ and $\mathfrak{h}$ is a Cartan subalgebra in $\mathfrak{g}$. □

Therefore, to prove the existence of a Cartan subalgebra in $\mathfrak{g}$ it is enough to show that there exists a Cartan subalgebra of $\mathfrak{g}_K$ defined over k. Assume that there exists a regular element h of $\mathfrak{g}_K$ which is rational over k. Then h is fixed by the action of the Galois group $\operatorname{Aut}_k(K)$. This in turn implies that $\mathfrak{h} = \mathfrak{g}(h, 0)$ is stable under the action $\operatorname{Aut}_k(K)$. By 2.1.4, $\mathfrak{h}$ is defined over k.

Therefore, it is enough to show that there exists a regular element of $\mathfrak{g}_K$ rational over k. This is a consequence of the following lemma.

5.2.5. Lemma. *Let $P \in K[X_1, X_2, \dots, X_n]$ be a polynomial such that*

$$P(\lambda_1, \lambda_2, \dots, \lambda_n) = 0 \text{ for all } \lambda_1, \lambda_2, \dots, \lambda_n \in k.$$

Then $P = 0$.

Proof. We prove the statement by induction in n. If $n = 1$, the statement is obvious since k is infinite. Assume that $n > 1$. Then we have

$$P(X_1, X_2, \dots, X_n) = \sum_{s=0}^{q} P_s(X_1, X_2, \dots X_{n-1}) X_n^s$$

for some $P_j \in K[X_1, X_2, \dots, X_{n-1}]$. Fix $\lambda_1, \lambda_2, \dots, \lambda_{n-1} \in k$. Then

$$0 = P(\lambda_1, \lambda_2, \dots, \lambda_n) = \sum_{s=0}^{q} P_s(\lambda_1, \lambda_2, \dots \lambda_{n-1}) \lambda_n^s$$

for all $\lambda_n \in k$. By the first part of the proof, it follows that $P_j(\lambda_1, \lambda_2, \dots \lambda_{n-1}) = 0$ for all $0 \le j \le q$. Since $\lambda_1, \lambda_2, \dots \lambda_{n-1} \in k$ are arbitrary, by the induction assumption $P_j = 0$ for $0 \le j \le q$, and $P = 0$. □

By the preceding lemma, a_ℓ cannot vanish identically on $\mathfrak{g}$. Therefore, a regular element rational over k must exist in $\mathfrak{g}_K$. This completes the proof of 5.2.3.

We now prove a weak converse of the above results. Let $\mathfrak{g}$ be a Lie algebra over an algebraically closed field k. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$. Since $\mathfrak{h}$ is nilpotent, for any $h \in \mathfrak{h}$, $\operatorname{ad} h|_{\mathfrak{h}} = \operatorname{ad}_{\mathfrak{h}} h$ is a nilpotent linear transformation by 1.5.4. Therefore, $\mathfrak{h} \subseteq \mathfrak{g}(h, 0)$. Clearly, the adjoint action of $\mathfrak{h}$ defines a representation ρ of $\mathfrak{h}$ on $\mathfrak{g}/\mathfrak{h}$. Moreover, $\mathfrak{h} = \mathfrak{g}(h, 0)$ if and only if $\rho(h)$ is a linear automorphism of $\mathfrak{g}/\mathfrak{h}$.

5.2.6. Lemma. *Let $\mathfrak{g}$ be a Lie algebra over an algebraically closed field k. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$. Then there exists $h \in \mathfrak{g}$ such that $\mathfrak{h} = \mathfrak{g}(h, 0)$.*

Later, in 5.5.1, we are going to see that h has to be regular.

Proof. As we remarked above, we have to show that there exists $h \in \mathfrak{h}$ such that $\rho(h)$ is a linear automorphism of $\mathfrak{g}/\mathfrak{h}$.

Since $\mathfrak{h}$ is nilpotent, by 1.6.3, there is a flag

$$\{0\} = V_0 \subset V_1 \subset \cdots \subset V_m = \mathfrak{g}/\mathfrak{h}$$

of invariant subspaces for ρ such that $\dim V_i = i$ for $0 \leq i \leq m$. Moreover, there exist linear forms $\alpha_1, \alpha_2, \ldots, \alpha_m$ on $\mathfrak{h}$ such that

$$\rho(x)v - \alpha_i(x)v \in V_{i-1} \text{ for any } v \in V_i$$

for $1 \leq i \leq m$. Hence, $\alpha_1(x), \alpha_2(x), \ldots, \alpha_m(x)$ are the eigenvalues of $\rho(x)$.

We claim that none of linear forms $\alpha_1, \alpha_2, \ldots, \alpha_m$ is equal to zero. Assume the opposite. Let $1 \leq k \leq m$ be such that $\alpha_1 \neq 0, \alpha_2 \neq 0, \ldots, \alpha_{k-1} \neq 0, \alpha_k = 0$. Then $\alpha_1 \cdot \alpha_2 \cdots \alpha_{k-1}$ is a nonzero polynomial on $\mathfrak{h}$. By 5.2.5, there exists $x \in \mathfrak{h}$ such that $\alpha_1(x) \cdot \alpha_2(x) \cdots \alpha_{k-1}(x) \neq 0$. Hence, we have $\alpha_1(x) \neq 0$, $\alpha_2(x) \neq 0$, $\ldots$, $\alpha_{k-1}(x) \neq 0$. It follows that $\rho(x)|_{V_{k-1}}$ is a linear automorphism of V_{k-1}, and $\rho(x)|_{V_k}$ is not. Hence, $V_k = V_{k-1} \oplus V'$ where V' is the one-dimensional kernel of $\rho(x)|_{V_k}$.

Let $v' \in V'$, $v' \neq 0$. We claim that $\rho(y)v' = 0$ for all $y \in \mathfrak{h}$.

To show this we first claim that

$$\rho(x)^p \rho(y)v' = \rho((\operatorname{ad} x)^p y)v' \text{ for all } y \in \mathfrak{h},$$

for any $p \in \mathbb{Z}_+$. The relation is obvious for $p = 0$. For $p = 1$, we have

$$\rho(x)\rho(y)v' = \rho(x)\rho(y)v' - \rho(y)\rho(x)v' = \rho([x, y])v'.$$

Therefore, by the induction assumption, we have

$$\rho(x)^p \rho(y)v' = \rho(x)^{p-1}\rho([x, y])v' = \rho((\operatorname{ad} x)^{p-1}[x, y])v' = \rho((\operatorname{ad} x)^p y)v',$$

and the above assertion follows.

Since $\mathfrak{h}$ is nilpotent, we have $(\operatorname{ad} x)^q y = 0$ for all $y \in \mathfrak{h}$ for sufficiently large q. Therefore,

$$\rho(x)^q \rho(y)v' = 0 \text{ for all } y \in \mathfrak{h}$$

for sufficiently large q. Therefore, $\rho(y)v'$ is in the nilspace of $\rho(x)$. Since $\rho(x)|_{V_{k-1}}$ is regular, we see that $\rho(y)v' \in V'$. On the other hand, since $\alpha_k = 0$, we have $\rho(y)V_k \subseteq V_{k-1}$. This finally implies that $\rho(y)v' = 0$ for all $y \in \mathfrak{h}$.

Let $z \in \mathfrak{g}$ be a representative of the coset $v' \in \mathfrak{g}/\mathfrak{h}$. Then the above result implies that $[y, z] \in \mathfrak{h}$ for all $y \in \mathfrak{h}$. Therefore, z is in the normalizer of $\mathfrak{h}$. Since $\mathfrak{h}$ is a Cartan subalgebra, $z \in \mathfrak{h}$ and $v' = 0$. Therefore, we have a contradiction.

It follows that all $\alpha_1, \alpha_2, \ldots, \alpha_m$ are nonzero. Therefore, by 5.2.5, there exists an element $h \in \mathfrak{h}$ such that $\alpha_1(h) \neq 0, \alpha_2(h) \neq 0, \ldots, \alpha_m(h) \neq 0$, i.e., $\rho(h)$ is regular. □

5.2.7. Corollary. *Let $\mathfrak{g}$ be a Lie algebra. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$. Then* $\dim \mathfrak{h} \geq \operatorname{rank} \mathfrak{g}$.

Proof. Assume first that $\mathfrak{g}$ is a Lie algebra over an algebraically closed field. Then by 5.2.6, $\mathfrak{g}(h, 0)$ for some $h \in \mathfrak{h}$. Therefore, $\dim \mathfrak{h} \geq \operatorname{rank} \mathfrak{g}$.

The general case follows from 5.2.4. □

Later, in 5.5.3, we are going to see that the inequality in the above result is actually an equality.

5.3. Cartan subalgebras in semisimple Lie algebras. In this section we specialize the discussion to semisimple Lie algebras.

5.3.1. Lemma. *Let $\mathfrak{g}$ be a semisimple Lie algebra and $\mathfrak{h}$ a Cartan subalgebra of $\mathfrak{g}$. Then $\mathfrak{h}$ is abelian.*

Proof. Assume first that k is algebraically closed. By 5.2.6, $h_0 \in \mathfrak{h}$ be such that $\mathfrak{h} = \mathfrak{g}(h_0, 0)$. Let $\lambda \neq 0$ and $x \in \mathfrak{g}(h_0, \lambda)$. Then, for $h \in \mathfrak{h}$, we have

$$\operatorname{ad}(x) \operatorname{ad}(h)(\mathfrak{g}(h_0, \mu)) \subset \operatorname{ad}(x)(\mathfrak{g}(h_0, \mu)) \subset \mathfrak{g}(h_0, \mu + \lambda).$$

If we choose a basis of $\mathfrak{g}$ corresponding to the decomposition $\mathfrak{g} = \bigoplus_{i=0}^{p} \mathfrak{g}(h_0, \lambda_i)$, where $\lambda_0 = 0, \lambda_1, \dots, \lambda_p$ are distinct eigenvalues of $\operatorname{ad}(h_0)$, we see that the corresponding block matrix of $\operatorname{ad}(x)\operatorname{ad}(h)$ has zero blocks on the diagonal. Therefore, $B(x, h) = \operatorname{tr}(\operatorname{ad}(x)\operatorname{ad}(h)) = 0$. Hence, it follows that $\mathfrak{h}$ is orthogonal to $\mathfrak{g}(h_0, \lambda_i)$ for any $1 \leq i \leq p$.

Since $\mathfrak{h}$ is nilpotent, it is also solvable. By 3.2.1, it follows that $\mathfrak{h}$ is orthogonal to $\mathcal{D}\mathfrak{h}$. This implies that $\mathcal{D}\mathfrak{h}$ is orthogonal to $\mathfrak{g}$. Since the Killing form is nondegenerate on $\mathfrak{g}$ by 4.1.1, it follows that $\mathcal{D}\mathfrak{h} = \{0\}$, i.e., $\mathfrak{h}$ is abelian.

The general case follows from 5.2.4. □

Since Cartan subalgebras are maximal nilpotent by 5.2.1, this implies the following result.

5.3.2. Corollary. *Cartan subalgebras in a semisimple Lie algebra are maximal abelian Lie subalgebras.*

5.3.3. Lemma. *Let $\mathfrak{g}$ be a semisimple Lie algebra over an algebraically closed field k. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$. Then all $h \in \mathfrak{h}$ are semisimple.*

Proof. By 5.3.2, $\mathfrak{h}$ is an abelian Lie subalgebra. Let $h \in \mathfrak{h}$. Let $h = s + n$ be its Jordan decomposition. Clearly, $\operatorname{ad} h|_{\mathfrak{h}} = 0$. Since $\operatorname{ad} s$ and $\operatorname{ad} n$ are the semisimple and nilpotent part of $\operatorname{ad} h$, they are polynomials without constant term in $\operatorname{ad} h$ by 3.1.4. Therefore $(\operatorname{ad} s)(\mathfrak{h}) = (\operatorname{ad} n)(\mathfrak{h}) = \{0\}$. Since $\mathfrak{h}$ is maximal abelian, we conclude that $s, n \in \mathfrak{h}$.

By 5.2.6, $\mathfrak{h} = \mathfrak{g}(h_0, 0)$ for some element $h_0 \in \mathfrak{h}$. As in the proof of 5.3.1, we see that $\mathfrak{h}$ is orthogonal to $\mathfrak{g}(h_0, \lambda)$ for eigenvalues $\lambda \neq 0$.

Let $y \in \mathfrak{h}$. Then y and n commute. Hence $\operatorname{ad} y$ and $\operatorname{ad} n$ commute and $\operatorname{ad}(y)\operatorname{ad}(n)$ is a nilpotent linear transformation. This in turn implies that $B(y, n) = 0$. Therefore, n is orthogonal to $\mathfrak{h}$. This implies that n is orthogonal to $\mathfrak{g}$. Since the Killing form is nondegenerate, $n = 0$. Therefore $h = s$ is semisimple. □

By 5.2.6, this has the following immediate consequence.

5.3.4. Corollary. *Let $\mathfrak{g}$ be a semisimple Lie algebra over an algebraically closed field k. Then all regular elements in $\mathfrak{g}$ are semisimple.*

PROOF. Let h be a regular element in $\mathfrak{g}$. By 5.1.6 and 5.1.7, $\mathfrak{g}(h,0)$ is a Cartan subalgebra of $\mathfrak{g}$. By 5.3.3, h must be semisimple. □

5.4. Elementary automorphisms. Let V be a linear space over the field k. Let $T \in \mathcal{L}(V)$ be a nilpotent linear transformation. Then

$$e^T = \sum_{p=0}^{\infty} \frac{1}{p!} T^p$$

is a well defined linear transformation on V.

5.4.1. LEMMA. *Let T be a nilpotent linear transformation on V. Then the map $\lambda \longmapsto e^{\lambda T}$ is a homomorphism of the additive group k into* $\mathrm{GL}(V)$.

PROOF. First, if $\lambda, \mu \in k$, we have

$$\begin{aligned} e^{(\lambda+\mu)T} &= \sum_{p=0}^{\infty} \frac{1}{p!}(\lambda+\mu)^p T^p = \sum_{p=0}^{\infty}\sum_{j=0}^{p} \frac{1}{p!}\binom{p}{j}\lambda^{p-j}\mu^j T^p \\ &= \sum_{j=0}^{\infty}\sum_{p=j}^{p} \frac{(p-j)!}{j!}\lambda^{p-j}\mu^j T^p = \sum_{j=0}^{\infty}\sum_{p=j}^{p} \frac{1}{(p-j)!j!}\lambda^{p-j}\mu^j T^p \\ &= \sum_{j=0}^{\infty}\sum_{p=0}^{\infty} \frac{1}{p!j!}\lambda^p\mu^j T^{p+j} = \left(\sum_{p=0}^{\infty}\frac{1}{p!}\lambda^p T^p\right)\left(\sum_{p=0}^{\infty}\frac{1}{p!}\mu^p T^p\right) = e^{\lambda T}e^{\mu T}. \end{aligned}$$

Therefore, the inverse of e^T is e^{-T} and $e^T \in \mathrm{GL}(V)$. Moreover, $\lambda \longmapsto e^T$ is a group homomorphism of the additive group k into $\mathrm{GL}(V)$. □

5.4.2. LEMMA. *Let $\mathfrak{g}$ be a Lie algebra. Let D be a nilpotent derivation of $\mathfrak{g}$. Then e^D is an automorphism of $\mathfrak{g}$.*

PROOF. Clearly, by 5.4.1, e^D is an automorphism of the linear space $\mathfrak{g}$. On the other hand, by induction one can easily establish that

$$D^p([x,y]) = \sum_{j=0}^{p}\binom{p}{j}[D^{p-j}x, D^j y].$$

Hence, we have

$$\begin{aligned} e^D([x,y]) &= \sum_{p=0}^{\infty}\frac{1}{p!}D^p[x,y] = \sum_{p=0}^{\infty}\sum_{j=0}^{p}\frac{1}{(p-j)!j!}[D^{p-j}x, D^j y] \\ &= \sum_{j=0}^{\infty}\sum_{p=j}^{\infty}\frac{1}{(p-j)!j!}[D^{p-j}x, D^j y] = \sum_{j=0}^{\infty}\sum_{p=0}^{\infty}\frac{1}{p!j!}[D^p x, D^j y] = [e^D x, e^D y] \end{aligned}$$

for any $x, y \in \mathfrak{g}$. Therefore e^D is an automorphism of $\mathfrak{g}$.

□

Let $\mathrm{Aut}(\mathfrak{g})$ be the group of all automorphisms of $\mathfrak{g}$. Let $x \in \mathfrak{g}$ be such that $\operatorname{ad} x$ is nilpotent. Then $e^{\operatorname{ad} x}$ is an automorphism of $\mathfrak{g}$. Denote by $\mathrm{Aut}_e(\mathfrak{g})$ the subgroup of $\mathrm{Aut}(\mathfrak{g})$ generated by the automorphisms of this form. The elements of $Aut_e(G)$ are called *elementary automorphisms*.

5.4.3. LEMMA. *The subgroup* $\mathrm{Aut}_e(\mathfrak{g})$ *is normal in* $\mathrm{Aut}(\mathfrak{g})$.

Proof. Let φ be an automorphism of $\mathfrak{g}$. Let $x \in \mathfrak{g}$ be such that $\operatorname{ad} x$ is nilpotent. Then $\operatorname{ad}(\varphi(x)) = \varphi \operatorname{ad} x \, \varphi^{-1}$ is also nilpotent. Therefore,

$$e^{\operatorname{ad}(\varphi(x))} = \sum_{p=0}^{\infty} \frac{1}{p!} \operatorname{ad}(\varphi(x))^p = \sum_{p=0}^{\infty} \frac{1}{p!} \varphi \, (\operatorname{ad} x)^p \, \varphi^{-1} = \varphi \, e^{\operatorname{ad} x} \varphi^{-1}$$

is an elementary automorphism of $\mathfrak{g}$. Hence $\varphi \operatorname{Aut}_e(\mathfrak{g}) \varphi^{-1} \subset \operatorname{Aut}_e(\mathfrak{g})$, i.e., $\operatorname{Aut}_e(\mathfrak{g})$ is a normal subgroup of $\operatorname{Aut}(\mathfrak{g})$. □

5.5. Conjugacy theorem. Let $\mathfrak{g}$ be a Lie algebra over an algebraically closed field k. Let $\mathfrak{h}$ be a Cartan subalgebra in $\mathfrak{g}$. Then by 5.2.6, there exists an element $h_0 \in \mathfrak{h}$ such that $\mathfrak{h} = \mathfrak{g}(h_0, 0)$. First we want to prove a stronger form of this result.

5.5.1. Lemma. *Let $\mathfrak{g}$ be a Lie algebra over an algebraically closed field k. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$. Then there exists a regular element $h \in \mathfrak{g}$ such that $\mathfrak{h} = \mathfrak{g}(h, 0)$.*

To prove this, it is enough to show that h_0 is regular. Consider the decomposition

$$\mathfrak{g} = \bigoplus_{i=0}^{p} \mathfrak{g}(h_0, \lambda_i)$$

where $\lambda_0 = 0, \lambda_1, \dots, \lambda_p$ are mutually different eigenvalues of $\operatorname{ad} h_0$. By 5.1.1, $\operatorname{ad} x_i$ is nilpotent for any $x_i \in \mathfrak{g}(h_0, \lambda_i)$, $1 \le i \le p$. Therefore, $e^{\operatorname{ad} x_i}$ are elementary automorphisms of $\mathfrak{g}$. It follows that we can define a map

$$F : \mathfrak{h} \times \mathfrak{g}(h_0, \lambda_1) \times \mathfrak{g}(h_0, \lambda_2) \times \cdots \times \mathfrak{g}(h_0, \lambda_p) \longrightarrow \mathfrak{g}$$

by

$$F(h, x_1, x_2, \dots, x_p) = e^{\operatorname{ad} x_1} e^{\operatorname{ad} x_2} \dots e^{adx_p} h$$

for $x_i \in \mathfrak{g}(h_0, \lambda_i)$, $1 \le i \le p$, $h \in \mathfrak{h}$.

This is clearly a polynomial map from $\mathfrak{h} \times \mathfrak{g}(h_0, \lambda_1) \times \mathfrak{g}(h_0, \lambda_2) \times \cdots \times \mathfrak{g}(h_0, \lambda_p)$ into $\mathfrak{g}$.

Let $T_{(h_0,0,0,\dots,0)}(F)$ be the differential of this map at $(h_0, 0, 0, \dots, 0)$.

5.5.2. Lemma. *The linear map $T_{(h_0,0,0,\dots,0)}(F) : \mathfrak{h} \times \mathfrak{g}(h_0, \lambda_1) \times \mathfrak{g}(h_0, \lambda_2) \times \cdots \times \mathfrak{g}(h_0, \lambda_p) \longrightarrow \mathfrak{g}$ is surjective.*

Proof. We have

$$F(h, 0, \dots, 0) = h \text{ and } F(h_0, 0, \dots, 0, x_i, 0, \dots, 0) = e^{\operatorname{ad} x_i} h_0$$

for any $1 \le i \le p$. Hence, we have

$$T_{(h_0,0,0,\dots,0)}(F)(h, 0, \dots, 0) = h$$

for any $h \in \mathfrak{h}$. Therefore, the differential of $T_{(h_0,0,0,\dots,0)}(F)$ is an isomorphism of $\mathfrak{h} \times \{0\} \times \cdots \times \{0\}$ onto $\mathfrak{h} \subset \mathfrak{g}$. Moreover, for $1 \le i \le p$, we have

$$F(h_0, 0, \dots, 0, x_i, 0, \dots, 0) = e^{\operatorname{ad} x_i}(h_0)$$

for any $x_i \in \mathfrak{g}(h_0, \lambda_i)$. Therefore, we have

$$T_{(h_0,0,0,\dots,0)}(F)(0, \dots, 0, x_i, 0, \dots, 0) = \operatorname{ad} x_i(h_0) = -\operatorname{ad}(h_0) x_i$$

for any $1 \le i \le p$. It follows that the differential of $T_{(h_0,0,0,\dots,0)}(F)$ is an isomorphism of $\{0\} \times \cdots \times \{0\} \times \mathfrak{g}(h_0, \lambda_i) \times \{0\} \times \cdots \times \{0\}$ onto $\mathfrak{g}(h_0, \lambda_i) \subset \mathfrak{g}$ for any $1 \le i \le p$. This clearly implies that the differential $T_{(h_0,0,0,\dots,0)}(F)$ is surjective. □

At this point we need a polynomial analogue of 1.1.3.5 which is proved in 5.6.2 in the next section. By this result, F is a dominant morphism. Hence, the image of F is dense in $\mathfrak{g}$. In particular, the set $\mathrm{Aut}_e(\mathfrak{g})\cdot\mathfrak{h}$ is dense in $\mathfrak{g}$. By 5.1.4, the set of all regular elements is also a dense Zariski open set in $\mathfrak{g}$. Therefore, these two sets have nonempty intersection. This implies that there is a $h \in \mathfrak{h}$ and $\varphi \in \mathrm{Aut}_e(\mathfrak{g})$ such that $\varphi(h)$ is regular. Since the set of all regular elements is invariant under $\mathrm{Aut}(\mathfrak{g})$, it follows that h is also regular. Therefore, $\mathfrak{g}(h,0)$ is a Cartan subalgebra of $\mathfrak{g}$ by 5.1.6 and 5.1.7. On the other hand, since $\mathfrak{h}$ is nilpotent, $\mathrm{ad}_{\mathfrak{h}} h$ is a nilpotent linear transformation. Hence, $\mathfrak{h} \subset \mathfrak{g}(h,0)$. Since $\mathfrak{h}$ is a maximal nilpotent Lie subalgebra by 5.2.1, it follows that $\mathfrak{h} = \mathfrak{g}(h,0)$. Therefore, $\dim\mathfrak{h} = \operatorname{rank}\mathfrak{g}$. This in turn implies that h_0 is regular. This completes the proof of 5.5.1.

In addition we see that the following result holds.

5.5.3. Proposition. *The dimension of all Cartan subalgebras in $\mathfrak{g}$ is equal to* $\operatorname{rank}\mathfrak{g}$.

Proof. This statement follows from 5.5.1 for Lie algebras over algebraically closed fields.

In general case, it follows from 5.2.4. □

Finally, we have the following conjugacy result.

5.5.4. Theorem. *Let $\mathfrak{g}$ be a Lie algebra over an algebraically closed field k. The the group $\mathrm{Aut}_e(\mathfrak{g})$ acts transitively on the set of all Cartan subalgebras of $\mathfrak{g}$.*

Proof. Let $\mathfrak{h}$ be a Cartan subalgebra in $\mathfrak{g}$. Then, by 5.5.1, $\mathfrak{h} = \mathfrak{g}(h_0,0)$ for some regular element $h_0 \in \mathfrak{h}$. Let

$$F : \mathfrak{h} \times \mathfrak{g}(h_0,\lambda_1) \times \mathfrak{g}(h_0,\lambda_2) \times \cdots \times \mathfrak{g}(h_0,\lambda_p) \longrightarrow \mathfrak{g}$$

be the map given by

$$F(h,x_1,x_2,\dots,x_p) = e^{\mathrm{ad}\,x_1} e^{\mathrm{ad}\,x_2} \dots e^{adx_p} h.$$

Then, as we remarked in the proof of 5.5.1, the polynomial map F is dominant. By **??**, the image of F contains a dense Zariski open set in $\mathfrak{g}$. Therefore, the set $\mathrm{Aut}_e(\mathfrak{g})\cdot\mathfrak{h}$ contains a dense Zariski open set in $\mathfrak{g}$.

Let $\mathfrak{h}$ and $\mathfrak{h}'$ be two Cartan subalgebras of $\mathfrak{g}$. Then, the sets $\mathrm{Aut}_e(\mathfrak{g})\cdot\mathfrak{h}$ and $\mathrm{Aut}_e(\mathfrak{g})\cdot\mathfrak{h}'$ contain dense Zariski open sets in $\mathfrak{g}$. Therefore, they contain a common regular element $\mathfrak{h}$. This implies that there exists a regular element $h \in \mathfrak{h}$, a regular element $h' \in \mathfrak{h}'$ and an elementary automorphism φ such that $\varphi(h) = h'$. As we remarked in the proof of 5.5.1, $\mathfrak{h} = \mathfrak{g}(h,0)$ and $\mathfrak{h}' = \mathfrak{g}(h',0)$. This in turn implies that

$$\varphi(\mathfrak{h}) = \varphi(\mathfrak{g}(h,0)) = \mathfrak{g}(\varphi(h),0) = \mathfrak{g}(h',0) = \mathfrak{h}'.$$

□

5.6. Dominant polynomial maps. Let V and W be two linear spaces over an algebraically closed field k. we denote by $R(V)$ and $R(W)$ the rings of polynomials with coefficients in k on V, resp. W.

A map $F : V \longrightarrow W$ is a *polynomial map* if $P \circ F \in R(V)$ for any $P \in R(W)$. If $F : V \longrightarrow W$ is a polynomial map, it induces a k-algebra homomorphism $F^* : R(W) \longrightarrow R(V)$ given by $F^*(P) = P \circ F$ for any $P \in R(W)$.

We say that a polynomial map $F : V \longrightarrow W$ is *dominant* if $F^* : R(W) \longrightarrow R(V)$ is injective.

5.6.1. LEMMA. *A polynomial map $F : V \longrightarrow W$ is dominant if and only if the image of F is Zariski dense in W.*

PROOF. Let U be a nonempty open set in W. Then there exists a polynomial P on W such that $W_P = \{w \in W \mid P(w) \neq 0\} \subset U$. If F is dominant, $P \circ F \neq 0$. Therefore, $W_P \cap \operatorname{im} F \neq \emptyset$. Therefore, $U \cap \operatorname{im} F$ is nonempty, i.e., $\operatorname{im} F$ is dense in W.

If $\operatorname{im} F$ is dense in W, for any nonzero polynomial P on W we have $W_P \cap \operatorname{im} F \neq \emptyset$. Therefore, there exists $v \in V$ such that $P(F(v)) \neq 0$, i.e., $F^*(P) \neq 0$. It follows that F^* is injective. □

Let $v \in V$. We can consider the polynomial map $G : h \longmapsto F(v+h) - F(v)$. Clearly, $G(0) = 0$, so the constant term of G is equal to 0. Let $\mathfrak{m}$ be the maximal ideal in $R(V)$ consisting of all polynomials vanishing at 0. Then, there exists a unique linear map $T_v(F) : V \longrightarrow W$ such that

$$F(v+h) - F(v) - T_v(F)(h) \in \mathfrak{m}^2 \otimes W.$$

This linear map is the *differential* of the polynomial map F.

We need the following version of 1.1.3.5.

5.6.2. PROPOSITION. *Let $F : V \longrightarrow W$ be a polynomial map and $v \in V$. Assume that $T_v(F) : V \longrightarrow W$ is a surjective linear map for some $v \in V$. Then F is dominant.*

PROOF. By an affine change of coordinates, we can assume that $v = 0$ and $F(v) = 0$. Therefore, we have $F(h) - T_0(F)(h) \in \mathfrak{m}^2 \otimes W$.

Let P be a nonzero polynomial on W. Then we can write $P = \sum_{q=0}^{\infty} P_q$, where P_q are homogeneous polynomials of degree q. Assume that $P_q = 0$ for $q < q_0$ and $P_{q_0} \neq 0$. Then $F^*(P)$ is a polynomial on V and $F^*(P) = \sum_{q=0}^{\infty} Q_q$, where Q_q are homogeneous polynomials of degree q. Clearly, $Q_q = 0$ for $q < q_0$ and $Q_{q_0} = P_{q_0} \circ T_0(F)$. Since $T_0(F)$ is a surjective linear map and $P_{q_0} \neq 0$, we have $Q_{q_0} \neq 0$. This in turn implies that $F^*(P) \neq 0$, and F^* is injective. □

Finally, we need the following basic result about dominant polynomial maps.

5.6.3. THEOREM. *Let $F : V \longrightarrow W$ be a dominant polynomial map. Then the image of F contains a nonempty Zariski open set in W.*

The proof of this result is based on some basic results from commutative algebra.

Since k is algebraically closed, by the Hilbert Nullstellensatz, the points in V and W are in bijection with the maximal ideals in rings of regular functions $R(V)$ and $R(W)$, respectively. Let $w \in W$, and N_w the maximal ideal of all polynomials on W vanishing in w. Then $F^*(N_w) \subset R(V)$. If the ideal generated by $F^*(N_w)$ is different from $R(V)$, there exists a maximal ideal M_v in $R(V)$ corresponding to $v \in V$, such that $F^*(N_w) \subset M_v$. Hence, for any $Q \in N_w$, we have $F^*(Q) = Q \circ F \in M_v$, i.e., $Q(F(v)) = 0$. This in turn implies that $F(v) = w$. Conversely, if $F(v) = w$ for some $v \in V$ and $w \in W$, we have $F^*(N_w) \subset M_v$, and the ideal generated by $F^*(N_w)$ is different from $R(V)$. Therefore, the image of F is characterized as the set of all $w \in W$ such that the ideal generated by $F^*(N_w)$ is different from $R(V)$.

Therefore, it is enough to find a nonzero polynomial $Q \in R(W)$ such that $Q(w) \neq 0$ implies that $F^*(N_w)$ doesn't generate $R(V)$. In other words, for any maximal ideal N in $R(W)$ such that $Q \notin N$, the ideal generated by $F^*(N)$ doesn't contain 1.

We are going to prove a slightly stronger statement: Let $A = R(V)$ and B a subalgebra of $R(V)$. Then, for any nonzero polynomial $P \in A$, there exists a nonzero polynomial $R \in B$ such that for any maximal ideal N in B not contianing R, the ideal generated by N in A doesn't contain P. The above statement follows immediately if we put $B = F^*(R(W))$, $R = Q \circ F$ and $P = 1$.

Clearly, A is a finitely generated algebra over B. Assume that $x_1, x_2, \dots, x_p$ are generators of A over B. Let B_k be the algebra generated over B by x_i, $1 \le i \le k$. Then $B_0 = B$ and $B_p = A$. We are going to show that there exists a family of nonzero polynomials $R_p = P, R_{p-1}, \dots, R_1, R_0$ such that any maximal ideal in B_{i-1} not containing R_{i-1} generates an ideal in B_i which doesn't contain R_i for any $1 \le i \le p$. Then $R = R_0$ satisfies the above statement.

To prove this statement, put $C = B_{i-1}$ and $y = x_i$. Then $D = B_i$ is the algebra generated by y over C. Consider the natural algebra homomorphism of $C[Y]$ into D which maps Y into y.

From the above discussion, to prove 5.6.3 it is enough to prove the following result.

5.6.4. Lemma. *Let S be a nonzero polynomial in D. Then there exist a nonzero polynomial $T \in C$ such that for any maximal ideal M in C which doesn't contain T, the ideal in D generated by M doesn't contain S.*

Proof. There are two possibilities: either the homomorphism of $C[Y]$ into D is an isomorphism, or it has a nontrivial kernel.

Assume first that this homomorphism is an isomorphism. Then $S \in D$ is a nonzero polynomial in y. Then $S = \sum_j a_j y^j$ where $a_j \in C$. Let j_0 be such that $T = a_{j_0} \neq 0$. If N is a maximal ideal in C, it generates an ideal in D which consists of all polynomials in y with coefficients in N. Therefore, if N doesn't contain T, S is not in this ideal.

Assume now that the homomorphism from $C[Y]$ into D has nonzero kernel. Let U be a nonzero polynomial in $C[Y]$ which is in the kernel of the natural homomorphism of $C[Y]$ into D. We can assume that the degree of U is minimal possible. Let $U = \sum_{i=0}^{n} b_j Y$ and $b_n \neq 0$. □

CHAPTER 5

Structure of semisimple Lie algebras

1. Root systems

1.1. Reflections. Let V be a finite dimensional linear space over a field k of characteristic 0. Let $\alpha \in V$. A linear automorphism $s \in \mathcal{L}(V)$ is a *reflection* with respect to α if:

(i) $s(\alpha) = -\alpha$;
(ii) $H = \{h \in V \mid s(h) = h\}$ satisfies $\dim H = \dim V - 1$.

Clearly, $s^2 = I$ and s is completely determined by α and H. The linear subspace H is called the *reflection hyperplane* of s.

Let V^* be the linear dual of V. As we remarked in 4.4.5, we have a linear isomorphism $\varphi : V^* \otimes V \longrightarrow \mathcal{L}(V)$ defined by

$$\varphi(f \otimes w)(v) = f(v)w \text{ for } f \in V^*, v, w \in V.$$

Consider $\alpha, \beta \in V$ and $f, g \in V^*$. Then we have

$$\begin{aligned}&(I + \varphi(f \otimes \alpha))(I + \varphi(g \otimes \beta))(v) = (I + \varphi(f \otimes \alpha))(v + g(v)\beta)\\ &= v + f(v)\alpha + g(v)\beta + f(\beta)g(v)\alpha = (I + \varphi(f \otimes \alpha) + \varphi(g \otimes \beta) + f(\beta)\varphi(g \otimes \alpha))(v)\end{aligned}$$

for any $v \in V$, i.e.,

$$(I + \varphi(f \otimes \alpha))(I + \varphi(g \otimes \beta)) = I + \varphi(f \otimes \alpha) + \varphi(g \otimes \beta) + f(\beta)\varphi(g \otimes \alpha).$$

1.1.1. Lemma. *Let $s \in \mathcal{L}(V)$. Then the following assertions are equivalent:*

(i) *s is a reflection with respect to α;*
(ii) *$s = I - \varphi(\alpha^* \otimes \alpha)$ for some $\alpha^* \in V^*$ with $\alpha^*(\alpha) = 2$;*
(iii) *$s^2 = I$ and $\operatorname{im}(I - s) = k\,\alpha$.*

If these conditions are satisfied, α^* is uniquely determined by s.

Proof. (i)⇒(ii) Let s be a reflection with respect to α, and H its reflection hyperplane. Then there exists a unique $\alpha^* \in V^*$ such that $H = \ker \alpha^*$ and $\alpha^*(\alpha) = 2$. In addition, we have

$$(I - \varphi(\alpha^* \otimes \alpha))(\alpha) = \alpha - \alpha^*(\alpha)\alpha = \alpha - 2\alpha = -\alpha.$$

and

$$(I - \varphi(\alpha^* \otimes \alpha))(h) = h - \alpha^*(h)\alpha = h$$

for any $h \in H$. Therefore, we have $s = I - \varphi(\alpha^* \otimes \alpha)$.

(ii)⇒(iii) We have $I - s = \varphi(\alpha^* \otimes \alpha)$. Therefore, $\operatorname{im}(I - s) = k\,\alpha$ since $\alpha^* \neq 0$. In addition,

$$s^2 = I - 2\varphi(\alpha^* \otimes \alpha) + \alpha^*(\alpha)\varphi(\alpha^* \otimes \alpha) = I.$$

(iii)⇒(i) For any $v \in V$, we have $(I - s)(v) = f(v)\alpha$ for some nonzero $f \in V^*$. Therefore, we have $s(v) = (I - \varphi(f \otimes \alpha))(v)$. It follows that

$$I = s^2 = (I - \varphi(f \otimes \alpha))^2 = I - 2\varphi(f \otimes \alpha) + f(\alpha)\varphi(f \otimes \alpha) = I + (f(\alpha) - 2)\varphi(f \otimes \alpha).$$

Since f is nonzero, $\varphi(f \otimes \alpha) \neq 0$, and it follows that $f(\alpha) = 2$. Let $H = \ker f$. Then $\dim H = \dim V - 1$ and $s(h) = h$ for any $h \in H$. On the other hand, we have

$$s(\alpha) = (I - \varphi(f \otimes \alpha))(\alpha) = \alpha - f(\alpha)\alpha = -\alpha$$

and s is a reflection with respect to α. □

1.1.2. Lemma. *Let α be a nonzero vector in V. Let R be a finite set of vectors in V which spans V. Then there exists at most one reflection s with respect to α such that $s(R) \subseteq R$.*

Proof. Let s and s' be two reflections satisfying the conditions of the lemma. Let $t = ss'$. Then t is a linear automorphism of V which maps R into itself. Since R is finite, $t : R \longrightarrow R$ is a bijection. Hence, t induces a permutation of R. Again, since R is finite, $t^n : R \longrightarrow R$ is the identity map for sufficiently large $n \in \mathbb{Z}_+$. Since R spans V, this implies that $t^n = I$.

Assume that

$$s = I - \varphi(f \otimes \alpha) \text{ and } s' = I - \varphi(f' \otimes \alpha)$$

with $f(\alpha) = f'(\alpha) = 2$. Then we have

$$t = ss' = I - \varphi(f \otimes \alpha) - \varphi(f' \otimes \alpha) + f(\alpha)\varphi(f' \otimes \alpha) = I - \varphi((f - f') \otimes \alpha).$$

If we put $g = f' - f$, we see that $t = I + \varphi(g \otimes \alpha)$. Moreover, we have $g(\alpha) = f'(\alpha) - f(\alpha) = 0$.

We claim that $t^p = I + p\varphi(g \otimes \alpha)$ for any $p \in \mathbb{N}$. Clearly, this is true for $p = 1$. Assume that it holds for $p = m$. Then, by the induction assumption, we have

$$\begin{aligned} t^{m+1} &= (I + m\varphi(g \otimes \alpha))(I + \varphi(g \otimes \alpha)) \\ &= I + m\varphi(g \otimes \alpha) + \varphi(g \otimes \alpha) + mg(\alpha)\varphi(g \otimes \alpha) = I + (m + 1)\varphi(g \otimes \alpha). \end{aligned}$$

This proves the claim.

It follows that

$$I = t^n = I + n\varphi(g \otimes \alpha)$$

for sufficiently large n. This in turn implies that $\varphi(g \otimes \alpha) = 0$ and $g = 0$. Therefore, $t = I$, and $s = s'$. □

1.2. Root systems. Let V be a finite dimensional linear space over a field k of characteristic 0. A finite subset R of V is a *root system* in V if:

(i) 0 is not in R;
(ii) R spans V;
(iii) for any $\alpha \in R$ there exists a reflection s_α with respect to α such that $s_\alpha(R) = R$;
(iv) for arbitrary $\alpha, \beta \in R$ we have

$$s_\alpha(\beta) = \beta + n\alpha$$

where $n \in \mathbb{Z}$.

The elements of R are called *roots* of V with respect to R.

Clearly, by 1.1.2, the reflection s_α is unique. We call it *the reflection with respect to root* α.

The dimension of V is called the *rank* of R and denoted by $\operatorname{rank} R$.

For any $\alpha \in R$ we have

$$s_\alpha = I - \varphi(\alpha^* \otimes \alpha)$$

for a unique $\alpha^* \in V^*$. The vector α^* is called the *dual root* of α.

The property (iv) is equivalent with

(iv)' for any $\alpha, \beta \in R$ we have $\alpha^*(\beta) \in \mathbb{Z}$.

We define $n(\beta, \alpha) = \alpha^*(\beta)$.

Clearly, $\alpha \in R$ implies that $-\alpha = s_\alpha(\alpha) \in R$.

1.2.1. Lemma. *Let R be a root system in V and α and β two proportional roots. Then $\beta = t\alpha$ where $t \in \{\pm\frac{1}{2}, \pm 1, \pm 2\}$.*

Proof. Let β be a root proportional to α. Then $\beta = t\alpha$ for some $t \in k^*$. Moreover, $\alpha^*(\beta) = t\alpha^*(\alpha) = 2t \in \mathbb{Z}$. Therefore, $t \in \frac{1}{2}\mathbb{Z}$. By replacing β with $-\beta$ we can assume that $t > 0$.

Let $\gamma = s\alpha$, $s \in \mathbb{Q}$, be a root such that s is maximal possible. By the above discussion, wee have $s \in \frac{1}{2}\mathbb{N}$ and $s \geq 1$. Then, $\gamma^*(\alpha) = \frac{1}{s}\gamma^*(\gamma) = \frac{2}{s}$ is an integer. Therefore, s is either 1 or 2. It follows that $\gamma = \alpha$ or $\gamma = 2\alpha$.

In the first case, $t \leq 1$ and $t = \{\frac{1}{2}, 1\}$. In the second, we can replace α with γ and conclude that $\beta = \frac{1}{2}\gamma = \alpha$ or $\beta = \gamma = 2\alpha$. ☐

Hence, for any root α, the set of all roots proportional to α is either $\{\alpha, -\alpha\}$, $\{\alpha, \frac{1}{2}\alpha, -\frac{1}{2}\alpha, -\alpha\}$ or $\{2\alpha, \alpha, -\alpha, -2\alpha\}$.

A root α is *indivisible* if $\frac{1}{2}\alpha \notin R$. A root system R is *reduced* if all its roots are indivisible.

Let α be an indivisible root such that $2\alpha \in R$. Then s_α is a reflection which maps 2α into -2α. By 1.1.2, we see that $s_\alpha = s_{2\alpha}$. Therefore,

$$s_\alpha = s_{2\alpha} = I - \varphi((2\alpha)^* \otimes 2\alpha) = I - \varphi(2(2\alpha)^* \otimes \alpha)$$

and $(2\alpha)^* = \frac{1}{2}\alpha^*$.

An *automorphism* of R is a linear automorphism t of V such that $t(R) = R$. All automorphisms of R form a subgroup of $\mathrm{GL}(V)$ which we denote by $\mathrm{Aut}(R)$. For $\alpha \in R$, s_α is an automorphism of R. The subgroup of $\mathrm{Aut}(R)$ generated by s_α, $\alpha \in R$, is called the *Weyl group* of R and denoted by $W(R)$.

Let $t \in \mathrm{Aut}(R)$. Then $t\, s_\alpha\, t^{-1}$ is in $\mathrm{Aut}(R)$, i.e., $(t\, s_\alpha\, t^{-1})(R) = R$. Moreover,

$$(t\, s_\alpha t^{-1})(t\alpha) = -t\alpha$$

and $(t\, s_\alpha\, t^{-1})(th) = th$ for any $h \in H$. Hence, $t\, s_\alpha\, t^{-1}$ fixes the hyperplane tH. It follows that $t\, s_\alpha\, t^{-1}$ is a reflection with respect to root $t\alpha$. By 1.1.2, we have

$$t\, s_\alpha\, t^{-1} = s_{t\alpha}.$$

1.2.2. Lemma. *Let α be a root in R and $t \in \mathrm{Aut}(R)$. Then*

(i) $t\, s_\alpha\, t^{-1} = s_{t\alpha}$;
(ii) *the dual root $(t\alpha)^*$ of $t\alpha$ is equal to $(t^{-1})^*\alpha^*$.*

Proof. By (i), we have

$$\begin{aligned} s_{t\alpha}(v) = (t\, s_\alpha\, t^{-1})(v) = t(t^{-1}v - \alpha^*(t^{-1}v)\alpha) \\ = v - \alpha^*(t^{-1}v)t\alpha = (I - \varphi((t^{-1})^*\alpha \otimes t\alpha))(v) \end{aligned}$$

for any $v \in V$. □

1.2.3. Proposition. (i) $\operatorname{Aut}(R)$ *and* $W(R)$ *are finite groups.*
(ii) $W(R)$ *is a normal subgroup of* $\operatorname{Aut}(R)$.

Proof. Any element of $\operatorname{Aut}(R)$ induces a permutation of R. Moreover, since R spans V, this map is an injective homomorphism of $\operatorname{Aut}(R)$ into the group of permutations of R. Therefore, $\operatorname{Aut}(R)$ is finite.

By 1.2.2, for any $\alpha \in R$ and $t \in \operatorname{Aut}(R)$ we have $t\, s_\alpha\, t^{-1} = s_{t\alpha}$. Therefore, the conjugation by t maps the generators of $W(R)$ into generators of $W(R)$. Hence, $t\, W(R)\, t^{-1} \subset W(R)$ for any $t \in \operatorname{Aut}(R)$ and $W(R)$ is a normal subgroup of $\operatorname{Aut}(R)$. □

We define on V an bilinear form

$$(v \mid v') = \sum_{\alpha \in R} \alpha^*(v)\alpha^*(v').$$

This bilinear form is $\operatorname{Aut}(R)$-invariant. In fact, if $t \in \operatorname{Aut}(R)$, by 1.2.2, we have

$$\begin{aligned} (tv \mid tv') = \sum_{\alpha \in R} \alpha^*(tv)\alpha^*(tv') = \sum_{\alpha \in R} (t^*\alpha^*)(v)(t^*\alpha^*)(v') \\ = \sum_{\alpha \in R} (t^{-1}\alpha)^*(v)(t^{-1}\alpha)^*(v') = \sum_{\alpha \in R} \alpha^*(v)\alpha^*(v') = (v \mid v'). \end{aligned}$$

We need now a simple result in the representation theory of finite groups.

1.2.4. Theorem. *Let G be a finite group and π its representation on a finite-dimensional linear space V over the field k. Let U be an invariant subspace for π. Then there is a direct complement U' of U which is also invariant under π.*

Proof. Let P be a projection of V onto U. Put

$$Q = \frac{1}{\operatorname{Card} G} \sum_{g \in G} \pi(g^{-1})P\pi(g).$$

Clearly, for any $v \in V$, $\pi(g^{-1})P\pi(g)v \in U$ for any $g \in G$. Hence, $Qv \in U$. Moreover, we have

$$\pi(g^{-1})P\pi(g)u = \pi(g^{-1})\pi(g)u = u$$

for any $u \in U$. Therefore, it follows that

$$Qu = \frac{1}{\operatorname{Card} G} \sum_{g \in G} \pi(g^{-1})P\pi(g)u = u$$

for any $u \in U$, and Q is a projection onto U. Clearly, we get

$$Q\pi(g) = \frac{1}{\operatorname{Card} G} \sum_{h \in G} \pi(h^{-1})P\pi(hg) = \frac{1}{\operatorname{Card} G} \sum_{h \in G} \pi(gh^{-1})P\pi(h) = \pi(g)Q$$

by replacing h with hg^{-1}, for any $g \in G$. Therefore, $\ker Q$ and $U = \operatorname{im} Q$ are invariant under π. Hence, the assertion follows for $U' = \ker Q$. □

1.2.5. LEMMA. *The invariant bilinear form $(v, v') \longmapsto (v \mid v')$ on V is nondegenerate.*

PROOF. Let U be the orthogonal to V with respect to this bilinear form. Then U is invariant under the action of $\operatorname{Aut}(R)$. Since $\operatorname{Aut}(R)$ is a finite group by 1.2.3, by 1.2.4 there exist an $\operatorname{Aut}(R)$-invariant direct complement U' of U.

Let $\alpha \in R$. Then U and U' are invariant subspaces for s_α. Therefore, the one-dimensional eigenspace of s_α for eigenvalue -1 must be either in U or in U'. This implies that either $\alpha \in U$ or $\alpha \in U'$. On the other hand, we have

$$(\alpha \mid \alpha) = \sum_{\beta \in R} \beta^*(\alpha)^2 = 4 + \sum_{\beta \in R - \{\alpha\}} \beta^*(\alpha)^2 > 0$$

since the terms in the last sum are nonnegative integers. Hence, $\alpha \notin U$. It follows that $\alpha \in U'$.

Since R spans V, $U' = V$ and $U = \{0\}$. Therefore, the bilinear form is nondegenerate. □

1.2.6. PROPOSITION. (i) *The set R^* of all dual roots of R is a root system in V^*.*

(ii) *For any root $\alpha \in R$, we have $s_{\alpha^*} = s_\alpha^*$.*

(iii) *The map $t \longmapsto (t^{-1})^*$ is a group isomorphism of* $\operatorname{Aut}(R)$ *onto* $\operatorname{Aut}(R^*)$. *This isomorphism maps $W(R)$ onto $W(R^*)$.*

(iv) *For any $\alpha \in R$, the dual root of α^* is equal to α.*

PROOF. Assume that R^* does not span V^*. Then, we can take $u \neq 0$ in V such that $\alpha^*(u) = 0$ for all $\alpha \in R$. this in turn implies that $(u \mid v) = 0$ for all $v \in V$. This contradicts 1.2.5. Hence, R^* spans V^*.

Let $\alpha \in R$. Then

$$(s_\alpha^* f)(v) = f(s_\alpha v) = f(v - \alpha^*(v)\alpha) = f(v) - \alpha^*(v) f(\alpha) = (f - f(\alpha)\alpha^*)(v)$$

for any $v \in V$ and $f \in V^*$. Let $\psi : V \otimes V^* \longrightarrow \mathcal{L}(V^*)$ be the natural linear isomorphism given by $\psi(v \otimes f)(g) = g(v) f$ for any $v \in V$ and $f, g \in V^*$. Then $s_\alpha^* = I - \psi(\alpha \otimes \alpha^*)$. By 1.1.1, it follows that s_α^* is a reflection with respect to α^*.

On the other hand, for any $t \in \operatorname{Aut}(R)$ and root $\alpha \in R$, by 1.2.2, we have $t^*\alpha^* = (t^{-1}\alpha)^*$. Therefore, $t^*(R^*) = R^*$. In particular, $s_\alpha^*(R^*) = R^*$ for any $\alpha \in R$. By 1.1.2, $s_{\alpha^*} = s_\alpha^*$ is the unique reflection with respect to α^* which permutes the elements of R^*.

Finally,

$$s_{\alpha^*}(\beta^*) = (I - \psi(\alpha \otimes \alpha^*))(\beta^*) = \beta^* - \beta^*(\alpha)\alpha^*.$$

Since $\beta^*(\alpha) \in \mathbb{Z}$ for any $\alpha^*, \beta^* \in R^*$, it follows that R^* is a root system in V^*.

This in turn implies that the dual root of α^* is equal to α for any $\alpha^* \in R^*$.

Moreover, we see that for any $t \in \operatorname{Aut}(R)$, $t^* \in \operatorname{Aut}(R^*)$. Therefore, $t \longmapsto (t^{-1})^*$ is a group homomorphism. Since $R^{**} = R$, it must be an isomorphism. In addition, this isomorphism maps s_α into s_{α^*} for any root $\alpha \in R$, hence it must map $W(R)$ onto $W(R^*)$. □

We say that R^* is the *dual root system* of R.

1.2.7. LEMMA. *Let R be a root system and R^* the dual root system. The following conditions are equivalent:*

(i) *The root system R is reduced.*

(ii) *The root system R^* is reduced.*

PROOF. As we remarked before, if $\alpha, 2\alpha \in R$, we have $\alpha^*, \frac{1}{2}\alpha^* \in R^*$. Therefore, R^* is not reduced. □

Let K be a field extension of k. Then $V_K = K \otimes_k V$ is a K-linear space, and we have the obvious inclusion map $V \longrightarrow V_K$ mapping $v \in V$ into $1 \otimes v$. This identifies the root system R of V with a subset in V_K. Clearly, R spans V_K. Also, the reflections s_α extend by linearity to V_K. So, R defines a root system in V_K. We say that this root system is obtained by *extension of scalars* from the original one.

Let $V_\mathbb{Q}$ be the linear subspace of V spanned by the roots in R over the field of rational numbers $\mathbb{Q}$.

1.2.8. LEMMA.

$$\dim_\mathbb{Q} V_\mathbb{Q} = \dim_k V.$$

PROOF. Let S be a subset of R. If S is linearly independent over k, it is obviously linearly independent over $\mathbb{Q}$. We claim that the converse also holds. Let S be linearly independent over $\mathbb{Q}$. Assume that S is linearly dependent over k. Then we would have a nonzero element $(t_\alpha; \alpha \in S)$ of k^S such that $\sum_{\alpha\in S} t_\alpha \alpha = 0$. This would imply that

$$\sum_{\alpha\in S} t_\alpha \beta^*(\alpha) = 0$$

for all $\beta \in R$. Hence, the rank of the matrix $(\beta^*(\alpha); \alpha \in S, \beta \in R)$ is $<$ Card S. Since the matrix $(\beta^*(\alpha); \alpha \in S, \beta \in R)$ has integral coefficients, this clearly implies that the above system has a nonzero solution $(q_\alpha; \alpha \in S)$ with $q_\alpha \in \mathbb{Q}$. Hence, we have

$$\sum_{\alpha\in S} q_\alpha \beta^*(\alpha) = 0.$$

for all $\beta^* \in R^*$. Since R^* is a root system in V^*, this implies that $\sum_{\alpha\in S} q_\alpha \alpha = 0$, contradicting our assumption. □

Therefore, the k-linear map $k \otimes_\mathbb{Q} V_\mathbb{Q} \longrightarrow V$ defined by $t \otimes v \longmapsto tv$ is an isomorphism of k-linear spaces. By the construction, R is in $V_\mathbb{Q}$. For any $\alpha \in R$, the reflection s_α permutes the elements of R. Therefore, s_α maps $V_\mathbb{Q}$ into itself. Let α^* be the dual root of α. Then $\alpha^*(\beta) \in \mathbb{Z}$ for any root $\beta \in R$, and α^* takes rational values on $V_\mathbb{Q}$. Therefore, its restriction to $V_\mathbb{Q}$ can be viewed as a linear form on $V_\mathbb{Q}$. Moreover, $s_\alpha(v) = v - \alpha^*(v)\alpha$ for $v \in V_\mathbb{Q}$, i.e., the restriction of s_α to $V_\mathbb{Q}$ is a reflection by 1.1.1. It follows that R can be viewed as a root system in $V_\mathbb{Q}$. Therefore, the root system R in V can be viewed as obtained by extension of scalars from the root system R in $V_\mathbb{Q}$.

This reduces the study of root systems over arbitrary field k to root systems over $\mathbb{Q}$. On the other hand, we can consider the field extension from $\mathbb{Q}$ to the field of real numbers $\mathbb{R}$. Clearly, the study of root systems in linear spaces over $\mathbb{Q}$ is equivalent to the study of root systems in linear spaces over $\mathbb{R}$. The latter can be studied by more geometric methods.

1.3. Strings. Let R be a root system in V. Following the discussion at the end of preceding section, we can assume that V is a real linear space.

1.3.1. Lemma. *Let R be a root system in V. Then*

$$(v \mid w) = \sum_{\alpha \in R} \alpha^*(v)\alpha^*(w)$$

is an $\mathrm{Aut}(R)$*-invariant inner product on V.*

Proof. The form $(v, w) \longmapsto (v \mid w)$ is bilinear and symmetric. We also proved that it is $\mathrm{Aut}(R)$-invariant.

Moreover, we have

$$(v \mid v) = \sum_{\alpha \in R} \alpha^*(v)^2 \geq 0.$$

In addition, $(v \mid v) = 0$ implies that $\alpha^*(v) = 0$ for all $\alpha^* \in R^*$. Since roots in R^* span V^*, this in turn implies that $v = 0$. □

In the following we assume that V is equipped with this inner product. With respect to it, $\mathrm{Aut}(R) \subset \mathrm{O}(V)$. In particular, s_α are orthogonal reflections. Hence, for any $\alpha \in R$, the reflection hyperplane H is orthogonal to α, i.e., $H = \{v \in V \mid (\alpha \mid v) = 0\}$. This implies that

$$s_\alpha(v) = v - \frac{2(\alpha \mid v)}{(\alpha \mid \alpha)}\alpha$$

for any $v \in V$. The inner product on V defines a natural isomorphism of V with V^*. Under this isomorphism, the dual root α^* corresponds to $\frac{2}{(\alpha|\alpha)}\alpha$ for any root $\alpha \in R$.

For any two roots α, β in R we put

$$n(\alpha, \beta) = \beta^*(\alpha) = 2\frac{(\alpha \mid \beta)}{(\beta \mid \beta)}.$$

Clearly, we have the following result.

1.3.2. Lemma. *The following conditions are equivalent:*

(i) *The roots α and β are orthogonal;*
(ii) $n(\alpha, \beta) = 0$*;*
(iii) $n(\beta, \alpha) = 0$.

Hence, if α and β are not orthogonal,

$$0 \neq n(\alpha, \beta)\, n(\beta, \alpha) = 4\frac{(\alpha \mid \beta)^2}{\|\alpha\|^2\|\beta\|^2} = 4\cos^2(\alpha, \beta) \in \mathbb{Z},$$

where (α, β) is the angle between roots α and β. This implies the following result.

1.3.3. Lemma. $n(\alpha, \beta)\, n(\beta, \alpha) \in \{0, 1, 2, 3, 4\}$.

In addition, is α and β are not orthogonal, we have

$$\frac{n(\beta, \alpha)}{n(\alpha, \beta)} = \frac{\|\beta\|^2}{\|\alpha\|^2}.$$

We can assume that α is the shorter root, i.e., $\|\alpha\| \leq \|\beta\|$. Then, we must have $|n(\alpha, \beta)| \leq |n(\beta, \alpha)|$.

Assume first that α and β are neither orthogonal nor proportional. Therefore, $0 < \cos^2(\alpha, \beta) < 1$. It follows that $n(\alpha, \beta)\, n(\beta, \alpha) \in \{1, 2, 3\}$. This leads to the following table.

$n(\alpha,\beta)$	$n(\beta,\alpha)$	(α,β)	
1	1	$\frac{\pi}{3}$	$\|\beta\| = \|\alpha\|$
-1	-1	$\frac{2\pi}{3}$	$\|\beta\| = \|\alpha\|$
1	2	$\frac{\pi}{4}$	$\|\beta\| = \sqrt{2}\|\alpha\|$
-1	-2	$\frac{3\pi}{4}$	$\|\beta\| = \sqrt{2}\|\alpha\|$
1	3	$\frac{\pi}{6}$	$\|\beta\| = \sqrt{3}\|\alpha\|$
-1	-3	$\frac{5\pi}{6}$	$\|\beta\| = \sqrt{3}\|\alpha\|$

If α and β are proportional, $\cos^2(\alpha,\beta) = 1$ and $n(\alpha,\beta)\, n(\beta,\alpha) = 4$. Therefore, we have the following table.

$n(\alpha,\beta)$	$n(\beta,\alpha)$	
2	2	$\beta = \alpha$
-2	-2	$\beta = -\alpha$
1	4	$\beta = 2\alpha$
-1	-4	$\beta = -2\alpha$

The following result follows immediately from the above tables.

1.3.4. Lemma. *Let α and β be two non-proportional roots in R such that $\|\alpha\| \leq \|\beta\|$. Then, $n(\alpha,\beta) \in \{-1, 0, 1\}$.*

1.3.5. Theorem. *Let $\alpha, \beta \in R$.*

(i) *If $n(\alpha,\beta) > 0$ and $\alpha \neq \beta$, then $\alpha - \beta$ is a root.*
(ii) *If $n(\alpha,\beta) < 0$ and $\alpha \neq -\beta$, then $\alpha + \beta$ is a root.*

Proof. By changing β into $-\beta$ we see that (i) and (ii) are equivalent. Hence, it is enough to prove (i). Let $n(\alpha,\beta) > 0$ and $\alpha \neq \beta$. Then, we see from the tables that either $n(\alpha,\beta) = 1$ or $n(\beta,\alpha) = 1$.

In the first case, we have

$$s_\beta(\alpha) = \alpha - n(\alpha,\beta)\beta = \alpha - \beta \in R.$$

In the second case, we have

$$s_\alpha(\beta) = \beta - n(\beta,\alpha)\alpha = \beta - \alpha \in R.$$

□

This result has the following obvious reinterpretation.

1.3.6. Corollary. *Let $\alpha, \beta \in R$.*

(i) *If $(\alpha \mid \beta) > 0$ and $\alpha \neq \beta$, then $\alpha - \beta$ is a root.*
(ii) *If $(\alpha \mid \beta) < 0$ and $\alpha \neq -\beta$, then $\alpha + \beta$ is a root.*
(iii) *If $\alpha - \beta, \alpha + \beta \notin R \cup \{0\}$, then α is orthogonal to β.*

If $\alpha, \beta \in R$ and $\alpha - \beta, \alpha + \beta \notin R \cup \{0\}$ we say that α and β are *strongly orthogonal*. By the above corollary, strongly orthogonal roots are orthogonal.

1.3.7. Proposition. *Let α, β be two roots not proportional to each other. Then:*

(i) *The set of integers $I = \{j \in \mathbb{Z} \mid \beta + j\alpha \in R\}$ is an interval $[-q, p]$ in $\mathbb{Z}$ which contains 0.*
(ii) *Let $S = \{\beta + j\alpha \mid j \in I\}$. Then $s_\alpha(S) = S$ and $s_\alpha(\beta + p\alpha) = \beta - q\alpha$.*
(iii) *$p - q = -n(\beta,\alpha)$.*

PROOF. Clearly $0 \in I$. Let p, resp. $-q$, be the largest, resp. smallest, element in I. Assume that the assertion doesn't hold. Then there would exist $r, s \in [-q, p]$, $r, s \in I$ such that $s > r + 1$ and $r + k \notin I$ for $1 \leq k \leq s - r - 1$. Since $\beta + r\alpha$ would be a root and $\beta + (r+1)\alpha$ would not be a root, we would have $(\alpha \mid \beta + r\alpha) \geq 0$ by 1.3.6. Also, $\beta + s\alpha$ would be a root and $\beta + (s-1)\alpha$ would not be a root, hence we would have $(\alpha \mid \beta + s\alpha) \leq 0$ by 1.3.6. On the other hand, we would have

$$0 \geq (\alpha \mid \beta + s\alpha) = (\alpha \mid \beta) + s(\alpha \mid \alpha) > (\alpha \mid \beta) + r(\alpha \mid \alpha) = (\alpha \mid \beta + r\alpha) \geq 0$$

what is clearly impossible. Therefore, we have a contradiction and (i) holds.

Clearly,

$$s_\alpha(\beta + j\alpha) = \beta - \alpha^*(\beta)\alpha - j\alpha = \beta - (j + n(\beta, \alpha))\alpha$$

for any $j \in \mathbb{Z}$. Therefore, $s_\alpha(S) = S$. The function $j \longmapsto -(j + n(\beta, \alpha))$ is a decreasing bijection of I onto I. Therefore, $-p - n(\beta, \alpha) = -q$ and $p - q = -n(\beta, \alpha)$. This proves (iii). In addition, we have

$$s_\alpha(\beta + p\alpha) = \beta - (p + n(\beta, \alpha))\alpha = \beta - q\alpha$$

and (ii) holds. □

The set S is called the *α-string determined by β*. The root $\beta - q\alpha$ is the *start* and $\beta + p\alpha$ is the *end* of the α-string S. The integer $p + q$ is the *length* of the α-string S and denoted by $\ell(S)$.

1.3.8. COROLLARY. *The length of an α-string can be* 0, 1, 2 *or* 3.

PROOF. Let S be an α-string and β its start. Then $q = 0$ in the notation of 1.3.7. By (iii), it follows that $p = -n(\beta, \alpha)$. Since α and β are not proportional, and $p \geq 0$, by the tables on page 150, we see that the possible values of $n(\beta, \alpha)$ are 0, -1, -2 and -3. □

We also observe the following fact.

1.3.9. COROLLARY. *Let α, β be two roots not proportional to each other. Assume that $\alpha + \beta$ is a root. Let S be the α-string determined by β, $\beta - q\alpha$ its start and $\beta + p\alpha$ its end. Then*

$$\frac{\|\alpha + \beta\|^2}{\|\beta\|^2} = \frac{q+1}{p}.$$

PROOF. Since β and $\beta + \alpha$ are in S the length of S is ≥ 1.

If $\ell(S) = 1$, β has to be the start of the string and $\beta + \alpha$ the end, i.e., $q = 0$ and $p = 1$. By 1.3.7, we see that $s_\alpha(\beta) = \beta + \alpha$. It follows that $\|\alpha + \beta\| = \|\beta\|$ and the equality hold.

If $\ell(S) = 2$ and β is the start of S, the end of the string is $\beta + 2\alpha$. Hence, $q = 0$ and $p = 2$. By 1.3.7, we have $s_\alpha(\beta) = \beta + 2\alpha$. Hence, $\alpha^*(\beta) = -2$ and $(\alpha \mid \beta) = -\|\alpha\|^2$. From the table on the page 150, it follows that

$$\|\alpha + \beta\|^2 = \|\alpha\|^2 + 2(\alpha \mid \beta) + \|\beta\|^2 = -\|\alpha\|^2 + \|\beta\|^2 = \frac{1}{2}\|\beta\|^2$$

and the equality holds.

If $\ell(S) = 2$ and β is not the start of S, $\beta + \alpha$ has to be the end of S. Hence, $q = 1$ and $p = 1$. If $\gamma = \beta - \alpha$, γ is the start of S, and by the preceding case

$$\|\alpha + \gamma\|^2 = \frac{1}{2}\|\gamma\|^2.$$

Since $s_\alpha(\gamma) = \beta + \alpha$ this implies that

$$\|\beta\|^2 = \frac{1}{2}\|\beta + \alpha\|^2$$

and the equality holds.

If $\ell(S) = 3$ and β is the start of S, the end of the string is $\beta + 3\alpha$. Hence, $q = 0$ and $p = 3$. Hence, we have $n(\beta, \alpha) = -3$ by 1.3.7. From the table on page 150, we see that $\|\beta\|^2 = 3\|\alpha\|^2$. Moreover, $\alpha^*(\beta) = -3$ and $(\alpha \mid \beta) = -\frac{3}{2}\|\alpha\|^2$. It follows that

$$\|\alpha + \beta\|^2 = \|\alpha\|^2 + 2(\alpha \mid \beta) + \|\beta\|^2 = -2\|\alpha\|^2 + \|\beta\|^2 = \frac{1}{3}\|\beta\|^2$$

and the equality holds.

If $\ell(S) = 3$ and $\beta - \alpha$ is the start of S, the end of the string is $\beta + 2\alpha$. Hence, $q = 1$ and $p = 2$. By 1.3.7, we have $s_\alpha(\beta) = \beta + \alpha$ and $\|\beta\| = \|\alpha + \beta\|$. This implies that the equality holds.

Finally, if $\ell(S) = 3$ and $\beta - 2\alpha$ is the start of S, the end of the string is $\beta + \alpha$. Hence, $q = 2$ and $p = 1$. By 1.3.7, $\gamma = s_\alpha(\beta + \alpha)$ is the start of S and $s_\alpha(\beta) = \gamma + \alpha$. This implies that

$$\frac{\|\alpha + \beta\|^2}{\|\beta\|^2} = \frac{\|\gamma\|^2}{\|\gamma + \alpha\|^2} = 3$$

by the first case. This implies that the equality holds. □

2. Root system of a semisimple Lie algebra

2.1. Roots. Let $\mathfrak{g}$ be a semisimple Lie algebra over an algebraically closed field k. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$. By 4.5.3.1 and 4.5.3.3, $\mathfrak{h}$ is a maximal abelian Lie subalgebra consisting of semisimple elements.

For any linear form $\alpha \in \mathfrak{h}^*$, we put

$$\mathfrak{g}_\alpha = \{x \in \mathfrak{g} \mid [h, x] = \alpha(h)x, h \in \mathfrak{h}\}.$$

Clearly, $\mathfrak{g}_\alpha$ is a linear subspace of $\mathfrak{g}$. Moreover, we have the following result.

2.1.1. Lemma. *Let $\alpha, \beta \in \mathfrak{h}^*$. Then*

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}.$$

Proof. Let $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_\beta$. Then, for any $H \in \mathfrak{h}$ we have

$$\begin{aligned}(\operatorname{ad} H)([x, y]) &= [(\operatorname{ad} H)x, y] + [x, (\operatorname{ad} H)y] \\ &= \alpha(H)[x, y] + \beta(H)[x, y] = (\alpha + \beta)(H)[x, y].\end{aligned}$$

Hence, we have $[x, y] \in \mathfrak{g}_{\alpha+\beta}$. □

2.1.2. Lemma. $\mathfrak{g}_0 = \mathfrak{h}$.

Proof. By 4.5.5.1, there exists a regular element $h_0 \in \mathfrak{h}$ such that $\mathfrak{h} = \mathfrak{g}(h_0, 0)$. Since h_0 is semisimple by 4.5.3.4, we see that

$$\mathfrak{h} = \{x \in \mathfrak{g} \mid [h_0, x] = 0\}.$$

Therefore, $\mathfrak{g}_0 \subset \mathfrak{h}$. On the other hand, since $\mathfrak{h}$ is abelian, for any $H, H' \in \mathfrak{h}$, we have $[H, H'] = 0$ and $\operatorname{ad}(H)(H') = 0$. Hence $H' \in \mathfrak{g}_0$. It follows that $\mathfrak{h} = \mathfrak{g}_0$. □

If $\alpha \neq 0$ and $\mathfrak{g}_\alpha \neq \{0\}$, α is a *root of* $\mathfrak{g}$ *with respect to* $\mathfrak{h}$. We denote by R the set of all roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$.

2.1.3. LEMMA.
$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha.$$

PROOF. Since $\mathfrak{h}$ is an abelian Lie algebra consisting of semisimple elements, the linear maps $\operatorname{ad} H$, $H \in \mathfrak{h}$ form a commuting family of semisimple linear maps on $\mathfrak{g}$. This immediately implies the statement. □

In particular, the set R is finite.

2.1.4. LEMMA. (i) *Let $\alpha, \beta \in R \cup \{0\}$ such that $\alpha + \beta \neq 0$. Then $\mathfrak{g}_\alpha$ is orthogonal to $\mathfrak{g}_\beta$ with respect to the Killing form.*

The restriction of the Killing form to $\mathfrak{g}_\alpha \times \mathfrak{g}_{-\alpha}$ induces a nondegenerate pairing.

The restriction of the Killing form to $\mathfrak{h} \times \mathfrak{h}$ is nondegenerate.

(ii) *Let $x \in \mathfrak{g}_\alpha$, $y \in \mathfrak{g}_{-\alpha}$ and $h \in \mathfrak{h}$. Then $[x, y] \in \mathfrak{h}$ and*
$$B(h, [x, y]) = \alpha(h) B(x, y).$$

PROOF. (i) Let $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_\beta$. Then, for any $\gamma \in R \cup \{0\}$, we have
$$(\operatorname{ad}(x) \operatorname{ad}(y))(\mathfrak{g}_\gamma) \subset \mathfrak{g}_{\gamma+\alpha+\beta}.$$
It follows that the block diagonal matrix of $\operatorname{ad}(x) \operatorname{ad}(y)$ with respect to the decomposition $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\gamma \in R} \mathfrak{g}_\gamma$ has zero blocks on the diagonal. Therefore, $B(x, y) = \operatorname{tr}(\operatorname{ad}(x) \operatorname{ad}(y)) = 0$.

This implies that $\mathfrak{g}_\beta$ is orthogonal to $\mathfrak{g}_\alpha$ for all $\beta \neq -\alpha$. Since the Killing form is nondegenerate on $\mathfrak{g}$ by 4.4.1.1, it follows that the pairing of $\mathfrak{g}_\alpha$ with $\mathfrak{g}_{-\alpha}$ is nonedgenerate for all $\alpha \in R \cup \{0\}$.

(ii) Clearly, we have $[x, y] \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subset \mathfrak{g}_0 = \mathfrak{h}$. Therefore, it follows that
$$B(h, [x, y]) = -B([x, h], y) = B([h, x], y) = \alpha(h) B(x, y).$$
□

2.1.5. PROPOSITION. *Let $\alpha \in R$. Then:*

(i) $\dim \mathfrak{g}_\alpha = 1$.

(ii) *The space $\mathfrak{h}_\alpha = [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subset \mathfrak{h}$ is one-dimensional. There exists a unique $H_\alpha \in \mathfrak{h}_\alpha$ such that $\alpha(H_\alpha) = 2$.*

(iii) *The subspace*
$$\mathfrak{s}_\alpha = \mathfrak{h}_\alpha \oplus \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}$$
is a Lie subalgebra of $\mathfrak{g}$.

(iv) *Let $X_\alpha \in \mathfrak{g}_\alpha$, $X_\alpha \neq 0$. Then there exists a unique element $X_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $[X_\alpha, X_{-\alpha}] = H_\alpha$. Let $\varphi : \mathfrak{sl}(2, k) \longrightarrow \mathfrak{g}$ be the linear map defined by*
$$\varphi(e) = X_\alpha, \quad \varphi(f) = X_{-\alpha}, \quad \varphi(h) = H_\alpha.$$
Then φ is a Lie algebra isomorphism of $\mathfrak{sl}(2, k)$ onto $\mathfrak{s}_\alpha$.

PROOF. Since the restriction of the Killing form B to $\mathfrak{h} \times \mathfrak{h}$ is nondegenerate by 2.1.4, there exists $h_\alpha \in \mathfrak{h}$ such that
$$B(h_\alpha, H) = \alpha(H) \text{ for all } H \in \mathfrak{h}.$$
Let $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_{-\alpha}$. Then, $[x, y] \in \mathfrak{h}$ and
$$B(H, [x, y]) = \alpha(H) B(x, y) = B(H, B(x, y) h_\alpha)$$

by 2.1.4.(ii). Therefore, by nondegeneracy of the Killing form on $\mathfrak{h} \times \mathfrak{h}$, it follows that

$$[x, y] = B(x, y)h_\alpha.$$

Hence, $\mathfrak{h}_\alpha$ is at most one-dimensional. On the other hand, since the Killing form induces a nondegenerate pairing of $\mathfrak{g}_\alpha$ with $\mathfrak{g}_{-\alpha}$ by 2.1.4.(i), we conclude that for any $x \in \mathfrak{g}_\alpha$, $x \neq 0$, there exists $y \in \mathfrak{g}_{-\alpha}$ such that $B(x, y) \neq 0$. This in turn implies that $[x, y] \neq 0$. Therefore, $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ is different from zero, and $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = kh_\alpha = \mathfrak{h}_\alpha$.

Clearly, we can choose x and y so that $B(x, y) = 1$. Then $[x, y] = h_\alpha$. Moreover, we have

$$[h_\alpha, x] = \alpha(h_\alpha)x, \quad [h_\alpha, y] = -\alpha(h_\alpha)y.$$

Then, the subspace $\mathfrak{n}$ spanned by x, y and h_α is a Lie subalgebra of $\mathfrak{g}$. Assume that $\alpha(h_\alpha) = 0$. Then $\mathcal{C}^1(\mathfrak{n}) = kh_\alpha$ and $\mathcal{C}^2(\mathfrak{n}) = 0$. Hence, $\mathfrak{n}$ is nilpotent. In particular, it is solvable and by 4.1.6.3, there exists a basis in which $\operatorname{ad} z$, $z \in \mathfrak{n}$, are represented by upper triangular matrices. In addition, in this basis, all matrices of $\operatorname{ad} z$, $z \in \mathcal{C}^1(\mathfrak{n})$, are nilpotent. In particular $\operatorname{ad} h_\alpha$ is nilpotent. This contradict the fact that h_α is semisimple. Therefore, $\alpha(h_\alpha) \neq 0$.

It follows that we can find $H_\alpha \in \mathfrak{h}_\alpha$ such that $\alpha(H_\alpha) = 2$. Moreover, it is clear that for any $X_\alpha \in \mathfrak{g}_\alpha$, $X_\alpha \neq 0$, one can find $X_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that

$$[X_\alpha, X_{-\alpha}] = H_\alpha.$$

Then we have

$$[H_\alpha, X_\alpha] = \alpha(H_\alpha)X_\alpha = 2X_\alpha \quad [H_\alpha, X_{-\alpha}] = -\alpha(H_\alpha)X_{-\alpha} = -2X_{-\alpha}.$$

Therefore, the linear map $\varphi : \mathfrak{sl}(2, k) \longrightarrow \mathfrak{g}$ defined by

$$\varphi(e) = X_\alpha, \quad \varphi(f) = X_{-\alpha}, \quad \varphi(h) = H_\alpha$$

is an isomorphism of $\mathfrak{sl}(2, k)$ onto the Lie subalgebra $\mathfrak{s}_\alpha$ spanned by X_α, $X_{-\alpha}$ and H_α.

Assume that $\dim \mathfrak{g}_\alpha > 1$. Let $y \in \mathfrak{g}_{-\alpha}$, $y \neq 0$. Then $x \longmapsto B(x, y)$ is a linear form on $\mathfrak{g}_\alpha$, and there exists $X_\alpha \in \mathfrak{g}_\alpha$, $X_\alpha \neq 0$, such that $B(X_\alpha, y) = 0$. We pick an $X_{-\alpha} \in \mathfrak{g}_{-\alpha}$ which satisfies the above conditions. then, the composition of φ with the adjoint representation defines a representation ρ of $\mathfrak{sl}(2, k)$ on $\mathfrak{g}$. Also, by the preceding discussion, we have

$$\rho(e)y = [X_\alpha, y] = B(X_\alpha, y)h_\alpha = 0$$

and

$$\rho(h)y = [H_\alpha, y] = -2\alpha(H_\alpha) = -2y.$$

It follows that y is a primitive vector of weight -2 for ρ. Since ρ is obviously finite-dimensional, this contradicts 4.4.7.2. Hence, $\dim \mathfrak{g}_\alpha = 1$.

Since $\dim \mathfrak{g}_{-\alpha} = 1$, the vector $X_{-\alpha}$ such that $[X_\alpha, X_{-\alpha}] = H_\alpha$ is unique. □

This result has the following consequences.

2.1.6. Corollary.

$$\dim \mathfrak{g} = \operatorname{rank} \mathfrak{g} + \operatorname{Card}(R).$$

2.1.7. Corollary. *For any $H, H' \in \mathfrak{h}$, we have*

$$B(H, H') = \sum_{\alpha \in R} \alpha(H)\alpha(H').$$

PROOF. Clearly, $\mathfrak{g}_\alpha$, $\alpha \in R \cup \{0\}$, are invariant for $\operatorname{ad}(H)$. Moreover, since $\mathfrak{h}$ is abelian, $\operatorname{ad}(H)$ induces 0 on $\mathfrak{h}$. Moreover, it induces multiplication by $\alpha(H)$ on $\mathfrak{g}_\alpha$, $\alpha \in R$. Since $\dim \mathfrak{g}_\alpha = 1$, the assertion follows. □

2.1.8. COROLLARY. *The set R spans $\mathfrak{h}^*$.*

PROOF. Let $H \in \mathfrak{h}$ be such that $\alpha(H) = 0$ for all $\alpha \in R$. Then we have $\operatorname{ad} H = 0$ and, by 4.1.3.2, this implies that $H = 0$. This clearly implies that R spans $\mathfrak{h}^*$. □

2.1.9. LEMMA. *Let $\alpha, \beta \in R$. Then*

(i) $\beta(H_\alpha) \in \mathbb{Z}$;
(ii) $B(H_\alpha, H_\beta) \in \mathbb{Z}$.

PROOF. Let $\varphi : \mathfrak{sl}(2,k) \longrightarrow \mathfrak{g}$ be the Lie algebra morphism satisfying

$$\varphi(e) = X_\alpha, \quad \varphi(f) = X_{-\alpha}, \quad \varphi(h) = H_\alpha,$$

which constructed in the proof of 2.1.5. Since the composition ρ of φ with the adjoint representation is a representation of $\mathfrak{sl}(2,k)$, the eigenvalues of $\rho(h)$ are integers by 4.4.7.4. Therefore, $\beta(H_\alpha) \in \mathbb{Z}$ for all $\beta \in R$. This proves (i).

(ii) follows from 2.1.7. □

2.1.10. THEOREM. *Let $\mathfrak{g}$ be a semisimple Lie algebra over k and $\mathfrak{h}$ a Cartan subalgebra in $\mathfrak{g}$. Let R be the set of all roots of $(\mathfrak{g}, \mathfrak{h})$. Then:*

(i) *R is a reduced root system in $\mathfrak{h}^*$;*
(ii) *the dual root system R^* of R is equal to $\{H_\alpha; \alpha \in R\}$.*

PROOF. By 2.1.8 we know that R spans $\mathfrak{h}^*$. Also, by 2.1.9, we know that $\beta(H_\alpha)$ are integers for any $\alpha, \beta \in R$.

Fix $\alpha, \beta \in R$. We claim that $\beta - \beta(H_\alpha)\alpha$ is also a root. Let $y \in \mathfrak{g}_\beta$, $y \neq 0$, and $p = \beta(H_\alpha)$. Let $\varphi : \mathfrak{sl}(2,k) \longrightarrow \mathfrak{g}$ be the Lie algebra morphism satisfying

$$\varphi(e) = X_\alpha, \quad \varphi(f) = X_{-\alpha}, \quad \varphi(h) = H_\alpha,$$

which was constructed in the proof of 2.1.5. Since the composition ρ of φ with the adjoint representation is a representation of $\mathfrak{sl}(2,k)$, and

$$\rho(h)y = [H_\alpha, y] = \beta(H_\alpha)y = py.$$

By 4.4.7.4, we have $z = \rho(f)^p y \neq 0$ if $p > 0$; and $z = \rho(e)^{-p} y \neq 0$ if $p < 0$. In both cases, $z \in \mathfrak{g}_{\beta - p\alpha}$. Hence, $\beta - p\alpha \in R$.

Since $\alpha(H_\alpha) = 2$, by 1.1.1, the linear map $s_\alpha = I - \varphi(H_\alpha \otimes \alpha)$ is a reflection with respect to α. Also, by the above discussion,

$$s_\alpha(\beta) = \beta - \beta(H_\alpha)\alpha \in R.$$

It follows that $s_\alpha(R) \subset R$. Hence, R is a root system. Also, H_α is the dual root of α for any $\alpha \in R$.

It remains to prove that R is reduced. Assume that $\alpha \in R$ is such that $2\alpha \in R$. Let $y \in \mathfrak{g}_{2\alpha}$, $y \neq 0$. If we consider the above representation ρ of $\mathfrak{sl}(2,k)$ on $\mathfrak{g}$, we see that

$$\rho(h)y = [H_\alpha, y] = 2\alpha(H_\alpha)y = 4y.$$

Also,

$$\rho(e)y = [X_\alpha, y] \subset [\mathfrak{g}_\alpha, \mathfrak{g}_{2\alpha}] \subset \mathfrak{g}_{3\alpha} = 0,$$

since R is a root system. It follows that y is a primitive vector for ρ. On the other hand, we have

$$4y = \rho(h)y = \rho([e,f])y = \rho(e)\rho(f)y = [X_\alpha, [X_{-\alpha}, y]].$$

Since $[X_{-\alpha}, y]$ is in $[\mathfrak{g}_{-\alpha}, \mathfrak{g}_{2\alpha}] \subset \mathfrak{g}_\alpha$, it must be proportional to X_α. Hence, the above commutator is zero, i.e., $4y = 0$ and $y = 0$ contradicting our assumption. It follows that R is reduced. □

Let $\alpha \in R$. Then $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ are one-dimensional. Moreover, we can find $X_\alpha \in \mathfrak{g}_\alpha$ and $X_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $[X_\alpha, X_{-\alpha}] = H_\alpha$. As we already remarked, the subspace $\mathfrak{s}_\alpha = \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ is a Lie subalgebra of $\mathfrak{g}$ and the linear map $\varphi : \mathfrak{sl}(2,k) \longrightarrow \mathfrak{s}_\alpha$ given by $\varphi(h) = H_\alpha$, $\varphi(e) = X_\alpha$ and $\varphi(f) = X_{-\alpha}$, is an isomorphism of Lie algebras.

2.1.11. PROPOSITION. (i) *Let $\alpha, \beta \in R$ be two non-proportional roots. Let S be the α-string determined by β and*

$$\mathfrak{g}_S = \bigoplus_{\gamma \in S} \mathfrak{g}_\gamma.$$

Then $\mathfrak{g}_S$ is an irreducible $\mathfrak{s}_\alpha$-submodule.

(ii) *If $\alpha, \beta \in R$ are such that $\alpha + \beta \in R$, then*

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}.$$

PROOF. (i) Clearly, by 2.1.5, $\dim \mathfrak{g}_S = \ell(S) + 1$.

Let ρ be the composition of φ with the adjoint representation. Let $\gamma \in S$. Then, we have

$$\rho(h)(\mathfrak{g}_\gamma) \subset [\mathfrak{h}, \mathfrak{g}_\gamma] = \mathfrak{g}_\gamma,\ \rho(e)(\mathfrak{g}_\gamma) \subset [\mathfrak{g}_\alpha, \mathfrak{g}_\gamma] = \mathfrak{g}_{\alpha+\gamma},\ \rho(f)(\mathfrak{g}_\gamma) \subset [\mathfrak{g}_{-\alpha}, \mathfrak{g}_\gamma] = \mathfrak{g}_{-\alpha+\gamma}.$$

Therefore, $\mathfrak{g}_S$ is invariant for ρ.

By 1.3.7, there exist $p, q \in \mathbb{Z}$ such that $I = \{j \in \mathbb{Z} \mid \beta + j\alpha \in S\} = [-q, p]$. This implies that $\ell(S) = p + q$ and $\dim \mathfrak{g}_S = p + q + 1$. For any nonzero $y \in \mathfrak{g}_{\beta+p\alpha}$, we have

$$\rho(e)y = [X_\alpha, y] \in \mathfrak{g}_{\beta+(p+1)\alpha} = \{0\},$$

since $\beta + (p+1)\alpha \notin R$. Also, by 1.3.7, we have

$$\rho(h)y = [H_\alpha, y] = (\beta + p\alpha)(H_\alpha)y = (\alpha^*(\beta) + 2p)y = (n(\beta, \alpha) + 2p)y = (p + q)y.$$

Hence, y is a primitive vector of weight $p + q$.

Let

$$\{0\} = V_0 \subset V_1 \subset \ldots V_{n-1} \subset V_n = \mathfrak{g}_S$$

be a maximal flag of invariant subspaces for the representation ρ. Then, there exists $1 \leq p \leq n$, such that $y \in V_p$ and $y \notin V_{p-1}$. Hence, the projection $\bar{y}$ of y into the quotient module V_p/V_{p-1} is a primitive vector of weight $p + q$. Since this module is irreducible, by 4.4.7.3, its dimension is equal to $p + q + 1$. Hence, it is equal to the dimension of $\mathfrak{g}_S$. It follows that $p = n = 1$ and $\mathfrak{g}_S$ is irreducible.

(ii) Assume that that $\alpha, \beta \in R$ and $\alpha + \beta \in R$. Since R is reduced, α and β are not proportional. Moreover, if S is the α-string determined by β, $\beta - q\alpha$ its start and $\beta + p\alpha$ its end, we see that $p \geq 1$. Let $\gamma = \beta + j\alpha \in S$. Then for any $y \in \mathfrak{g}_\gamma$, we have

$$\rho(h)y = [H_\alpha, y] = \gamma(H_\alpha)y = (\beta(H_\alpha) + 2j)y = (n(\beta, \alpha) + 2j)y = (q - p + 2j)y.$$

Since $\mathfrak{g}_S$ is irreducible by (i), by 4.4.7.3, the only primitive vectors in it have weight $p+q$, i.e., correspond to $j=p$. Therefore, $\mathfrak{g}_\beta$ doesn't contain any primitive vectors and $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \rho(e)(\mathfrak{g}_\beta) \neq 0$. On the other hand, $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta]$ is contained in the one-dimensional subspace $\mathfrak{g}_{\alpha+\beta}$. Therefore, we have $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$. □

2.2. Structure constants. From the preceding discussion, we see that a semisimple Lie algebra $\mathfrak{g}$ is spanned by its Cartan subalgebra $\mathfrak{h}$ and the root spaces $\mathfrak{g}_\alpha$, $\alpha \in R$. Moreover, for each α in R, we can pick a vector X_α which spans $\mathfrak{g}_\alpha$ such that $[X_\alpha, X_{-\alpha}] = H_\alpha$. In this case,

$$[H, X_\alpha] = \alpha(H) X_\alpha$$

for any $H \in \mathfrak{h}$ and $\alpha \in R$. Therefore, the structure of $\mathfrak{g}$ is completely determined by the commutators $[X_\alpha, X_\beta]$ where $\alpha, \beta \in R$ and $\beta \neq -\alpha$. Clearly, in this case, we have

$$[X_\alpha, X_\beta] = \begin{cases} 0, & \text{if } \alpha+\beta \notin R; \\ N_{\alpha,\beta} X_{\alpha+\beta}, & \text{if } \alpha+\beta \in R; \end{cases}$$

where $N_{\alpha,\beta}$ are nonzero elements in k by 2.1.11. These *structure constants* completely determine the additional structure of $\mathfrak{g}$.

Let $\alpha \in R$. Let $\varphi : \mathfrak{sl}(2,k) \longrightarrow \mathfrak{g}$ be the Lie algebra morphism satisfying $\varphi(e) = X_\alpha$, $\varphi(f) = X_{-\alpha}$ and $\varphi(h) = H_\alpha$ which was constructed in the proof of 2.1.5. Let β be another root such that $\alpha+\beta$ is a root. Let S be the α-string determined by β. As before, let $\beta - q\alpha$ be the start and $\beta + p\alpha$ be the end of S. By 2.1.11 we know that $\mathfrak{g}_S$ is an irreducible $\mathfrak{sl}(2,k)$-module of dimension $\ell(S)+1 = p+q+1$. Since $\beta+(p+1)\alpha$ is not a root, $[X_\alpha, X_{\beta+p\alpha}] = 0$ and $v = X_{\beta+p\alpha}$ is a primitive vector of this module. If we put

$$v_k = \frac{1}{k!}(\operatorname{ad}\varphi(f))^k v = \frac{1}{k!}(\operatorname{ad} X_{-\alpha})^k X_{\beta+p\alpha} \in \mathfrak{g}_{\beta+(p-k)\alpha}$$

for $0 \leq k \leq p+q$, we see from 4.4.7.3 that they form a basis of $\mathfrak{g}_S$. In addition, we have

$$\operatorname{ad} X_{-\alpha}(\operatorname{ad} X_\alpha(v_p)) = \operatorname{ad}\varphi(f)(\operatorname{ad}\varphi(e)(v_p)) = (q+1)\operatorname{ad}\varphi(f)(v_{p-1}) = p(q+1)v_p.$$

Since v_p is proportional to X_β, this immediately implies that

$$\operatorname{ad} X_{-\alpha}(\operatorname{ad} X_\alpha) X_\beta = p(q+1) X_\beta.$$

On the other hand, we have

$$\operatorname{ad} X_{-\alpha}(\operatorname{ad} X_\alpha) X_\beta = [X_{-\alpha}, [X_\alpha, X_\beta]] = N_{\alpha,\beta}[X_{-\alpha}, X_{\alpha+\beta}] = N_{\alpha,\beta} N_{-\alpha,\alpha+\beta} X_\beta.$$

Therefore, we established the following result.

2.2.1. Lemma. *Let α, β be two roots such that $\alpha+\beta \in R$. Then*

$$N_{\alpha,\beta} N_{-\alpha,\alpha+\beta} = p(q+1).$$

By the invariance of the Killing form, we have

$$\begin{aligned} B(H_\alpha, H_\alpha) = B([X_\alpha, X_{-\alpha}], H_\alpha) &= -B([X_{-\alpha}, X_\alpha], H_\alpha) \\ &= -B(X_\alpha, [H_\alpha, X_{-\alpha}]) = 2B(X_\alpha, X_{-\alpha}). \end{aligned}$$

Hence, we have the following result.

2.2.2. Lemma. *Let $\alpha \in R$. Then*

$$\frac{1}{2} B(H_\alpha, H_\alpha) = B(X_\alpha, X_{-\alpha}).$$

In addition, by the invariance of the Killing form, we have

$$N_{-\alpha,\alpha+\beta}B(X_\beta, X_{-\beta}) = B([X_{-\alpha}, X_{\alpha+\beta}], X_{-\beta})$$
$$= -B(X_{\alpha+\beta}, [X_{-\alpha}, X_{-\beta}]) = -N_{-\alpha,-\beta}B(X_{\alpha+\beta}, X_{-\alpha-\beta}).$$

Therefore, by 2.2.2, we have the following result.

2.2.3. LEMMA. *Let α, β be two roots such that $\alpha + \beta \in R$. Then we have*

$$N_{-\alpha,\alpha+\beta}B(H_\beta, H_\beta) = -N_{-\alpha,-\beta}B(H_{\alpha+\beta}, H_{\alpha+\beta}).$$

By 2.1.7, the restriction of the Killing form on $\mathfrak{h}\times\mathfrak{h}$ is invariant under the action of $\mathrm{Aut}(R^*)$. Consider the linear subspace of $\mathfrak{h}^*$ spanned by R over $\mathbb{Q}$. By 2.1.7, the Killing form restricted to this subspace is an inner product, which agrees with the one we used in 1.3. Under this inner product, the dual root of H_α is identified with $\alpha = \frac{2H_\alpha}{(H_\alpha|H_\alpha)}$. Hence, we have

$$\|\alpha\| = \frac{2}{(H_\alpha|H_\alpha)}\|H_\alpha\| = 2\frac{1}{\|H_\alpha\|},$$

This implies that

$$B(H_\alpha, H_\alpha) = 4\frac{1}{\|\alpha\|^2}.$$

Therefore, the above result implies that

$$\frac{N_{-\alpha,\alpha+\beta}}{\|\beta\|^2} = -\frac{N_{-\alpha,-\beta}}{\|\alpha+\beta\|^2}$$

and

$$\frac{N_{-\alpha,-\beta}}{N_{-\alpha,\alpha+\beta}} = -\frac{\|\alpha+\beta\|^2}{\|\beta\|^2} = -\frac{q+1}{p}$$

by 1.3.9. By multiplying this with 2.2.1 we get

$$N_{\alpha,\beta}N_{-\alpha,-\beta} = -(q+1)^2.$$

2.2.4. LEMMA. *Let α, β be two roots such that $\alpha + \beta \in R$. Then we have*

$$N_{\alpha,\beta}N_{-\alpha,-\beta} = -(q+1)^2.$$

We are going to see later that we can chose X_α, $\alpha \in R$, in such way that $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$ for all $\alpha, \beta \in R$. This implies that in this situation $N_{\alpha,\beta}^2 = (q+1)^2$, i.e., the structure constants $N_{\alpha,\beta}$ are up to a sign equal to $q+1$. In particular, they are integers.

3. Bases of root systems and Weyl groups

3.1. Indivisible roots. Let R be a root system in V. Let R_0 be the set of all indivisible roots in R.

3.1.1. LEMMA. *Let R be a root system in V.*

(i) *The set R_0 of all indivisible roots α in R is a root system in V.*
(ii) *The root system R_0 is reduced.*
(iii) *The Weyl group $W(R_0)$ iis equal to $W(R)$.*

PROOF. Clearly, the set of indivisible roots in R spans V. Let α be in R_0 and $\beta \in R$. Assume that $s_\beta(\alpha)$ is divisible. Then we would have $s_\beta(\alpha) = 2\gamma$ for some root γ in R. This in turn would imply that $\alpha = 2s_\beta(\gamma)$, contrary to our assumption. Hence $s_\beta(\alpha)$ is indivisible. It follows that $s_\alpha(R_0) = R_0$.

(ii) This is clear from the definition.

(iii) Let α be a indivisible root in R such that 2α is a root. Then $s_\alpha = s_{2\alpha}$. Therefore, $W(R)$ and $W(R_0)$ have the same set of generators. □

3.2. Bases of root systems. A *basis* B of a root system R is a finite set of roots such that

(i) B is a basis of V;

(ii) any root α in R is

$$\alpha = \sum_{\beta \in B} m_\beta \beta$$

where m_β, $\beta \in B$, are integers and have the same sign.

We want to show that bases of root systems exist. First, we show that we can assume that R is a reduced root system.

3.2.1. LEMMA. *Let R be a root system and R_0 the root system of indivisible roots in R.*

(i) *If B is a basis of the root system R all roots in B are indivisible and B is a basis of R_0.*

(ii) *If B is a basis of R_0, B is a basis of R.*

PROOF. (i) Assume that α is a root in B. If α is not indivisible, $\alpha = 2\beta$ for $\beta \in R$. Hence, we would have $\beta = \frac{1}{2}\alpha$, contradicting the definition of basis. Hence all roots in B are indivisible. Clearly, B is a basis of R_0.

(ii) is evident. □

Now we can assume that V is a real vector space. Also, we can assume that V is equipped with an inner product which is invariant under $\mathrm{Aut}(R)$.

We start with a simple result.

3.2.2. LEMMA. *Let B be a finite set of vectors in V such that $\alpha, \beta \in B$ implies that $(\alpha \mid \beta) \leq 0$ for $\alpha \neq \beta$. Assume that there exists a vector $v \in V$ such that $(v \mid \alpha) > 0$ for all $\alpha \in B$. Then B is linearly independent.*

PROOF. Assume that B is linearly dependent. Let $\sum_{\alpha \in B} c_\alpha \alpha = 0$ be a nontrivial linear combination of vectors in B. Let $B' = \{\alpha \in B \mid c_\alpha > 0\}$ and $B'' = \{\alpha \in B \mid c_\alpha < 0\}$. Clearly, either B' or B'' must be nonempty. Let

$$u = \sum_{\alpha \in B'} c_\alpha \alpha = - \sum_{\alpha \in B''} c_\alpha \alpha.$$

Then, by our assumptions, we have

$$(u \mid u) = - \sum_{\alpha \in B'} \sum_{\beta \in B''} c_\alpha c_\beta (\alpha|\beta) \leq 0.$$

Therefore, we must have $u = 0$. Assume that B' is nonempty. Then

$$0 = (v \mid u) = \sum_{\alpha \in B'} c_\alpha (v \mid \alpha).$$

Since all terms in the sum are strictly positive, we have a contradiction. If B'' is nonempty, the proof is analogous. □

Now we prove the existence of bases of root systems.

3.2.3. Lemma. *Let R be a root system in V and α an indivisible root in R. Then there exists a basis B of R containing α.*

Proof. As we remarked, we can assume that R is reduced. Let V' be the subspace orthogonal to α. Any root β in R different from α or $-\alpha$ is not orthogonal to V'. Therefore, there exists a vector $u \in V'$ such that $(u \mid \beta) \neq 0$ for all $\beta \in R$ different from α or $-\alpha$. Assume that a is the smallest number of the set $\{|(u \mid \beta)| \mid \beta \in R - \{-\alpha, \alpha\}\}$. Let $\epsilon < \frac{a}{2}$. By continuity, there exists a neighborhood U of u in V such that $v \in U$ implies that $|(v \mid \alpha)| < \epsilon$ and $|(v \mid \beta)| > \epsilon$ for $\beta \in R - \{-\alpha, \alpha\}$. We pick $v \in U$ such that $(v \mid \alpha) > 0$. Then we have $0 < (v \mid \alpha) < \epsilon$ and $|(v \mid \beta)| > \epsilon$ for $\beta \in R - \{-\alpha, \alpha\}$.

Put $R_v^+ = \{\beta \in R \mid (v \mid \beta) > 0\}$. Clearly, R is a disjoint union of R_v^+ and $-R_v^+$. Moreover, α is in R_v^+.

We say that a root β in R_v^+ is *indecomposable* if it is not equal to sum of two vectors $\gamma, \delta \in R_v^+$. Let B be the set of all indecomposable roots in R_v^+. We claim that B is a basis of the root system R.

The set $F = \{(v \mid \beta) \mid \beta \in R_v^+\}$ is a finite set of positive numbers. By our construction, $c = (v \mid \alpha)$ is the smallest element of this set. For any other root β in R_v^+ we have $(v \mid \beta) > c$. If α is decomposable, we would have $\alpha = \beta + \gamma$ where $\beta, \gamma \in R_v^+$. This would imply that

$$c = (v \mid \alpha) = (v \mid \beta) + (v \mid \gamma)$$

with $(v \mid \beta)$ and $(v \mid \gamma)$ positive numbers strictly less than c, what is impossible. Therefore α is indecomposable, i.e., $\alpha \in B$.

Let β and γ be two different roots in B. Assume that $(\beta \mid \gamma) > 0$. Then, by 1.3.6, $\delta = \beta - \gamma$ is a root. Assume that δ is in R_v^+. This would imply that $\beta = \gamma + \delta$. This is impossible, since β is indecomposable. Therefore, $-\delta$ must be in R_v^+. But in this case we would have $\gamma = \beta + (-\delta)$, what contradicts the indecomposability of γ. Hence, $(\beta \mid \gamma) \leq 0$. By 3.2.2, the set B is linearly independent.

Now we prove that any root in R_v^+ is a sum of roots in B. The proof is by induction on the ordering on the finite set F. We already proved that $(v \mid \beta) = c$ implies that $\beta = \alpha$. Assume that the statement holds for all roots β in R_v^+ such that $(v \mid \beta) < d$ for some $d \in F$. Let β be a root such that $(v \mid \beta) = d$. If β is in B we are done. Otherwise, β is decomposable and by the above argument $\beta = \gamma + \delta$ with $(v \mid \gamma)$ and $(v \mid \delta)$ positive numbers strictly less than d. By the induction assumption, γ and δ are sums of roots in B. Therefore, β is a sum of roots in B.

If β is a root in $-R_v^+$, $-\beta$ is a sum of roots in B. Therefore B satisfies the property (ii). On the other hand, R spans V, so B is a basis of the vector space V. □

The consequence of this lemma is the following result.

3.2.4. Theorem. *Let R be a root system. Then there exists a basis B of R.*

3.3. Positive roots. Let B be a basis of R. We denote by R^+ the set of all roots in R which are sums of roots in B. The set R^+ is the *set of positive roots* in R attached to B. Obviously we have that $R = R^+ \cup (-R^+)$ and $R^+ \cap (-R^+) = \emptyset$.

3.3.1. LEMMA. *The basis B is completely determined by R^+.*

PROOF. Assume that we have two bases B and B' which determine the same positive root system. Let α be an element in B. Then α is a sum of roots in B'. Analogously, any $\beta \in B'$ is a sum of roots in B. If the first sum has more than one nonzero term, we would have a contradiction. Therefore, α has to be in B'. This implies $B \subset B'$. Therefore $B = B'$. □

A set of roots $P \subset R$ is *closed* if $\alpha + \beta \in P$ for any $\alpha, \beta \in P$ such that $\alpha + \beta$ is a root. If B is a basis of R and R^+ the corresponding set of positive roots, R^+ is a closed set of roots.

3.3.2. PROPOSITION. *Let P be a closed set of roots such that $P \cap (-P) = \emptyset$. Then there exists a basis B of R such that P is a subset of the set of positive roots R^+ attached to B.*

PROOF. Let $\beta = \alpha_1 + \alpha_2 + \cdots + \alpha_q$ for $\alpha_i \in P$ and $q \geq 1$. We claim that $\beta \neq 0$. The proof is by induction in q. If $q = 1$, the assertion is obvious.

Assume that $q > 1$ and $\beta = 0$. Then $\alpha_1 = -(\alpha_2 + \cdots + \alpha_q)$ and

$$0 < (\alpha_1 \mid \alpha_1) = -\sum_{i=2}^{q} (\alpha_1 | \alpha_i).$$

Therefore, there exists $2 \leq j \leq q$ such that $(\alpha_1 \mid \alpha_j) < 0$. By 1.3.6, this implies that $\alpha_1 + \alpha_j$ is a root. Since α_1 and α_j are in P and P is closed, the root $\alpha_1 + \alpha_j \in P$. Therefore, we have

$$0 = \beta = (\alpha_1 + \alpha_j) + \sum_{i=2, i \neq j} \alpha_i$$

is a sum of $q - 1$ roots in P. This contradicts the induction assumption.

We claim that there exists $\gamma \in P$ such that $(\gamma \mid \alpha) \geq 0$ for all $\alpha \in P$. Assume the opposite. Then for any $\beta \in P$ there would exist $\alpha \in P$ such that $(\beta \mid \alpha) < 0$. By 1.3.6, this would imply that $\beta + \alpha$ is a root, and $\beta + \alpha \in P$. This would imply that there exists a sequence (α_n) in P such that $\alpha_1 + \alpha_2 + \cdots + \alpha_n \in P$ for any $n \in \mathbb{N}$. Since P is a finite set, we would have $\alpha_p + \alpha_{p+1} + \cdots + \alpha_q = 0$ for some $p < q$ in $\mathbb{N}$. This would contradict the above statement.

Therefore, there exists a nonzero vector $v \in V$ such that $(\alpha \mid v) \geq 0$ for all $\alpha \in P$.

Now we can complete the proof by induction in $\operatorname{rank} R$. The statement is obvious if $\operatorname{rank} R = 1$. Assume that $\operatorname{rank} R > 1$. Let R_1 be the set of roots perpendicular to v. Denote by V_1 the linear subspace of V spanned by R_1. Clearly, R_1 is a root system in V_1 and $\operatorname{rank} R_1 < \operatorname{rank} R$. Moreover, $P \cap R_1$ is a closed subset of R_1 satisfying the assumptions. Hence, there exists a basis B_1 of $R_!$ such that the corresponding set of positive roots contains $P_1 = P \cap R_1$. This in turn implies that there exists a vector $u \in V_1$ such that $(\alpha \mid u) > 0$ for $\alpha \in B_1$. Hence, $(\alpha \mid u) > 0$ for all α in P_1. By the construction, $|(v \mid \beta)| > 0$ for all roots in $\beta \in R - R_1$. Assume that $\epsilon > 0$ is such that $|(v|\beta)| > \epsilon$ for all $\beta \in R - R_1$. By shortening u, we can assume that $|(u \mid \beta)| < \epsilon$ for all $\beta \in R - R_1$. Then

$$|(v + u \mid \beta)| \geq ||(v \mid \beta)| - |(u \mid \beta)|| > 0$$

for $\beta \in R - R_1$. On the other hand,

$$(v + u \mid \beta) = (u \mid \beta) \neq 0$$

for $\beta \in R_1$. Therefore, $(v+u \mid \beta) \neq 0$ for all $\beta \in R$. By the proof of 3.2.3, we know that there exists a basis B of R such that $R^+ = \{\alpha \in R \mid ((v+u \mid \alpha) > 0\}$.

On the other hand, for $\beta \in P_1$ we have

$$(v+u \mid \beta) = (u \mid \beta) > 0$$

and $P_1 \subset R^+$. If $\beta \in P - P_1$, we have

$$(v+u \mid \beta) = (v \mid \beta) + (u \mid \beta) > \epsilon - \epsilon = 0$$

and $\beta \in R^+$. It follows that $P \subset R^+$. □

3.3.3. Corollary. *Let P be a closed subset of the root system R such that $P \cap (-P) = \emptyset$ and $P \cup (-P) = R$. Then there exists a basis B of R such that P is equal to the set of positive roots R^+ attached to B.*

Proof. By 3.3.2 there exists a basis B such that $P \subset R^+$. On the other hand, R is a disjoint union of P and $-P$, so $\operatorname{Card}(P) = \frac{1}{2}\operatorname{Card}(R) = \operatorname{Card}(R^+)$. Hence, we must have $P = R^+$. □

3.4. Weyl groups and bases. Let $\mathcal{B}$ is the set of all bases of R. For any basis B and automorphism A of R, $A(B)$ is again a basis of R. Therefore, the group $\operatorname{Aut}(R)$ acts on the set $\mathcal{B}$. We want to study this action. As we remarked, we can assume that R is a reduced root system.

First we need some preparation.

3.4.1. Lemma. *Let R be a reduced root system. Let α be a root in B. Then s_α permutes $R^+ - \{\alpha\}$.*

Proof. Let γ be a positive root different from α. Then

$$\gamma = \sum_{\beta \in B - \{\alpha\}} m_\beta \beta + m_\alpha \alpha$$

where m_α and m_β are nonnegative integers. Moreover, since γ is not α, at least one of m_β is strictly positive.

Now, we have

$$s_\alpha(\gamma) = \sum_{\beta \in B - \{\alpha\}} m_\beta s_\alpha(\beta) - m_\alpha \alpha = \sum_{\beta \in B - \{\alpha\}} m_\beta \beta - \left(\sum_{\beta \in B - \{\alpha\}} m_\beta \alpha\check{}(\beta) + m_\alpha \right) \alpha.$$

Since at least one m_β is strictly positive, $s_\alpha(\gamma)$ has to be a positive root different from α. □

Let R be a reduced root system. Let ρ be the half-sum of roots in R^+. Then we have the following result.

3.4.2. Corollary. *For any simple root $\alpha \in B$ we have*

(i)
$$s_\alpha(\rho) = \rho - \alpha;$$

(ii) $\alpha\check{}(\rho) = 1$.

Proof. By 3.4.1, s_α permutes elements of $R^+ - \{\alpha\}$ and maps α into $-\alpha$. Hence, we have $s_\alpha(2\rho) = 2\rho - 2\alpha$ what implies the first statement. It immediately implies the second. □

Let $W(R)$ be the Weyl group of R. For a basis B of R consider the subgroup W_B of $W(R)$ generated by reflections s_α, $\alpha \in B$.

3.4.3. Lemma. *Let $v \in V$. Then there exists an element w of W_B such that $(w \cdot v \mid \alpha) \geq 0$ for any $\alpha \in B$.*

Proof. Let $w \in W_S$ be an element such that $(w \cdot v \mid \rho)$ is maximal possible. Then, for any α in B, we have

$$(w \cdot v \mid \rho) \geq ((s_\alpha w) \cdot v \mid \rho) = (w \cdot v \mid s_\alpha \rho) = (w \cdot v \mid \rho - \alpha) = (w \cdot v \mid \rho) - (w \cdot v \mid \alpha)$$

by 3.4.2. This in turn implies that $(w \cdot v \mid \alpha) \geq 0$ for all $\alpha \in B$. □

Now we can prove the following result.

3.4.4. Lemma. *The group W_B acts transitively on $\mathcal{B}$.*

Proof. Let B' be another basis of R. Then there exists a vector $v \in V$ such that $(v \mid \beta) > 0$ for all $\beta \in B'$. This implies that $(v \mid \beta) > 0$ for all positive roots with respect to B', and therefore $(v \mid \beta) \neq 0$ for all roots in R. By 3.4.3, there exists $w \in W_B$ such that $(w \cdot v \mid \alpha) \geq 0$ for all $\alpha \in B$. This in turn implies that $(v \mid w^{-1}\alpha) > 0$ for all $\alpha \in B$. Therefore, $(v \mid \alpha) > 0$ for all roots $\alpha \in w^{-1}(B)$. It follows that B' and $w^{-1}(B)$ determine the same set of positive roots in R. Therefore, by 3.3.1, we must have $B' = w(B)$. □

This result has two important consequences.

3.4.5. Theorem. *The Weyl group $W(R)$ acts transitively on $\mathcal{B}$.*

3.4.6. Theorem. *Let B be a basis of R. Then $W(R)$ is generated by s_α, $\alpha \in B$.*

Proof. Let α be a root in R. Then by 3.2.3, there exists a basis B' of R containing α. By 3.4.4, there exists $w \in W_B$ such that $w(B') = B$. In particular, $w\alpha \in B$ and $s_{w\alpha} \in W_B$. Since $s_{w\alpha} = w s_\alpha w^{-1}$, we see that $s_\alpha = w^{-1} s_{w\alpha} w \in W_B$. Therefore, for any root α, the reflection s_α is in W_B. This implies that $W_B = W(R)$. □

3.5. Length function on the Weyl group. Let R be a reduced root system and B its basis. Denote by R^+ the corresponding set of positive roots.

Let $S = \{s_\alpha; \alpha \in B\}$. By 3.4.6, the reflections in S generate the Weyl group $W(R)$ of R.

Therefore, any $w \in W(R)$ can be written as a product $w = s_{\alpha_1} s_{\alpha_2} \cdots s_{\alpha_q}$ for some elements $\alpha_1, \alpha_2, \cdots, \alpha_q$ in B. We say that the *length* $\ell(w) = \ell_S(w)$ is the smallest number $q \geq 0$ such that w is the product of q elements of S. We say that a *reduced expression* of w is any sequence $(s_1, s_2, \cdots, s_q)$ of elements of S such that $w = s_1 s_2 \cdots s_q$ and $q = \ell(w)$.

Then 1 is the unique element of $W(R)$ of length 0 and S consists of elements of length 1.

The function $\ell : W(R) \longrightarrow \mathbb{Z}_+$ is called the *length function* on $W(R)$ with respect to S.

3.5.1. Lemma. *We have*

(i) *$\ell(w) = \ell(w^{-1})$ for any $w \in W(R)$;*
(ii) *$\ell(ww') \leq \ell(w) + \ell(w')$ for any $w, w' \in W$;*
(iii) *$|\ell(w) - \ell(w')| \leq \ell(ww'^{-1})$ for any $w, w' \in W$.*

PROOF. (i) Let $w = s_1 s_2 \cdots s_q$ with $q = \ell(w)$. Then $w^{-1} = s_q s_{q-1} \cdots s_1$. Hence $\ell(w^{-1}) \leq \ell(w)$. This in turn implies that $\ell(w) = \ell((w^{-1})^{-1}) \leq \ell(w^{-1}) \leq \ell(w)$. It follows that $\ell(w) = \ell(w^{-1})$.

(ii) Let $w, w' \in W(R)$. Let $(s_1, s_2, \cdots, s_p)$ and $(s'_1, s'_2 \cdots, s'_q)$ be the reduced expressions of w and w' respectively. Then $ww' = s_1 s_2 \cdots s_p s'_1 s'_2 \cdots s'_q$ and $\ell(ww') \leq p + q = \ell(w) + \ell(q)$.

(iii) From (ii) we see that $\ell(w) \leq \ell(ww'^{-1}) + \ell(w')$. Hence we have $\ell(w) - \ell(w') \leq \ell(ww'^{-1})$. Switching w and w' we get $\ell(w') - \ell(w) \leq \ell(w'w^{-1})$. By (i), we have that $\ell(w'w^{-1}) = \ell((ww'^{-1})^{-1}) = \ell(ww'^{-1})$. This implies (iii). □

3.5.2. THEOREM. *For any $w \in W(R)$ we have*

$$\ell(w) = \operatorname{Card}(R^+ \cap (-wR^+)).$$

PROOF. We prove this statement by induction in $\ell(w)$. Clearly the statement is true for $w = 1$, i.e., for $\ell(w) = 0$.

Let $w \in W(R)$ and $\alpha \in B$. Then

$$\begin{aligned} R^+ \cap (-s_\alpha w R^+) &= s_\alpha(s_\alpha R^+ \cap (-wR^+)) \\ &= s_\alpha((R^+ - \{\alpha\}) \cap (-wR^+)) \cup s_\alpha(\{-\alpha\} \cap (-wR^+)), \end{aligned}$$

since s_α permutes the roots in $R^+ - \{\alpha\}$ by 3.4.1. There are two possibilities:

(a) $\alpha \in wR^+$. Then

$$(R^+ - \{\alpha\}) \cap (-wR^+) = R^+ \cap (-wR^+)$$

and

$$\operatorname{Card}(R^+ \cap (-s_\alpha w R^+)) = \operatorname{Card}(R^+ \cap (-wR^+)) + 1.$$

(b) $\alpha \in -wR^+$. Then

$$\operatorname{Card}(R^+ \cap (-s_\alpha w R^+)) = \operatorname{Card}(R^+ \cap (-wR^+)) - 1.$$

Now we can prove the induction step. Assume that there exists $k \in \mathbb{Z}_+$ such that the assertion holds for all $v \in W$ such that $\ell(v) \leq k$. Let $w \in W(R)$ be an element with $\ell(w) = k + 1$. Then there exists $\alpha \in B$ and $w' \in W$ such that $w = s_\alpha w'$ and $\ell(w) = \ell(w') + 1$. By the induction assumption, the assertion holds for w', i.e.,

$$\operatorname{Card}(R^+ \cap (-w'R^+)) = k.$$

Assume that

$$\operatorname{Card}(R^+ \cap (-wR^+)) \neq k + 1.$$

By the above discussion, we must have $\alpha \in -w'R^+$. Therefore, ${w'}^{-1}\alpha \notin R^+$. Assume that $w' = s_{\alpha_1} s_{\alpha_2} \ldots s_{\alpha_k}$ is a reduced expression of w'. Then we have $s_{\alpha_k} \ldots s_{\alpha_2} s_{\alpha_1} \alpha \notin R^+$. Since $\alpha \in B \subset R^+$, there exists $1 \leq j \leq k$ such that

$$s_{\alpha_j} \ldots s_{\alpha_1} \alpha \in R^+ \text{ and } s_{\alpha_{j+1}} \ldots s_{\alpha_1} \alpha \notin R^+.$$

Since reflection with respect to the simple root α_{j+1} permutes $R^+ - \{\alpha_{j+1}\}$ we see that

$$s_{\alpha_j} \ldots s_{\alpha_1} \alpha = \alpha_{j+1}.$$

Hence, we have

$$s_{\alpha_{j+1}} = s_{s_{\alpha_j} \ldots s_{\alpha_1} \alpha} = (s_{\alpha_j} \ldots s_{\alpha_1}) s_\alpha (s_{\alpha_j} \ldots s_{\alpha_1})^{-1},$$

and

$$s_{\alpha_1} \ldots s_{\alpha_j} s_{\alpha_{j+1}} = s_\alpha s_{\alpha_1} \ldots s_{\alpha_j}.$$

Therefore,

$$w = s_\alpha w' = s_\alpha s_{\alpha_1} \dots s_{\alpha_k} = s_\alpha s_{\alpha_1} \dots s_{\alpha_j} s_{\alpha_{j+1}} \dots s_{\alpha_k} = s_{\alpha_1} \dots s_{\alpha_j} s_{\alpha_{j+2}} \dots s_{\alpha_k},$$

what contradicts $\ell(w) = k+1$. Therefore, the statement must hold for w. □

The following result follows from the proof of the theorem.

3.5.3. Corollary. *Let $w \in W$ and $\alpha \in B$. Then either*
(i) $\alpha \in wR^+$ *and* $\ell(s_\alpha w) = \ell(w) + 1$, *or*
(ii) $\alpha \in -wR^+$ *and* $\ell(s_\alpha w) = \ell(w) - 1$.

Moreover, we have the following simple observation.

3.5.4. Corollary. *Let $w \in W(R)$ be such that $w(B) = B$. Then $w = 1$.*

Proof. By our assumption we have $w(R^+) = R^+$. Therefore $R \cap (-wR^+) = \emptyset$. By 3.5.2, it follows that $\ell(w) = 1$. Therefore, $w = 1$. □

This result has the following important consequence.

3.5.5. Theorem. *The Weyl group $W(R)$ acts simply transitively on $\mathcal{B}$.*

3.6. The longest element. Assume that R is a reduced root system. Let $w \in W(R)$. Then, by 3.5.2, we have

$$\ell(w) \le \operatorname{Card}(R^+) = \frac{1}{2} \operatorname{Card}(R).$$

Since $W(R)$ acts simply transitively on set $\mathcal{B}$ of all bases in R, there exists a unique element $w_0 \in W$ such that $w_0(B) = -B$. Therefore, we have $w_0(R^+) = -R^+$. By 3.5.2, we have $\ell(w_0) = \operatorname{Card}(R^+ \cap (-w_0(R^+))) = \operatorname{Card}(R^+)$. Any element w of that length has to satisfy $-w(R^+) = R^+$, i.e., $w(R^+) = -R^+$ and $w(B) = -B$. Since $W(R)$ acts simply transitively on $\mathcal{B}$, we have $w = w_0$. It follows that w_0 is the unique element of maximal length.

Moreover, $w_0^2(B) = B$. Hence, $w_0^2 = 1$, i.e., $w_0 = w_0^{-1}$.

In addition, we have

$$\ell(ww_0) = \operatorname{Card}(R^+ \cap (-ww_0(R^+))) = \operatorname{Card}(R^+ \cap (w(R^+)))$$
$$= \operatorname{Card}(R^+ - (R^+ \cap (-w(R^+)))) = \operatorname{Card}(R^+) - \operatorname{Card}(R^+ \cap (-w(R^+))) = \ell(w_0) - \ell(w)$$

for any $w \in W$.

This in turn implies that

$$\ell(w_0 w) = \ell((w_0 w)^{-1}) = \ell(w^{-1} w_0) = \ell(w_0) - \ell(w^{-1}) = \ell(w_0) - \ell(w)$$

for any $w \in W$.

4. Relations defining semisimple Lie algebra

4.1. Triangular decomposition of semisimple Lie algebras. Let $\mathfrak{g}$ be a semisimple Lie algebra over an algebraically closed field k. Let $\mathfrak{h}$ a Cartan subalgebra of $\mathfrak{g}$. Let R be the root system of $(\mathfrak{g}, \mathfrak{h})$ in $\mathfrak{h}^*$. Fix a basis B of R and denote by R^+ the corresponding set of positive roots in R.

Let

$$\mathfrak{n}_+ = \bigoplus_{\alpha \in R^+} \mathfrak{g}_\alpha \text{ and } \mathfrak{n}_- = \bigoplus_{\alpha \in R^+} \mathfrak{g}_{-\alpha}.$$

4.1.1. PROPOSITION. (i) *The subspaces* $\mathfrak{n}_+$ *and* $\mathfrak{n}_-$ *are Lie subalgebras of* $\mathfrak{g}$.

(ii) *we have a vector space decomposition*

$$\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+.$$

The direct sums $\tilde{\mathfrak{b}}_+ = \mathfrak{h} \oplus \tilde{\mathfrak{n}}_+$ *and* $\tilde{\mathfrak{b}}_- = \mathfrak{h} \oplus \tilde{\mathfrak{n}}_-$ *are Lie subalgebras of* $\tilde{\mathfrak{g}}_W$. *Moreover* $\tilde{\mathfrak{n}}_+$ *is an ideal in* $\tilde{\mathfrak{b}}_+$ *and* $\tilde{\mathfrak{n}}_-$ *is an ideal in* $\tilde{\mathfrak{b}}_-$.

PROOF. Let $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_\beta$ with $\alpha, \beta \in R^+$. Then either $[x, y] = 0$ or $[x, y] \in \mathfrak{g}_{\alpha+\beta}$ with $\alpha + \beta \in R$. Since R^+ is closed, in the latter case we have $\alpha + \beta \in R^+$ and $[x, y] \in \mathfrak{n}_+$. By linearity, we see that $\mathfrak{n}_+$ is a Lie subalgebra of $\mathfrak{g}$.

Since $-B$ is also a basis of R and $-R^+$ is the corresponding set of positive roots, the above statement also proves the assertion about $\mathfrak{n}_-$.

Finally,

$$\mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+ = \mathfrak{h} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha = \mathfrak{g}$$

by 2.1.3. □

Let α be a positive root. Then $\alpha = \sum_{\beta \in B} m_\beta \beta$ with $m_\beta \in \mathbb{N}$. We define $m(\alpha) = \sum_{\alpha \in B} m_\alpha$. Clearly m is a function from R^+ into $\mathbb{Z}_+$. If $\alpha, \beta \in R^+$ such that $\alpha + \beta \in R^+$, we have $m(\alpha + \beta) = m(\alpha) + m(\beta)$. For any $k \in \mathbb{N}$, we denote

$$\mathfrak{n}_k = \bigoplus_{m(\alpha) \geq k} \mathfrak{g}_\alpha.$$

Clearly, $(\mathfrak{n}_k, k \in \mathbb{N})$ is a decreasing filtration of $\mathfrak{n}_+$. We have $\mathfrak{n}_1 = \mathfrak{n}_+$ and $\mathfrak{n}_k = \{0\}$ for large enough k.

Let $\alpha, \beta \in R^+$ be such that $m(\beta) \geq k$. Let $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_\beta$. Then $[x, y] = 0$ or $[x, y] \in \mathfrak{g}_{\alpha+\beta}$. Since $m(\alpha + \beta) = m(\alpha) + m(\beta) \geq m(\alpha) + 1$, we see that $[x, y] \in \mathfrak{n}_{k+1}$. By linearity we conclude that $[\mathfrak{n}_+, \mathfrak{n}_k] \subset \mathfrak{n}_{k+1}$. Therefore, $\mathfrak{n}_k$ are ideals in $\mathfrak{n}_+$. We claim that $\mathcal{C}^k(\mathfrak{n}_+) \subset \mathfrak{n}_k$ for any $k \in \mathbb{N}$. Clearly, $\mathfrak{n}_1 = \mathfrak{n}_+ = \mathcal{C}^0(\mathfrak{n}_+)$. Assume that the statement holds for k. Then we have

$$\mathcal{C}^{k+1}(\mathfrak{n}_+) = [\mathfrak{n}_+, \mathcal{C}^k(\mathfrak{n}_+)] \subset [\mathfrak{n}_+, \mathfrak{n}_k] \subset \mathfrak{n}_{k+1}.$$

By induction, this proves our claim. Since $\mathfrak{n}_k = \{0\}$ for large k, we see that $\mathcal{C}^k(\mathfrak{n}_+) = \{0\}$ for large k, i.e., $\mathfrak{n}_+$ is a nilpotent Lie algebra.

By switching B with $-B$, we also see that $\mathfrak{n}_-$ is a nilpotent Lie algebra. Therefore we have the following result.

4.1.2. PROPOSITION. *The Lie subalgebras* $\mathfrak{n}_+$ *and* $\mathfrak{n}_-$ *are nilpotent.*

4.1.3. LEMMA. *Let* $\alpha_1, \alpha_2, \dots, \alpha_n$ *be positive roots such that* $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_n$ *is a positive root. Then there exists a permutation* π *of* $\{1, 2, \dots, n\}$ *such that* $\alpha_{\pi(1)} + \alpha_{p(2)} + \cdots + \alpha_{\pi(k)}$ *is a positive root for any* $1 \leq k \leq n$.

PROOF. As before, we can assume that R is a real root system in V and V is equipped with an invariant inner product.

The statement is obvious for $n = 1$. Assume that the statement holds for all $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_k$ for $k < n$. Let $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_n$. Assume that $(\alpha \mid \alpha_i) \leq 0$ for all $1 \leq j \leq n$. Then

$$0 < (\alpha \mid \alpha) = \sum_{i=1}^{n} (\alpha \mid \alpha_j) \leq 0,$$

what is impossible. Hence, there exists $1 \leq j \leq n$ such that $(\alpha \mid \alpha_j) < 0$. Hence, by 1.3.6, we know that $\alpha - \alpha_j$ is a root. Hence $\alpha = \beta + \alpha_j$ and $\beta = \alpha_1 + \cdots + \alpha_{j-1} + \alpha_{j+1} + \cdots + \alpha_n$. Since the induction assumption applies to β, the assertion follows. □

4.1.4. THEOREM. (i) *Let X_α be nonzero vectors in $\mathfrak{g}_\alpha$ for $\alpha \in B$. Then $(X_\alpha; \alpha \in B)$ generate the Lie algebra $\mathfrak{n}_+$.*

(ii) *Let Y_α be nonzero vectors in $\mathfrak{g}_{-\alpha}$ for $\alpha \in B$. Then $(Y_\alpha; \alpha \in B)$ generate the Lie algebra $\mathfrak{n}_-$.*

(iii) *The vectors $(X_\alpha, Y_\alpha, H_\alpha; \alpha \in B)$ generate the Lie algebra $\mathfrak{g}$.*

PROOF. Let α be a root in R^+. Assume that $m(\alpha) = k$. Then α is a sum of k roots in B, i.e., $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_k$. By 4.1.3, we can pick these simple roots in such way that $\alpha_1 + \alpha_2 + \cdots + \alpha_j$ is a root in R^+ for any $1 \leq j \leq k$. In particular, $\beta = \alpha_1 + \alpha_2 + \cdots + \alpha_{k-1}$ is a root in R^+ and $\alpha = \beta + \alpha_k$. Therefore, by **??** we have $[\mathfrak{g}_\beta, \mathfrak{g}_{\alpha_k}] = \mathfrak{g}_\alpha$.

Consider now the Lie subalgebra $\mathfrak{u}$ of $\mathfrak{n}_+$ generated by X_α, $\alpha \in B$. By the definition, it contains all root subspaces $\mathfrak{g}_\beta$ with $m(\beta) = 1$. Assume that $\mathfrak{u}$ contains all root subspaces $\mathfrak{g}_\beta$ with $m(\beta) < k$. Let $\alpha \in R^+$ be such that $m(\alpha) = k$. Then by the above argument, $\alpha = \beta + \alpha_k$ with $m(\beta) < k$ and $\mathfrak{g}_\alpha = [\mathfrak{g}_\beta, \mathfrak{g}_{\alpha_k}] \subset \mathfrak{u}$. Hence, $\mathfrak{u}$ contains all root subspaces $\mathfrak{g}_\beta$ with $m(\beta) \leq k$. By induction it follows that $\mathfrak{u} = \mathfrak{n}_+$.

The proof for $\mathfrak{n}_-$ follows by replacing the basis B with $-B$. Finally, (iii) follows from 2.1.3. □

4.2. Serre identities. The following remarkable identities were observed by Serre.

4.2.1. LEMMA. *Let α and β be two different roots in B.*

(i) *Let $X_\alpha \in \mathfrak{g}_\alpha$, $X_\beta \in \mathfrak{g}_\beta$. Then we have*

$$(\operatorname{ad} X_\alpha)^{-n(\beta,\alpha)+1}(X_\beta) = 0.$$

(ii) *Let $X_{-\alpha} \in \mathfrak{g}_{-\alpha}$, $X_{-\beta} \in \mathfrak{g}_{-\beta}$. Then we have*

$$(\operatorname{ad} X_{-\alpha})^{-n(\beta,\alpha)+1}(X_{-\beta}) = 0.$$

PROOF. Since α and β are in B, $\alpha - \beta$ is not a root and we must have $n(\alpha, \beta) \leq 0$ by 1.3.5. Therefore, the exponent in above relation is always positive. Since

$$n(-\beta, -\alpha) = (-\alpha)^*(-\beta) = \alpha^*(\beta) = n(\beta, \alpha)$$

for any two roots α and β, the statement (ii) follows from (i) by switching the basis B with basis $-B$.

Consider the α-string S defined by β. Since $\beta - \alpha$ is not a root, this string starts with β. Assume that $\beta + p\alpha$ is its end. Then by 1.3.7, we have $p = -n(\beta, \alpha)$. Therefore, the vector $(\operatorname{ad} X_\alpha)^{-n(\beta,\alpha)}(X_\beta) = (\operatorname{ad} X_\alpha)^p(X_\beta) \in \mathfrak{g}_{\beta+p\alpha}$ and

$$(\operatorname{ad} X_\alpha)^{-n(\beta,\alpha)+1}(X_\beta) \subset (\operatorname{ad} X_\alpha)(\mathfrak{g}_{\beta+p\alpha}) = \{0\}$$

since $\beta + (p+1)\alpha$ is not a root. □

4.3. Lie algebra defined by Weyl relations. Let R be a reduced root system in a vector space V defined over an algebraically closed field k. Let r be the rank of R. Let $B = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$ be the basis of R. We denote by $n(i,j) = n(\alpha_i, \alpha_j)$, $1 \le i, j \le r$, the Cartan matrix of R with respect to the basis B.

Let $\tilde{\mathfrak{g}}$ be the free Lie algebra over k generated by elements H_i, X_i, Y_i, $1 \le i \le r$. Clearly, there exists an endomorphism Ω of $\tilde{\mathfrak{g}}$ such that $\Omega(H_i) = -H_i$, $\Omega(X_i) = Y_i$ and $\Omega(Y_i) = X_i$ for all $1 \le i \le r$. By definition, we have $\Omega^2 = 1$. Therefore, Ω is an automorphism.

Define the ideal $\mathfrak{i}$ in $\tilde{\mathfrak{g}}$ generated by relations:

(W1) $[H_i, H_j]$ for all $1 \le i, j \le r$;
(W2) $[H_i, X_j] - n(j,i)X_j$ for all $1 \le i, j \le r$;
(W3) $[H_i, Y_j] + n(j,i)Y_j$ for all $1 \le i, j \le r$;
(W4) $[X_i, Y_j] - \delta_{i,j} H_i$ for all $1 \le i, j \le r$.

We call there expressions on the right side of the above equations the *Weyl relations*.

In this section we want to describe the structure of the Lie algebra $\tilde{\mathfrak{g}}_W$ which is the quotient of the free Lie algebra $\tilde{\mathfrak{g}}$ by the ideal $\mathfrak{i}$ generated by Weyl relations.

First we remark that the Weyl relations are invariant under the action of Ω. Therefore, Ω induces an automorphism of ω of the Lie algebra $\tilde{\mathfrak{g}}_W$. Moreover, we have $\omega^2 = 1$.

Let U be an r-dimensional complex vector space and $T(U)$ its tensor algebra. Let $e_1, e_2, \dots, e_r$ be a basis of U. Then a basis of $T(U)$ is given by $e_I = e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_m}$ for any multiindex $I = (i_1, i_2, \dots, i_m)$ with $1 \le i_k \le r$, $1 \le k \le m$, for any $m \in \mathbb{Z}_+$. To define a representation of the free Lie algebra $\tilde{\mathfrak{g}}$ on $T(U)$ one arbitrarily assigns the action of the generators $(H_i, X_i, Y_i; 1 \le i \le r)$ on $T(U)$.

For a multiindex $I = (i_1, i_2, \dots, i_m)$, we put

$$n(I,j) = n(i_1,j) + n(i_2,j) + \cdots + n(i_m,j)$$

for all $1 \le j \le r$.

First we define the action of H_i by

$$H_i \cdot 1 = 0$$

and

$$H_i \cdot e_I = n(I,i) e_I$$

for any multiindex I. Then we define the action of X_i by

$$X_i \cdot 1 = e_i$$

and

$$X_i \cdot e_I = e_i \otimes e_I$$

for all I.

Finally we define the action of Y_i inductively by

$$Y_i \cdot 1 = 0.$$

If $I = (i_1, i_2, \dots, i_m)$ is a nontrivial multiindex, we put $J = (i_2, \dots, i_m)$, and define

$$Y_i \cdot e_I = e_{i_1} \otimes Y_i \cdot e_J - \delta_{i,i_1} n(J,i) e_J$$

for all such I.

4.3.1. Lemma. *The ideal* $\mathfrak{i}$ *acts trivially on* $T(U)$.

PROOF. Since the action of H_i, $1 \le i \le r$, is given by diagonal matrices, they obviously commute. Hence they satisfy the relation (W1).

Now we consider (W2). First, we have

$$[H_i, X_j] \cdot 1 = H_i X_j \cdot 1 - X_j H_i \cdot 1 = H_i \cdot e_j = n(j,i) e_j = n(j,i) X_j \cdot 1.$$

also, we have

$$\begin{aligned}[H_i, X_j] \cdot e_I &= H_i X_j e_I - X_j H_i e_I = H_i(e_j \otimes e_I) - n(I,i) X_j e_I \\ &= (n(j,i) + n(I,i))(e_j \otimes e_I) - n(I,i)(e_j \otimes e_I) = n(j,i) X_j \cdot e_I.\end{aligned}$$

Therefore, (W2) holds.

Now we consider (W4). We have

$$[X_i, Y_j] \cdot 1 = X_i Y_j \cdot 1 - Y_j X_i \cdot 1 = -Y_j \cdot e_i = -e_i \otimes Y_i \cdot 1 = 0.$$

Moreover, we have

$$\begin{aligned}[X_i, Y_j] \cdot e_I = X_i Y_j \cdot e_I - Y_j X_i \cdot e_I = e_i \otimes Y_j \cdot e_I - Y_j(e_i \otimes e_I) \\ = \delta_{i,j} n(I,i) e_I = \delta_{i,j} H_i \cdot e_I.\end{aligned}$$

This proves (W4).

To prove (W3) we first establish the identity

$$H_i Y_j \cdot e_I = (n(I,i) - n(j,i)) Y_j \cdot e_I$$

for all I.

First we observe that

$$H_i Y_j \cdot 1 = 0 = -n(j,i) Y_j \cdot 1.$$

Assume that the identity holds for all basis elements with $m < k$. By definition, for $I' = (i_1, i_2, \dots, i_k)$ and $J' = (i_2, \dots, i_k)$, we have

$$\begin{aligned}H_i Y_j \cdot e_{I'} &= H_i \cdot (e_{i_1} \otimes Y_j \cdot e_{J'}) - \delta_{i_1,j} n(I',j) H_i \cdot e_{J'} \\ &= H_i \cdot (e_{i_1} \otimes Y_j \cdot e_{J'}) - \delta_{i_1,j} n(J',j) n(J',i) \cdot e_{J'}.\end{aligned}$$

Moreover, by (W2) and the induction assumption, we have

$$\begin{aligned}H_i \cdot (e_{i_1} \otimes Y_j \cdot e_{J'}) &= H_i X_{i_1} Y_j \cdot e_{J'} = [H_i X_{i_1}] Y_j \cdot e_{J'} + X_{i_1} H_i Y_j \cdot e_{J'} \\ &= n(i_1, i) X_{i_1} Y_j \cdot e_{J'} + e_{i_1} \otimes H_i Y_j \cdot e_{J'} \\ &= n(i_1, i)(e_{i_1} \otimes Y_j \cdot e_{J'}) + (n(J',i) - n(j,i))(e_{i_1} \otimes Y_j \cdot e_{J'}) \\ &= (n(I',i) - n(j,i))(e_{i_1} \otimes Y_j \cdot e_{J'}).\end{aligned}$$

Substituting this into the above relation we get

$$\begin{aligned}&H_i Y_j \cdot e_{I'} = (n(I',i) - n(j,i))(e_{i_1} \otimes Y_j \cdot e_{J'}) - \delta_{i_1,j} n(J',j) n(J',i) e_{J'} \\ &= (n(I',i) - n(j,i))(e_{i_1} \otimes Y_j \cdot e_{J'} - \delta_{i_1,j} n(J',j) e_{J'}) = (n(I',i) - n(j,i)) Y_j \cdot e_{I'}.\end{aligned}$$

This completes the proof of the above identity.

Therefore, it follows that

$$\begin{aligned}[H_i, Y_j] \cdot e_I &= H_i Y_j \cdot e_I - Y_j H_i \cdot e_I \\ &= (n(I,i) - n(j,i)) Y_j \cdot e_I - n(I,i) Y_j \cdot e_I = -n(j,i) Y_j \cdot e_I.\end{aligned}$$

This completes the proof of (W3). □

Therefore, the action of $\tilde{\mathfrak{g}}$ on $T(U)$ factors through $\tilde{\mathfrak{g}}_W$, i.e., the action on $T(U)$ satisfies Weyl relations (W1) to (W4). Consider the vector subspace $\mathfrak{h}$ spanned by H_i, $1 \leq i \leq r$, in $\tilde{\mathfrak{g}}_W$. By (W1), this subspace is an abelian Lie subalgebra. Let $(\alpha_1^*, \alpha_2^*, \dots, \alpha_r^*)$ be the dual roots in V^* corresponding to $(\alpha_1, \alpha_2, \dots, \alpha_r)$. Then they form a basis of V^*, and we define a linear map $f \longmapsto H$ of V^* into $\tilde{\mathfrak{g}}_W$ by mapping $f = \sum_{i=1}^r c_i \alpha_i^*$ into $H = \sum_{i=1}^r c_i H_i$ for any $c_i \in k$. Then the action of H on $T(U)$ satisfies

$$H \cdot e_j = \sum_{i=1}^r c_i H_i \cdot e_j = \left(\sum_{i=1}^r c_i n(j,i) \right) e_j$$

for all $1 \leq j \leq r$. If $H = 0$, we must have

$$0 = \sum_{i=1}^r c_i n(j,i) = \sum_{i=1}^r c_i \alpha_i^*(\alpha_j) = \left(\sum_{i=1}^r c_i \alpha_i^* \right)(\alpha_j)$$

for all $1 \leq j \leq r$. Since $(\alpha_1, \alpha_2, \dots, \alpha_r)$ is a basis of V, it follows that $\sum_{i=1}^r c_i \alpha_i^* = 0$. This in turn implies that $c_i = 0$ for all $1 \leq i \leq r$, i.e., $f = 0$. Therefore, the linear map $f \longmapsto H$ is injective. Hence, we can identify V^* with its image $\mathfrak{h}$ in $\tilde{\mathfrak{g}}_W$. Moreover, we can view $(H_i, 1 \leq i \leq r)$ as a basis of $\mathfrak{h}$. Also, we can view α_i as linear forms on $\mathfrak{h}$ which take value $n(i,j)$ on H_j, $1 \leq j \leq r$. This in turn implies that for any multiindex $I = (i_1, i_2, \dots, i_m)$ we can define linear forms $\alpha_I = \alpha_{i_1} + \alpha_{i_2} + \dots + \alpha_{i_m}$ which satisfy $\alpha_I(H_j) = n(I,j)$ for all $1 \leq j \leq r$.

By (W2), (W3) and (W4), for any $1 \leq i \leq r$, the vectors H_i, X_i, Y_i in $\tilde{\mathfrak{g}}_W$ satisfy the relations

$$[H_i, X_i] = 2X_i, \quad [H_i, Y_i] = -2Y_i, \quad [X_i, Y_i] = H_i.$$

Therefore, the span of H_i, X_i and Y_i is a Lie subalgebra of $\tilde{\mathfrak{g}}_W$. Moreover, we can define a morphism φ of the simple Lie algebra $\mathfrak{sl}(2,k)$ onto this Lie subalgebra by $\varphi(H) = H_i$, $\varphi(E) = X_i$ and $\varphi(F) = Y_i$. This morphism is not zero, since $H_i \neq 0$. Since $\mathfrak{sl}(2,k)$ is simple, φ is an isomorphism. Therefore, $X_i \neq 0$ and $Y_i \neq 0$ in $\tilde{\mathfrak{g}}_W$ for any $1 \leq i \leq r$.

Consider the adjoint action of $\mathfrak{h}$ on $\tilde{\mathfrak{g}}_W$. Let $\lambda \in \mathfrak{h}^*$. We define the *weight subspace*

$$\tilde{\mathfrak{g}}_{W,\lambda} = \{x \in \tilde{\mathfrak{g}}_W \mid (\operatorname{ad} H)x = \lambda(H)x \text{ for all } H \in \mathfrak{h}\}.$$

The vector $\lambda \in \mathfrak{h}^*$ is a *weight* of $\tilde{\mathfrak{g}}_W$ if $\tilde{\mathfrak{g}}_W \neq \{0\}$. Since $\operatorname{ad} H$ is a derivation, we conclude that for any two $\lambda, \mu \in \mathfrak{h}^*$, we have

$$[\tilde{\mathfrak{g}}_{W,\lambda}, \tilde{\mathfrak{g}}_{W,\mu}] \subset \tilde{\mathfrak{g}}_{W,\lambda+\mu}.$$

If S is a set of weights of the adjoint representation is *closed*, i.e., such that for any $\alpha, \beta \in S$ such that $\alpha + \beta$ is a weight we have $\alpha + \beta \in S$, we see that

$$\tilde{\mathfrak{g}}_{W,S} = \bigoplus_{\lambda \in S} \tilde{\mathfrak{g}}_{W,\lambda}$$

is a Lie subalgebra of $\tilde{\mathfrak{g}}_W$.

Since $\mathfrak{h} \subset \tilde{\mathfrak{g}}_{W,0}$, $X_i \in \tilde{\mathfrak{g}}_{W,\alpha_i}$ and $Y_i \in \tilde{\mathfrak{g}}_{W,-\alpha_i}$, for all $1 \leq i \leq r$, we see that the generators of $\tilde{\mathfrak{g}}_W$ are in weight subspaces of $\tilde{\mathfrak{g}}_W$. Therefore, from the above discussion we have the following result.

4.3.2. Lemma. *If P is the set of all weights of $\tilde{\mathfrak{g}}_W$, we have*

$$\tilde{\mathfrak{g}}_W = \bigoplus_{\lambda \in P} \tilde{\mathfrak{g}}_{W,\lambda}.$$

Clearly, we have

$$\omega(\tilde{\mathfrak{g}}_{W,\lambda}) = \tilde{\mathfrak{g}}_{W,-\lambda}$$

for any weight λ.

Let P_+ be the set of nonzero weights which are sums of elements of B, and $P_- = -P_+$. Then P_+ and P_- are closed. Therefore, the sums of corresponding weight subspaces are Lie subalgebras of $\tilde{\mathfrak{g}}_W$.

Denote by $\tilde{\mathfrak{n}}_+$ the Lie subalgebra of $\tilde{\mathfrak{g}}_W$ generated by X_i, $1 \le i \le r$. Since X_i are in weight subspaces corresponding to B, we see that $\tilde{\mathfrak{n}}_+$ is contained in the Lie subalgebra $\tilde{\mathfrak{g}}_{W,P_+}$.

Also, denote by $\tilde{\mathfrak{n}}_-$ the Lie subalgebra of $\tilde{\mathfrak{g}}_W$ generated by Y_i, $1 \le i \le r$. Since Y_i are in weight subspaces corresponding to $-B$, we see that $\tilde{\mathfrak{n}}_-$ is contained in the Lie subalgebra $\tilde{\mathfrak{g}}_{W,P_-}$.

Since $\omega(X_i) = Y_i$ for $1 \le i \le r$, we have

$$\omega(\tilde{\mathfrak{n}}_+) = \tilde{\mathfrak{n}}_-.$$

Finally we have the following triangular decomposition of $\tilde{\mathfrak{g}}_W$.

4.3.3. Theorem. *We have*

$$\tilde{\mathfrak{g}}_W = \tilde{\mathfrak{n}}_- \oplus \mathfrak{h} \oplus \tilde{\mathfrak{n}}_+.$$

The direct sums $\tilde{\mathfrak{b}}_+ = \mathfrak{h} \oplus \tilde{\mathfrak{n}}_+$ and $\tilde{\mathfrak{b}}_- = \mathfrak{h} \oplus \tilde{\mathfrak{n}}_-$ are Lie subalgebras of $\tilde{\mathfrak{g}}_W$. The Lie subalgebra $\tilde{\mathfrak{n}}_+$ is an ideal in $\tilde{\mathfrak{b}}_+$. The Lie subalgebra $\tilde{\mathfrak{n}}_-$ is an ideal in $\tilde{\mathfrak{b}}_-$.

Proof. From the above discussion, we know that the sum $\mathfrak{k} = \tilde{\mathfrak{n}}_- \oplus \mathfrak{h} \oplus \tilde{\mathfrak{n}}_+$ on the right side is direct. Clearly, it contains all generators of $\tilde{\mathfrak{g}}_W$. Hence, we just need to show that it is a Lie subalgebra of $\tilde{\mathfrak{g}}_W$.

To show that $\mathfrak{k}$ is a Lie subalgebra, it is enough to show that $\mathfrak{k}$ is invariant for the adjoint action of all generators of $\tilde{\mathfrak{g}}_W$. Since $\mathfrak{k}$ is invariant under ω, and ω maps X_i into Y_i for $1 \le i \le r$, it is enough to show that $\mathfrak{k}$ is invariant for $\operatorname{ad} H_i$ and $\operatorname{ad} Y_i$, $1 \le i \le r$.

Let $I = (i_1, i_2, \dots, i_m)$ be a finite sequence of $1 \le i_k \le r$. We put $X_I = [X_{i_1}, [X_{i_2}, [\cdots, [X_{i_{m-1}}, X_{i_m}] \cdots]]]$. The elements X_I span a vector subspace of $\tilde{\mathfrak{n}}_+$ which is invariant for all $\operatorname{ad} X_i$, $1 \le i \le r$. Therefore, it must be equal to $\tilde{\mathfrak{n}}_+$, i.e., $\tilde{\mathfrak{n}}_+$ is spanned by all X_I.

We claim that

$$\operatorname{ad} H_i(X_I) = \alpha_I(H_i) X_I$$

for all I. The proof is by induction in m. For $m = 1$ this is the relation (W2). Assume that the statement holds for $m < k$. Then for $I = (i_1, i_2, \dots, i_k)$ we have $X_I = [X_{i_1}, X_J]$ with $J = (i_2, \dots, i_m)$. Then we have

$$\begin{aligned} \operatorname{ad} H_i(X_I) = \operatorname{ad} H_i([X_{i_1}, X_J]) &= [\operatorname{ad} H_i(X_{i_1}), X_J] + [X_{i_1}, \operatorname{ad} H_i(X_J)] \\ &= (\alpha_{i_1}(H_i) + \alpha_J(H_i))[X_{i_1}, X_J] = \alpha_I(H_i) X_I \end{aligned}$$

by the induction assumption. This establishes the above identity.

This identity implies that $\operatorname{ad} H_i(\tilde{\mathfrak{n}}_+) \subset \tilde{\mathfrak{n}}_+$ for $1 \le i \le r$. By applying ω we see that $\operatorname{ad} H_i(\tilde{\mathfrak{n}}_-) \subset \tilde{\mathfrak{n}}_-$ for $1 \le i \le r$. Therefore, we see that $\mathfrak{k}$ is invariant for $\operatorname{ad} H_i$, $1 \le i \le r$.

This discussion also implies that $[\mathfrak{h}, \tilde{\mathfrak{n}}_+] \subset \tilde{\mathfrak{n}}_+$. Therefore, $\tilde{\mathfrak{b}}_+$ is a Lie subalgebra of $\tilde{\mathfrak{g}}_W$ and $\tilde{\mathfrak{n}}_+$ is its ideal. The statements about $\tilde{\mathfrak{b}}_-$ follow by applying ω.

Clearly, we have $\operatorname{ad} Y_i(\tilde{\mathfrak{n}}_-) \subset \tilde{\mathfrak{n}}_-$ for all $1 \leq i \leq r$. Moreover,

$$\operatorname{ad} Y_i(H_j) = -\operatorname{ad} H_j(Y_i) = n(i,j)Y_i$$

by (W3) and $\operatorname{ad} Y_i(\mathfrak{h}) \subset \mathfrak{k}$ for $1 \leq i \leq r$.

Now we want to prove that $\operatorname{ad} Y_i(\tilde{\mathfrak{n}}_+) \subset \mathfrak{k}$ for $1 \leq i \leq r$. It is enough to show that $\operatorname{ad} Y_i(X_I) \in \tilde{\mathfrak{b}}_+$ for all I. The proof is by induction in m. For $m = 1$ this follows from the relation (W4). In general, we have

$$\begin{aligned}\operatorname{ad} Y_i(X_I) &= \operatorname{ad} Y_i([X_{i_1}, X_J]) = [\operatorname{ad} Y_i(X_{i_1}), X_J] + [X_{i_1}, \operatorname{ad} Y_i(X_J)] \\ &= -\delta_{i,i_1}[H_i, X_J] + [X_{i_1}, \operatorname{ad} Y_i(X_J)] = -\delta_{i,i_1}\alpha_J(H_i)X_J + [X_{i_1}, \operatorname{ad} Y_i(X_J)].\end{aligned}$$

The first term is obviously in $\tilde{\mathfrak{n}}_+$ and the second is in the Lie subalgebra $\tilde{\mathfrak{b}}_+$ by the induction assumption.

Therefore, we see that $\operatorname{ad} Y_i(\mathfrak{k}) \subset \mathfrak{k}$ for all $1 \leq i \leq r$. Hence $\tilde{\mathfrak{g}}_W = \mathfrak{k}$. □

4.3.4. Corollary. (i) *The nonzero weights of* $\tilde{\mathfrak{g}}_W$ *are in* $P_+ \cup P_-$.
(ii) *The weight subspaces of* $\tilde{\mathfrak{g}}_W$ *are finite dimensional.*
(iii) *For any* $1 \leq i \leq r$ *and* $m \in \mathbb{Z}$, $\tilde{\mathfrak{g}}_{W,m\alpha_i} = \{0\}$ *for* $m \neq -1, 0, 1$.
(iv) *Moreover,*

$$\dim \tilde{\mathfrak{g}}_{W,\alpha_i} = \dim \tilde{\mathfrak{g}}_{W,-\alpha_i} = 1.$$

Proof. The statement (i) follows immediately from 4.3.3. Moreover, if λ is in S_+, the corresponding weight subspace is in $\tilde{\mathfrak{n}}_+$. By the calculation in the proof of the theorem, it is spanned by the vectors X_I such that $\alpha_I = \lambda$. This condition determines I up to the ordering of indices. Therefore, we see that the number of possible I is finite. This in turn implies that the dimension of $\tilde{\mathfrak{g}}_{W,\lambda}$ is finite.

If $\lambda = m\alpha_i$ with $m \in \mathbb{N}$, the multiindex I in the above argument must be a sequence of m copies of i. Clearly, the corresponding X_I are 0 except in the case $m = 1$. The statement for negative m follows by application of the automorphism ω. □

4.4. Lie algebra defined by adding Serre relations. Consider the elements

$$\Theta^+_{i,j} = (\operatorname{ad} X_i)^{-n(j,i)+1}(X_j)$$

and

$$\Theta^-_{i,j} = (\operatorname{ad} Y_i)^{-n(j,i)+1}(Y_j)$$

in $\tilde{\mathfrak{g}}_W$ for $1 \leq i, j \leq r$ and $i \neq j$. Clearly, $\omega(\Theta^+_{i,j}) = \Theta^-_{i,j}$.

4.4.1. Lemma. *Let* V *be a vector space and* X, Y *and* H *three linear maps on* V *satisfying the relations* $[H, X] = 2X$, $[H, Y] = -2Y$ *and* $[X, Y] = H$. *Then, for any* $p \in \mathbb{N}$, *we have*

$$YX^p = X^pY - pX^{p-1}(H + (p-1)I).$$

PROOF. The formula is correct for $p = 1$. Moreover, if the statement holds for p, we have

$$\begin{aligned} YX^{p+1} &= YX^pX = X^pYX - pX^{p-1}(HX + (p-1)X) \\ &= X^p[Y,X] + X^{p+1}Y - pX^{p-1}([H,X] + XH + (p-1)X) \\ &= X^{p+1}Y - X^pH - pX^{p-1}(2X + XH + (p-1)X) \\ &= X^{p+1}Y - X^pH - pX^p(H + (p+1)I) = X^{p+1}Y - (p+1)X^p(H + pI). \end{aligned}$$

Hence the statement follows by induction. □

4.4.2. LEMMA. *For any $1 \le i, j \le r$, $i \ne j$, we have*

$$\operatorname{ad} Y_k(\Theta^+_{i,j}) = 0$$

for all $1 \le k \le r$.

PROOF. Assume first that $k \ne i$. Then we have $[X_i, Y_k] = 0$ by (W4) and $\operatorname{ad} Y_k$ and $\operatorname{ad} X_i$ commute. Therefore, by using again (W4), it follows that

$$\begin{aligned} \operatorname{ad} Y_k(\Theta^+_{i,j}) &= \operatorname{ad} Y_k(\operatorname{ad} X_i)^{-n(j,i)+1}(X_j) \\ &= (\operatorname{ad} X_i)^{-n(j,i)+1}([Y_k, X_j]) = -\delta_{j,k}(\operatorname{ad} X_i)^{-n(j,i)+1}(H_j). \end{aligned}$$

If $j \ne k$, the statement follows.

If $j = k$ we get

$$\operatorname{ad} Y_j(\Theta^+_{i,j}) = -(\operatorname{ad} X_i)^{-n(j,i)+1}(H_j).$$

Since we have

$$[H_j, X_i] = \alpha_i(H_j)X_i = n(i,j)X_i$$

by (W2), if roots α_i and α_j are orthogonal, $n(i,j) = 0$ and H_j and X_j commute, i.e., $\operatorname{ad} X_i(H_j) = 0$. This implies our statement in this case. If roots are not orthogonal, $n(i,j) < 0$. Therefore, by the table ??, we have $n(j,i) \le -1$ and $-n(j,i) + 1 \ge 2$. On the other hand, we have

$$(\operatorname{ad} X_i)^2(H_j) = [X_i, [X_i, H_j]] = -\alpha_i(H_j)[X_i, X_i] = 0.$$

This completes the argument in the case $k \ne i$.

It remains to consider the case $k = i$. In this case, by 4.4.1, we have

$$\begin{aligned} &\operatorname{ad} Y_i(\Theta^+_{i,j}) = (\operatorname{ad} Y_i)(\operatorname{ad} X_i)^{-n(j,i)+1}X_j \\ &= (\operatorname{ad} X_i)^{-n(j,i)+1}[Y_i, X_j] + (n(j,i)-1)(\operatorname{ad} X_i)^{-n(j,i)}([H_i, X_j] - n(j,i)X_j) = 0 \end{aligned}$$

using Weyl relations (W2) and (W4). □

Let $\tilde{\mathfrak{a}}_+$ be the ideal in $\tilde{\mathfrak{n}}_+$ generated by elements $\Theta^+_{i,j}$ for $1 \le i, j \le r$ and $i \ne j$. Let $\tilde{\mathfrak{a}}_-$ be the ideal in $\tilde{\mathfrak{n}}_-$ generated by elements $\Theta^-_{i,j}$ for $1 \le i, j \le r$ and $i \ne j$. Then ω maps $\tilde{\mathfrak{a}}_+$ onto $\tilde{\mathfrak{a}}_-$.

4.4.3. LEMMA. (i) *The vector subspace $\tilde{\mathfrak{a}}_+$ is an ideal in $\tilde{\mathfrak{g}}_W$.*
(ii) *The vector subspace $\tilde{\mathfrak{a}}_-$ is an ideal in $\tilde{\mathfrak{g}}_W$.*

PROOF. The second statement follows from the first by applying the automorphism ω.

To prove (i) we use the notation from the proof of 4.3.3. We denote by $\mathcal{J}$ the set of all multiindices ending in $-n(j,i)+1$ copies of i followed by j, for $1 \le i, j \le r$

with $i \neq j$. Clearly, all elements X_I attached to the multiindices in $\mathcal{J}$ are in $\tilde{\mathfrak{a}}_+$. Moreover, their linear span is invariant under the action of $\operatorname{ad} X_i$. Therefore, it is equal to $\mathfrak{a}_+$. From the calculation in the proof of 4.3.3, we know that all of these vectors are in weight subspaces of $\tilde{\mathfrak{g}}_W$, i.e., they are eigenvectors of $\operatorname{ad} H_i$ for all $1 \leq i \leq r$. Therefore, $\operatorname{ad} H_i(\tilde{\mathfrak{a}}_+) \subset \tilde{\mathfrak{a}}_+$ for all $1 \leq i \leq r$.

Consider now the action of $\operatorname{ad} Y_i$. By 4.4.2, this element annihilates all Θ^+_{jk}, $1 \leq j, k \leq r$ and $j \neq k$. Let I be a multiindex in $\mathcal{J}$. Then we have two options, either I corresponds to $\Theta^+_{j,k}$ for some $1 \leq j, k \leq r$, $j \neq k$, or I is a multiindex corresponding to

$$X_I = (\operatorname{ad} X_{i_1} \operatorname{ad} X_{i_2} \cdots \operatorname{ad} X_{i_m})(\Theta^+_{j,k}) = \operatorname{ad} X_{i_1}(X_J)$$

where J is the multiindex obtained from I by removing the first index i_1.

In the first case, we already know that $\operatorname{ad} Y_i(X_I) = 0 \in \tilde{\mathfrak{a}}_+$. In the second case, we have, as in the proof of 4.3.3, that

$$\operatorname{ad} Y_i(X_I) = -\delta_{i,i_1}\alpha_J(H_i)X_J + \operatorname{ad} X_{i_1}(\operatorname{ad} Y_i)(X_J).$$

Since J is in $\mathcal{J}$, but shorter than I, we conclude that $\operatorname{ad} Y_i(X_I) \in \tilde{\mathfrak{a}}_+$ by induction in m. Therefore, we have $\operatorname{ad} Y_i(\tilde{\mathfrak{a}}_+) \subset \tilde{\mathfrak{a}}_+$ for all $1 \leq i \leq k$.

Since $\tilde{\mathfrak{a}}_+$ is invariant for all $\operatorname{ad} X_i$, $\operatorname{ad} Y_i$ and $\operatorname{ad} H_i$, it is an ideal in $\tilde{\mathfrak{g}}_W$. □

Therefore, $\tilde{\mathfrak{a}}_+ \oplus \tilde{\mathfrak{a}}_-$ is an ideal in $\tilde{\mathfrak{g}}_W$. Let $\tilde{\mathfrak{g}}_S$ be the quotient of $\tilde{\mathfrak{g}}_W$ with respect to the ideal $\tilde{\mathfrak{a}}_+ \oplus \tilde{\mathfrak{a}}_-$. Then we have

$$\tilde{\mathfrak{g}}_S = \tilde{\mathfrak{n}}_+/\tilde{\mathfrak{a}}_+ \oplus \mathfrak{h} \oplus \tilde{\mathfrak{n}}_-/\tilde{\mathfrak{a}}_-.$$

Since Serre relations are mapped to each other under the action of ω, the automorphism ω maps $\tilde{\mathfrak{a}}_+$ onto $\tilde{\mathfrak{a}}_-$ and leaves $\tilde{\mathfrak{a}}_+ \oplus \tilde{\mathfrak{a}}_-$ invariant. Therefore, it induces an automorphism ϖ of $\tilde{\mathfrak{g}}_S$.

A linear map T on a vector space V is *locally nilpotent* if $T^n v = 0$ for some $n \in \mathbb{N}$ for every $v \in V$.

4.4.4. Lemma. (i) *The linear maps* $\operatorname{ad} X_i$*,* $1 \leq i \leq r$*, on* $\tilde{\mathfrak{g}}_S$ *are locally nilpotent.*

(ii) *The linear maps* $\operatorname{ad} Y_i$*,* $1 \leq i \leq r$*, on* $\tilde{\mathfrak{g}}_S$ *are locally nilpotent.*

Proof. Since the automorphism ϖ maps X_i into Y_i for all $1 \leq i \leq r$, it is enough to prove (i).

By (W4), we see that $\operatorname{ad} X_i(Y_j) = 0$ for $1 \leq i, j \leq r$ and $i \neq j$. Moreover, we have

$$(\operatorname{ad} X_i)^3(Y_i) = (\operatorname{ad} X_i)^2(H_i) = -2 \operatorname{ad} X_i(X_i) = 0$$

for all $1 \leq i \leq r$. Hence $(\operatorname{ad} X_i)^p(Y_j) = 0$ for p large enough and $1 \leq i, j \leq r$.

Also, we have

$$(\operatorname{ad} X_i)^2(H_j) = -\alpha_i(H_j)(\operatorname{ad} X_i)(X_i) = 0$$

for all $1 \leq i, j \leq r$.

Finally, $\operatorname{ad} X_i(X_i) = 0$, $1 \leq i \leq r$, and by Serre relations we have $(\operatorname{ad} X_i)^p(X_j) = 0$ for sufficiently large p and $1 \leq i, j \leq r$, $i \neq j$. Therefore, $(\operatorname{ad} X_i)^p(X_j) = 0$ for sufficiently large p and $1 \leq i, j \leq r$.

Assume that $u, v \in \tilde{\mathfrak{g}}_S$ such that $(\operatorname{ad} X_i)^p u = 0$ and $(\operatorname{ad} X_i)^q v = 0$. By induction we prove that

$$(\operatorname{ad} X_i)^m [u, v] = \sum_{p=0}^{m} \binom{m}{p} [(\operatorname{ad} X_i)^p u, (\operatorname{ad} X_i)^{m-p} v]$$

for all $m \in \mathbb{N}$. Therefore, we have $(\operatorname{ad} X_i)^{p+q}[u, v] = 0$. It follows that the linear subspace on which $\operatorname{ad} X_i$ is locally nilpotent is a Lie subalgebra of $\tilde{\mathfrak{g}}_S$. Since it contains all generators of $\tilde{\mathfrak{g}}_S$, $\operatorname{ad} X_i$ are locally nilpotent on $\tilde{\mathfrak{g}}_S$ for all $1 \leq i \leq r$. □

Therefore one can define linear maps $\exp(t \operatorname{ad} X_i)$ and $\exp(t \operatorname{ad} Y_i)$ on $\tilde{\mathfrak{g}}_S$ for $t \in \mathbb{C}$ and $1 \leq i \leq r$. These maps are automorphisms of the Lie algebra $\tilde{\mathfrak{g}}_S$.

For $1 \leq i \leq r$ we define the automorphism

$$\theta_i = e^{\operatorname{ad} X_i} e^{-\operatorname{ad} Y_i} e^{\operatorname{ad} X_i}.$$

We have

$$e^{\operatorname{ad} X_i}(H_j) = H_j + [X_i, H_j] = H_j - n(i,j) X_i.$$

by (W2). Moreover, we have

$$\begin{aligned} e^{-\operatorname{ad} Y_i} e^{\operatorname{ad} X_i}(H_j) &= e^{-\operatorname{ad} Y_i}(H_j - n(i,j) X_i) \\ &= H_j - n(i,j) X_i - [Y_i, H_j] + n(i,j)[Y_i, X_i] - \frac{1}{2} n(i,j)[Y_i, [Y_i, X_i]] \\ &= H_j - n(i,j) X_i - n(i,j) Y_i - n(i,j) H_i + n(i,j) Y_i = H_j - n(i,j) H_i - n(i,j) X_i \end{aligned}$$

Hence, it follows that

$$\begin{aligned} \theta_i(H_i) &= e^{\operatorname{ad} X_i}(H_j - n(i,j) H_i - n(i,j) X_i) \\ &= H_j - n(i,j) H_i - n(i,j) X_i - n(i,j) X_i + 2n(i,j) X_i \\ &= H_j - n(i,j) H_i = H_j - \alpha_i(H_j) H_i = s_{\alpha_i}(H_j). \end{aligned}$$

Therefore, the reflections corresponding to roots α_i, $1 \leq i \leq r$, on $\mathfrak{h}$ are induced by automorphisms of $\tilde{\mathfrak{g}}_S$. Since $W(R)$ is generated by these reflections, any element of $W(R)$ is induced by an automorphism of $\tilde{\mathfrak{g}}_S$.

4.4.5. LEMMA. (i) *The set of nonzero weights in $\tilde{\mathfrak{g}}_S$ is invariant under $W(R)$.*

(ii) *Let $w \in W(R)$ and λ a weight of $\tilde{\mathfrak{g}}_S$. Then $w\lambda$ is also a weight of $\tilde{\mathfrak{g}}_S$ and*

$$\dim \tilde{\mathfrak{g}}_{S, w\lambda} = \dim \tilde{\mathfrak{g}}_{S,\lambda}$$

PROOF. By the above discussion, there exists an automorphism A of $\tilde{\mathfrak{g}}_S$ such that $A(\mathfrak{h}) = \mathfrak{h}$ and $A|_{\mathfrak{h}} = w$. Therefore, for any $X \in \tilde{\mathfrak{g}}_{S,\lambda}$ we have

$$[H, A(X)] = A([A^{-1}H, X]) = \lambda(w^{-1}H) AX = (w\lambda)(H) A(X)$$

for any $H \in \mathfrak{h}$. Hence, we have $A(\tilde{\mathfrak{g}}_{S,\lambda}) = \tilde{\mathfrak{g}}_{S,w\lambda}$, and $w\lambda$ is a weight of $\tilde{\mathfrak{g}}_S$. □

4.4.6. LEMMA. *Let R be a root system in V. Let λ be a sum of roots which is not proportional to any root in R. Then there exists an element $W(R)$ such that $w\lambda = \sum_{\beta \in B} m_\beta \beta$ where $m_\beta \in \mathbb{Z}$ and $m_\beta > 0$ and $m_\gamma < 0$ for some $\beta, \gamma \in B$.*

PROOF. We can assume that V is a real vector space. Since λ is not proportional to any root α, λ must be nonzero. Let H be a hyperplane orthogonal to λ. Using the assumption, the linear forms defined by inner products with the roots in R are nonzero on H. Therefore, there exists a vector v in H such that $(\alpha \mid v) \neq 0$ for all $\alpha \in R$. By the proof of 3.2.3, there exists a basis B' of R such that $(\alpha \mid v) > 0$ for $\alpha \in B'$. By 3.4.4, there exists $w \in W(R)$ such that $w(B') = B$. Clearly, we have $w\lambda = \sum_{\beta \in B} m_\beta \beta$. Moreover, we have

$$0 = (\lambda \mid v) = (w\lambda \mid wv) = \sum_{\beta \in B} m_\beta (\beta \mid wv) = \sum_{\beta \in B} m_\beta (w^{-1}\beta \mid v).$$

Since $w^{-1}\beta$ are in B', we have $(w^{-1}\beta \mid v) > 0$ for all $\beta \in B$. Since λ is nonzero, the above equality can hold only if some m_β are strictly positive and some are strictly negative. □

4.4.7. LEMMA. (i) *The set of all nonzero weights of* $\tilde{\mathfrak{g}}_S$ *is equal to* R.
(ii) *For any* $\alpha \in R$*, the weight subspace* $\tilde{\mathfrak{g}}_{S,\alpha}$ *is one-dimensional.*
(iii) *The Lie algebra* $\tilde{\mathfrak{g}}_S$ *is finite dimensional.*

PROOF. (i) Let λ be a nonzero weight of $\tilde{\mathfrak{g}}_S$. Then it is also a nonzero weight of $\tilde{\mathfrak{g}}_W$. By 4.3.4, it is a sum of positive roots or the negative of a sum of positive roots.

By 4.4.5, $w\lambda$ is also a nonzero weight of $\tilde{\mathfrak{g}}_S$ for any $w \in W(R)$. Therefore, by 4.4.6 and the above remark, λ must be proportional to a root $\alpha \in R$.

Assume first that $\lambda = m\alpha_i$. Then, by 4.3.4, we see that $m = \pm 1$, i.e., $\lambda \in R$. On the other hand, if α is not in B, by 3.2.3 and 3.4.4, there is a $w \in W(R)$ such that $w\alpha$ is in B. By 4.4.5, $w\lambda$ is also a weight of $\tilde{\mathfrak{g}}_S$. Hence, by the first part of the argument, the weight $w\lambda$ of $\tilde{\mathfrak{g}}_S$ is in R. This implies that λ is in R in general.

If α is a root in R, by 3.2.3, α is in a basis B' of R. By 3.4.4, there exists $w \in W$ such that $w\alpha$ is in B. Therefore, $w\alpha$ is a weight of $\tilde{\mathfrak{g}}_S$. By 4.4.5, α is a weight of $\tilde{\mathfrak{g}}_S$. Hence, the set of all nonzero weights of $\tilde{\mathfrak{g}}_S$ is equal to R.

(ii) Let $\alpha \in R$. Then there exists $w \in W(R)$ such that $w\alpha \in B$. Hence, by 4.3.4 and 4.4.5, we have

$$\dim \tilde{\mathfrak{g}}_{S,\alpha} = \dim \tilde{\mathfrak{g}}_{S,w\alpha} = 1.$$

By 4.3.2, (i) and (ii) we have

$$\dim \tilde{\mathfrak{g}}_S = \dim \mathfrak{h} + \operatorname{Card} R.$$

□

4.4.8. LEMMA. *Let* α *be a root in* R*. Then*

$$\mathfrak{s}_\alpha = \tilde{\mathfrak{g}}_{S,\alpha} \oplus \tilde{\mathfrak{g}}_{S,-\alpha} \oplus [\tilde{\mathfrak{g}}_{S,\alpha}, \tilde{\mathfrak{g}}_{S,-\alpha}]$$

is a Lie subalgebra isomorphic to $\mathfrak{sl}(2,k)$.

PROOF. This follows from the Weyl relations if $\alpha \in B$. Let $\alpha \in R$ be arbitrary. Then by 3.2.3 and 3.4.4 there exists $w \in W(R)$ such that $\alpha = w\alpha_i$ for some $1 \leq i \leq r$. By the above discussion, there exists an automorphism ϕ of $\tilde{\mathfrak{g}}_S$ such that ϕ induces w on $\mathfrak{h}$. This in turn implies that $\phi(\mathfrak{s}_{\alpha_i}) = \mathfrak{s}_\alpha$. □

4.4.9. LEMMA. *The Lie algebra* $\tilde{\mathfrak{g}}_S$ *is semisimple.*

PROOF. Let $\mathfrak{c}$ be an abelian ideal in $\tilde{\mathfrak{g}}_S$. Since $\mathfrak{c}$ is invariant of the adjoint action of $\mathfrak{h}$ it is a direct sum of its weight subspaces. Assume that μ is a nonzero weight of $\mathfrak{h}$. Then, μ must be a weight of $\tilde{\mathfrak{g}}_S$. By 4.4.7, μ must be a root α in R. Since the root subspaces are one-dimensional by 4.4.7, it follows that $\tilde{\mathfrak{g}}_{S,\alpha} \subset \mathfrak{c}$. On the other hand, the Lie subalgebra $\mathfrak{s}_\alpha$ is simple by 4.4.8, hence $\mathfrak{c} \cap \mathfrak{s}_\alpha = \{0\}$, and we have a contradiction. It follows that $\mathfrak{c} \subset \mathfrak{h}$. Since $\mathfrak{c}$ is an ideal, all roots α_i must vanish on $\mathfrak{c}$, i.e., $\mathfrak{c} = \{0\}$. □

Clearly, $\mathfrak{h}$ is a nilpotent Lie subalgebra of $\tilde{\mathfrak{g}}_W$. Let $\mathfrak{m}$ be the normalizer of $\mathfrak{h}$. Then $\mathfrak{m}$ is a direct sum of its weight subspaces. Assume that λ is a nonzero weight of $\mathfrak{m}$. Then λ is a root in R. Let X be a nonzero vector in the root subspace $\tilde{\mathfrak{g}}_{W,\alpha}$. Then $X \in \mathfrak{m}$ by 4.4.7. Moreover, we have $\operatorname{ad} X(H) = -[H, X] = -\alpha(H)X$ for all $H \in \mathfrak{h}$. Since $\mathfrak{m}$ normalizes $\mathfrak{h}$, it follows that $X \in \mathfrak{h}$ contradicting our choice. Therefore, $\mathfrak{m} = \mathfrak{h}$ and $\mathfrak{h}$ is equal to its normalizer. Hence, $\mathfrak{h}$ is a Cartan subalgebra of $\tilde{\mathfrak{g}}_S$. This in turn implies that R is the root system of $\tilde{\mathfrak{g}}_S$.

Hence, for a given complex reduced root system R we constructed a complex semisimple Lie algebra $\mathfrak{g} = \tilde{\mathfrak{g}}_S$ with Cartan subalgebra $\mathfrak{h}$ and root system R.

4.4.10. THEOREM. *The Lie algebra $\tilde{\mathfrak{g}}_S$ is a finite dimensional semisimple Lie algebra over k. The Lie subalgebra $\mathfrak{h}$ is its Cartan subalgebra. The root system of $(\tilde{\mathfrak{g}}_S, \mathfrak{h})$ is equal to R.*

This gives a construction of a semisimple Lie algebra attached to a given reduced root system R.

5. Semisimple Lie algebras and their root systems

5.1. The isomorphism theorem. Assume that $\mathfrak{g}$ is a semisimple Lie algebra with Cartan subalgebra $\mathfrak{h}$ and root system R. Fix a basis B of R. Since Weyl and Serre relations hold in $\mathfrak{g}$ by 4.2.1, there exists a surjective morphism φ of $\tilde{\mathfrak{g}}_S$ onto $\mathfrak{g}$ which is the identity on $\mathfrak{h}$. Since the dimensions of $\mathfrak{g}$ and $\tilde{\mathfrak{g}}_S$ are the same, this morphism is an isomorphism.

Therefore, we proved the following theorem.

5.1.1. THEOREM. *Let $\mathfrak{g}$ and $\mathfrak{g}'$ are two semisimple Lie algebras over k with Cartan subalgebras $\mathfrak{h}$ and $\mathfrak{h}'$ respectively over an algebraically closed field k. Let R and R' be the root systems of $(\mathfrak{g}, \mathfrak{h})$ and $(\mathfrak{g}', \mathfrak{h}')$ respectively. Let φ be a linear isomorphism of $\mathfrak{h}$ into $\mathfrak{h}'$ such that its adjoint is an isomorphism of root system R' with R. Then φ extends to an isomorphism of Lie algebras.*

Therefore, the classification of semisimple Lie algebras over k is equivalent to the classification of reduced root systems.

5.2. Automorphisms of semisimple Lie algebras. Let $\mathfrak{g}$ be a semisimple Lie algebra over an algebraically closed field k. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$. Let $\operatorname{Aut}(\mathfrak{g})$ be the automorphism group of $\mathfrak{g}$. Let

$$\operatorname{Aut}(\mathfrak{g}, \mathfrak{h}) = \{T \in \operatorname{Aut}(\mathfrak{g}) \mid T(\mathfrak{h}) \subset \mathfrak{h}\}$$

be the subgroup of $\operatorname{Aut}(\mathfrak{g})$ consisting of automorphisms leaving the Cartan subalgebra $\mathfrak{h}$ invariant.

For any $T \in \operatorname{Aut}(\mathfrak{g}, \mathfrak{h})$, he restriction $T|_{\mathfrak{h}}$ of T to $\mathfrak{h}$ is an element of $\operatorname{GL}(\mathfrak{h})$. Clearly, this map is a homomorphism of $\operatorname{Aut}(\mathfrak{g}, \mathfrak{h})$ into $\operatorname{GL}(\mathfrak{h})$. Composing it with the homomorphism $C \longmapsto (C^*)^{-1}$ of $\mathrm{GL}(\mathfrak{h})$ into $\mathrm{GL}(\mathfrak{h}^*)$, we get a homomorphism $\varphi : \mathrm{Aut}(\mathfrak{g}, \mathfrak{h}) \longrightarrow \mathrm{GL}(\mathfrak{h}^*)$.

5.2.1. LEMMA. *Let T be an element in $\mathrm{Aut}(\mathfrak{g}, \mathfrak{h})$. Let α be a root in R. Then $\varphi(T)(\alpha)$ is also also a root in R.*

PROOF. Let $\mathfrak{g}_\alpha$ be the root subspace of $\mathfrak{g}$ corresponding to α. Let $X \in \mathfrak{g}_\alpha$. Then we have

$$[H, TX] = T[T^{-1}H, X] = \alpha(T^{-1}H)TX = (\varphi(T)\alpha)(H)TX$$

for all $H \in \mathfrak{h}$. Hence $\varphi(T)(\alpha) \in R$. □

Therefore, $\varphi(A)$ is in $\mathrm{Aut}(R)$. This implies that φ is a homomorphism of $\mathrm{Aut}(\mathfrak{g}, \mathfrak{h})$ into $\mathrm{Aut}(R)$.

From 5.1.1, we conclude that the following result holds.

5.2.2. THEOREM. *Let $\mathfrak{g}$ be a semisimple Lie algebra and $\mathfrak{h}$ its Cartan subalgebra. Then the homomorphism $\varphi : \mathrm{Aut}(\mathfrak{g}, \mathfrak{h}) \longrightarrow \mathrm{Aut}(R)$ is surjective.*

5.3. Chevalley systems. A choice of nonzero $X_\alpha \in \mathfrak{g}_\alpha$, $\alpha \in R$, is called a *Chevalley system* of $\mathfrak{g}$ if the linear map $\omega : \mathfrak{g} \longmapsto \mathfrak{g}$ defined by $\omega|\mathfrak{h} = -1_\mathfrak{h}$ and $\omega(X_\alpha) = -X_{-\alpha}$, $\alpha \in R$, is an automorphism of $\mathfrak{g}$.

5.3.1. LEMMA. *The following conditions are equivalent:*

(i) *X_α is a Chevalley system in $\mathfrak{g}_\alpha$;*
(ii) *for any $\alpha, \beta \in R$, such that $\alpha + \beta \in R$, we have $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$.*

PROOF. If (i) holds, for any $\alpha, \beta \in R$ such that $\alpha + \beta \in R$, we have

$$\begin{aligned} N_{-\alpha,-\beta}X_{-\alpha-\beta} = [X_{-\alpha}, X_{-\beta}] &= [\omega(X_\alpha), \omega(X_\beta)] \\ &= \omega([X_\alpha, X_\beta]) = N_{\alpha,\beta}\omega(X_{\alpha+\beta}) = -N_{\alpha,\beta}X_{-\alpha-\beta}. \end{aligned}$$

This clearly implies that $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$.

If (ii) holds, we have

$$\omega([H, X_\alpha]) = \alpha(H)\omega(X_\alpha) = -\alpha(H)X_{-\alpha} = [H, X_{-\alpha}] = [\omega(H), \omega(X_\alpha)];$$

and

$$\omega([X_\alpha, X_{-\alpha}]) = \omega(H_\alpha) = -H_\alpha = H_{-\alpha} = [X_{-\alpha}, X_\alpha] = [\omega(X_\alpha), \omega(X_{-\alpha})]$$

for all $H \in \mathfrak{h}$ and $\alpha \in R$. Moreover, for any $\alpha, \beta \in R$ such that $\alpha + \beta \in R$, we have

$$\omega([X_\alpha, X_\beta]) = N_{\alpha,\beta}\, \omega(X_{\alpha+\beta}) = N_{-\alpha,-\beta}\, X_{-\alpha-\beta} = [X_{-\alpha}, X_{-\beta}] = [\omega(X_\alpha), \omega(X_\beta)].$$

Hence, ω is an automorphism of $\mathfrak{g}$. □

5.3.2. THEOREM. *Let $\mathfrak{g}$ be a semisimple Lie algebra over an algebraic closed field k. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$. Then there exists a Chevalley system of $\mathfrak{g}$.*

PROOF. Clearly $-1_{\mathfrak{h}^*}$ is an automorphism of the root system R. By 5.2.2, there exists an automorphism ω of $\mathfrak{g}$ such that $\omega|\mathfrak{h} = -1_\mathfrak{h}$. For any $\alpha \in R$, we clearly have $\omega(X_\alpha) \in \mathfrak{g}_{-\alpha}$, so $\omega(X_\alpha) = -t_\alpha X_{-\alpha}$ with $t_\alpha \in k^*$. Moreover, we have

$$t_\alpha t_{-\alpha} H_\alpha = t_\alpha t_{-\alpha}[X_\alpha, X_{-\alpha}] = [\omega(X_{-\alpha}), \omega(X_\alpha)] = \omega([X_{-\alpha}, X_\alpha]) = \omega(H_{-\alpha}) = H_\alpha$$

for any $\alpha \in R$. Hence, we have $t_\alpha t_{-\alpha} = 1$ for all $\alpha \in R$. We pick square roots u_α of t_α such that $u_\alpha u_{-\alpha} = 1$. Put $Y_\alpha = u_\alpha^{-1} X_\alpha$ for $\alpha \in R$. Then, we have

$$[Y_\alpha, Y_{-\alpha}] = u_\alpha^{-1} u_{-\alpha}^{-1} [X_\alpha, X_{-\alpha}] = H_\alpha$$

for all $\alpha \in R$. Moreover,

$$\omega(Y_\alpha) = u_\alpha^{-1}\omega(X_\alpha) = -u_\alpha^{-1} t_\alpha X_{-\alpha} = -u_\alpha X_{-\alpha} = -u_{-\alpha}^{-1} X_{-\alpha} = -Y_{-\alpha}$$

for any $\alpha \in R$. Hence, Y_α, $\alpha \in R$, is a Chevalley system. □

By combining 2.2.4 with 5.3.2, we get the following immediate consequence.

5.3.3. Corollary. *Let X_α, $\alpha \in R$, be a Chevalley system in $\mathfrak{g}$. Let α, β be two roots such that $\alpha + \beta \in R$. Let $\beta - q\alpha$ be the start of the α-string determined by β. Then*

$$N_{\alpha,\beta} = \pm(q+1).$$

Hence, there is a basis of $\mathfrak{g}$ such that all structure constants are integers.

CHAPTER 6

Classification of compact Lie groups

1. Compact semisimple Lie groups

1.1. Real forms. In this section we specalize the results of 4.2 to the case of the field $\mathbb{R}$ of real numbers imbedded into the filed $\mathbb{C}$ of complex numbers.

Let $\mathfrak{g}$ be a complex Lie algebra. A $\mathbb{R}$-structure on $\mathfrak{g}$ is called a *real form* of $\mathfrak{g}$. If $\mathfrak{g}_0$ is a real Lie algebra, the Lie algebra $\mathfrak{g} = \mathbb{C} \otimes_{\mathbb{R}} \mathfrak{g}_0$ is called the *complexification* of $\mathfrak{g}_0$. Clearly, $\mathfrak{g}_0$ is a real form of $\mathfrak{g}$.

The Galois group $\mathrm{Aut}_{\mathbb{R}}(\mathbb{C})$ of all $\mathbb{R}$-linear automorphisms of $\mathbb{C}$ consists of two elements: the identity and the complex conjugation $z \longmapsto \bar{z}$. If $\mathfrak{g}$ is a complex Lie algebra and $\mathfrak{g}_0$ its real form, the Galois group $\mathrm{Aut}_{\mathbb{R}}(\mathbb{C})$ acts naturally on $\mathfrak{g}$ and $\mathfrak{g}_0$ is the set of all fixed points of this action. We denote by σ the action of the complex conjugation in $\mathrm{Aut}_{\mathbb{R}}(\mathbb{C})$ on $\mathfrak{g}$. Clearly, σ satisfies the following properties:

(i) σ is $\mathbb{R}$-linear map;
(ii) $\sigma(\lambda x) = \bar{\lambda}\sigma(x)$ for any $\lambda \in \mathbb{C}$ and $x \in \mathfrak{g}$;
(iii) $\sigma([x, y]) = [\sigma(x), \sigma(y)]$ for any $x, y \in \mathfrak{g}$;
(iv) $\sigma^2 = 1_{\mathfrak{g}}$.

Such map is called a *conjugation* of $\mathfrak{g}$.

On the other hand, if σ is a conjugation of $\mathfrak{g}$, then the set of its fixed points is a $\mathbb{R}$-linear subspace $\mathfrak{g}_0$ of $\mathfrak{g}$. By (ii), $\mathfrak{g}_0$ is a real Lie subalgebra of $\mathfrak{g}_0$. Let $x \in \mathfrak{g}$. Then $x_1 = \frac{1}{2}(x+\sigma(x))$ and $x_2 = -\frac{i}{2}(x-\sigma(x))$ are in $\mathfrak{g}_0$, and $x = x_1 + ix_2$. Therefore, the complex linear map $\mathbb{C} \otimes_{\mathbb{R}} \mathfrak{g}_0 \longrightarrow \mathfrak{g}$ is surjective. On the other hand, if $x = x' + ix''$ for $x', x'' \in \mathfrak{g}_0$, we have $x + \sigma(x) = 2x' = 2x_1$ and $x - \sigma(x) = 2ix'' = 2ix_2$, i.e., $x' = x_1$ and $x'' = x_2$. Hence, it follows that the map $\mathbb{C} \otimes_{\mathbb{R}} \mathfrak{g}_0 \longrightarrow \mathfrak{g}$ is also injective, i.e., $\mathfrak{g}_0$ is a real form of $\mathfrak{g}$. Moreover, the conjugation σ is equal to the action of the nontrivial element of the Galois group with respect to this real form. Therefore, the study of real forms of $\mathfrak{g}$ is equivalent to the study of conjugations of $\mathfrak{g}$.

Let K be a connected semisimple Lie group. Then by 3.1.8.9, K is a compact group if and only if the Killing form on $\mathfrak{k}$ is negative definite.

We say that a real semisimple Lie algebra is *compact*, if its Killing form is negative definite.

If $\mathfrak{g}$ is a complex semisimple Lie algebra and $\mathfrak{g}_0$ its real form, we say that $\mathfrak{g}_0$ is a *compact real form* if the Killing form is negative definite on $\mathfrak{g}_0$. We say that the corresponding conjugation of $\mathfrak{g}$ is a *compact conjugation* of $\mathfrak{g}$.

1.1.1. Lemma. *Let σ be a conjugation on a complex semisimple Lie algebra $\mathfrak{g}$. Then, the following statements are equivalent:*

(i) *the conjugation σ is compact;*
(ii) *the form $(x, y) \longmapsto -B(x, \sigma(y))$ is an inner product on $\mathfrak{g}$.*

PROOF. By the construction, the form $(x, y) \longmapsto -B(x, \sigma(y))$ is linear in the first and antilinear in the second variable.

If it is an inner product, its restriction to $\mathfrak{g}_0 \times \mathfrak{g}_0$ is obviously positive definite and also equal to the Killing form of $\mathfrak{g}_0$. Therefore, $\mathfrak{g}_0$ is a compact form.

Conversely, if $\mathfrak{g}_0$ is a compact form and $x \in \mathfrak{g}$, $x = x_1 + ix_2$ with $x_1, x_2 \in \mathfrak{g}_0$. Therefore,

$$\begin{aligned} -B(x, \sigma(x)) &= -B(x_1 + ix_2, x_1 - ix_2) \\ &= -B(x_1, x_1) + iB(x_1, x_2) - B(x_2, x_2) - B(x_2, x_1) \\ &= -(B(x_1, x_1) + B(x_2, x_2)) \geq 0 \end{aligned}$$

since the Killing form is negative definite on $\mathfrak{g}_0 \times \mathfrak{g}_0$. Moreover, $B(x, \sigma(x)) = 0$ implies $B(x_1, x_1) = B(x_2, x_2) = 0$ and $x_1 = x_2 = 0$. Therefore, $(x, y) \longmapsto -B(x, \sigma(y))$ is positive definite. □

Let σ be a conjugation of $\mathfrak{g}$ and φ an automorphism of $\mathfrak{g}$. Then $\varphi \circ \sigma \circ \varphi^{-1}$ is a conjugation of $\mathfrak{g}$. Therefore, the automorphism group $\operatorname{Aut}(\mathfrak{g})$ acts on the set of all conjugations of $\mathfrak{g}$.

Assume that the Lie algebra $\mathfrak{g}$ is semisimple. Since the Killing form is invariant under the action of the automorphisms of $\mathfrak{g}$, this action of $\operatorname{Aut}(\mathfrak{g})$ on the set of all conjugations of $\mathfrak{g}$ preserves the subset of compact conjugations.

1.2. Existence of compact forms. Let $\mathfrak{g}$ be a complex semisimple Lie algebra. In this section we are going to construct a compact form of $\mathfrak{g}$. In the following section we shall prove that this compact form is unique up to an inner automorphism.

Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$. Let R be the corresponding root system in $\mathfrak{h}^*$. We fix a Chevalley system of $\mathfrak{g}$ attached to $\mathfrak{h}$. Let $\mathfrak{h}_0$ be the real linear subspace of $\mathfrak{h}$ spanned by H_α, $\alpha \in R$. Then $\mathfrak{h}_0$ is a real form of $\mathfrak{h}$.

The real span $\mathfrak{g}_0$ of $\mathfrak{h}_0$ and X_α, $\alpha \in R$, is a $\mathbb{R}$-structure on the complex linear space $\mathfrak{g}$. Moreover, since the structure constants $N_{\alpha,\beta}$ are real, we see that $\mathfrak{g}_0$ is a real Lie algebra, i.e., it is a real form of $\mathfrak{g}$. This real form is called a *normal form* of $\mathfrak{g}$. We denote the corresponding conjugation by σ.

Let ω be the automorphism of $\mathfrak{g}$ described in 5.3. Since ω leaves $\mathfrak{g}_0$ invariant, it is clear that $\sigma \circ \omega = \omega \circ \sigma$. Put $\tau = \sigma \circ \omega$. Then

$$\tau^2 = \sigma \circ \omega \circ \sigma \circ \omega = \sigma \circ \omega \circ \omega \circ \sigma = 1_{\mathfrak{g}}$$

and τ is clearly a conjugation of $\mathfrak{g}$.

1.2.1. LEMMA. *The conjugation τ is compact.*

PROOF. By 1.1.1, we have to check that $(x, y) \longmapsto -B(x, \tau(y))$ is positive definite.

First, we remark that a root α takes real values on $\mathfrak{h}_0$. This implies that

$$\alpha(\tau(H)) = \alpha(\sigma(\omega(H))) = -\alpha(\sigma(H)) = -\overline{\alpha(H)}$$

for any $H \in \mathfrak{h}$. Hence, by 2.1.7, we have

$$-B(H, \tau(H')) = -\sum_{\alpha \in R} \alpha(H)\alpha(\tau(H')) = \sum_{\alpha \in R} \alpha(H)\overline{\alpha(H')}$$

for any $H, H' \in \mathfrak{h}$. So, $(H, H') \longmapsto -B(H, \tau(H'))$ is an inner product on $\mathfrak{h}$.

If $H \in \mathfrak{h}$ and $\alpha \in R$, we have

$$-B(H, \tau(X_\alpha)) = B(H, X_{-\alpha}) = 0$$

by 2.1.4. Hence, all $\mathfrak{g}_\alpha$, $\alpha \in R$, are orthogonal to $\mathfrak{h}$.

Moreover, if $\alpha, \beta \in R$, $\beta \neq \alpha$, we have

$$-B(X_\alpha, \tau(X_\beta)) = B(X_\alpha, X_{-\beta}) = 0$$

by 2.1.4. Hence, all $\mathfrak{g}_\alpha$, $\alpha \in R$, are orthogonal to each other.

Finally, by 2.2.2 and 2.1.7, we have

$$-B(X_\alpha, \tau(X_\alpha)) = B(X_\alpha, X_{-\alpha}) = \frac{1}{2} B(H_\alpha, H_\alpha) \geq 0$$

for any $\alpha \in R$. Hence, $(x, y) \longmapsto -B(x, \tau(y))$ is positive definite. □

Therefore, the real Lie subalgebra $\mathfrak{u}_0$ consisting of fixed points of τ is a compact real form of $\mathfrak{g}$.

By the above discussion, $\mathfrak{u}_0$ is spanned by iH_α, $X_\alpha - X_{-\alpha}$, $i(X_\alpha + X_{-\alpha})$, $\alpha \in R$.

1.2.2. Theorem. *Let $\mathfrak{g}$ be a complex semisimple Lie subalgebra. Then, the real Lie subalgebra $\mathfrak{u}_0$ is a compact form of $\mathfrak{g}$.*

1.3. Uniqueness of compact forms. In this section we prove the uniqueness of compact forms of complex semisimple Lie algebras (up to conjugacy by inner automorphisms). This proof doesn't depend on the structure theory of semisimple Lie algebras like the existence result proved in the last section.

1.3.1. Theorem. *Let τ be a compact conjugation of a complex semisimple Lie algebra $\mathfrak{g}$ and σ an arbitrary conjugation of $\mathfrak{g}$. Then there exists $\varphi \in \mathrm{Aut}(\mathfrak{g})$ such that $\varphi \circ \tau \circ \varphi^{-1}$ and σ commute.*

Proof. Define an inner product on $\mathfrak{g}$ by

$$(x \mid y) = -B(x, \tau y)$$

for $x, y \in \mathfrak{g}$. Let $A = \sigma\tau$. Then A is an automorphism of $\mathfrak{g}$. Hence it leaves the Killing form invariant, and we have

$$(Ax|y) = -B(\sigma\tau x, \tau y) = -B(x, (\sigma\tau)^{-1}\tau y) = -B(x, \tau\sigma\tau y) = -B(x, \tau A y) = (x|Ay)$$

for any $x, y \in \mathfrak{g}$. Hence, A is a selfadjoint linear operator on $\mathfrak{g}$. Let $X_1, X_2, \dots, X_n$ be an orthonormal basis of $\mathfrak{g}$ in which A is represented by a diagonal matrix, i.e.,

$$AX_i = \lambda_i X_i$$

for $1 \leq i \leq n$ and $\lambda_i \in \mathbb{R}$. Put $C = A^2$. Then C is a positive selfadjoint operator and

$$CX_i = \lambda_i^2 X_i$$

for $1 \leq i \leq n$. Let $t \in \mathbb{R}$. Then we put

$$C^t X_i = (\lambda_i^2)^t X_i$$

for $1 \leq i \leq n$.

Let $c_{ijk} \in \mathbb{C}$, $1 \leq i, j, k \leq n$, be the structure constants of $\mathfrak{g}$, i.e.,

$$[X_i, X_j] = \sum_{k=1}^{n} c_{ijk} X_k$$

for $1 \leq i.j \leq n$. Clearly, a linear automorphism S of $\mathfrak{g}$ is an automorphism of the Lie algebra $\mathfrak{g}$ if and only if

$$[SX_i, SX_j] = \sum_{k=1}^{n} c_{ijk} SX_k$$

for any $1 \leq i, j \leq k$. Since A is an automorphism of $\mathfrak{g}$, it follows that

$$\lambda_i \lambda_j \sum_{k=1}^{n} c_{ijk} X_k = \lambda_i \lambda_j [X_i, X_j] = [AX_i, AX_j] = \sum_{k=1}^{n} c_{ijk} AX_k = \sum_{k=1}^{n} c_{ijk} \lambda_k X_k$$

for $1 \leq i.j \leq n$. This in turn implies that

$$(\lambda_i \lambda_j - \lambda_k) c_{ijk} = 0$$

for $1 \leq i, j, k \leq n$. In other words, if $c_{ijk} \neq 0$, we must have $\lambda_i \lambda_j = \lambda_k$. This in turn implies that

$$(\lambda_i^2)^t (\lambda_j^2)^t = (\lambda_k^2)^t$$

for any $t \in \mathbb{R}$. Hence, we have

$$\begin{aligned}[C^t X_i, C^t X_j] = (\lambda_i^2)^t (\lambda_j^2)^t [X_i, X_j] = (\lambda_i^2)^t (\lambda_j^2)^t \sum_{k=1}^{n} c_{ijk} X_k \\ = \sum_{k=1}^{n} c_{ijk} (\lambda_k^2)^t X_k = \sum_{k=1}^{n} c_{ijk} C^t X_k\end{aligned}$$

for any $1 \leq i, j \leq n$; and C^t is an automorphism of $\mathfrak{g}$ for any $t \in \mathbb{R}$.

This clearly implies that

$$\tau' = C^t \circ \tau \circ C^{-t}$$

is a compact conjugation of $\mathfrak{g}$. Now, we have

$$\tau \circ A \circ \tau = \tau(\sigma\tau)\tau = \tau\sigma = A^{-1}.$$

Hence, it follows that

$$\tau \circ C \circ \tau = \tau \circ A^2 \circ \tau = (\tau \circ A \circ \tau) \circ (\tau \circ A \circ \tau) = (A^{-1})^2 = C^{-1}.$$

This finally implies that

$$\tau \circ C = C^{-1} \circ \tau.$$

It follows that

$$C^{-1}(\tau X_i) = \tau C X_i = \lambda_i^2 \tau X_i$$

for $1 \leq i \leq n$. Hence τX_i is an eigenvector of C^{-1} for the eigenvalue λ_i^2. Hence, we have

$$C^{-t} \tau X_i = (\lambda_i^2)^t \tau X_i = \tau C^t X_i$$

for $1 \leq i \leq n$; i.e.,

$$\tau \circ C^t \circ \tau = C^{-t}$$

for any $t \in \mathbb{R}$. This implies that

$$\sigma\tau' = \sigma(C^t \circ \tau \circ C^{-t}) = (\sigma\tau) \circ (\tau \circ C^t \circ \tau) \circ C^{-t} = AC^{-2t}$$

and

$$\tau'\sigma = (\sigma\tau')^{-1} = (AC^{-2t})^{-1} = C^{2t} A^{-1}.$$

It $t = \frac{1}{4}$, we have

$$\sigma\tau' = AC^{-\frac{1}{2}}$$

and
$$\tau'\sigma = C^{\frac{1}{2}}A^{-1}.$$
This in turn implies that
$$\sigma\tau' X_i = AC^{-\frac{1}{2}}X_i = \lambda_i(\lambda_i^2)^{-\frac{1}{2}}X_i = \operatorname{sgn}(\lambda_i)X_i$$
and
$$\tau'\sigma X_i = C^{\frac{1}{2}}A^{-1}X_i = (\lambda_i^2)^{\frac{1}{2}}\lambda_i^{-1}X_i = \operatorname{sgn}(\lambda_i)X_i$$
for $1 \leq i \leq n$. It follows that
$$\sigma\tau' = \tau'\sigma.$$
□

This result has a simple consequence.

1.3.2. Corollary. *Let τ and τ' be two compact conjugations of a complex semisimple Lie algebra $\mathfrak{g}$. Then there exists $\varphi \in \operatorname{Aut}(\mathfrak{g})$ such that $\tau' = \varphi \circ \tau \circ \varphi^{-1}$.*

Proof. From the preceding theorem, we can find $\varphi \in \operatorname{Aut}(\mathfrak{g})$ such that $\tau'' = \varphi\circ\tau'\circ\varphi^{-1}$ commutes with τ. As in the proof of the preceding theorem, we can view $\mathfrak{g}$ as an inner product space equipped with the inner product $(x \mid y) = -B(x, \tau(y))$ Let $A = \tau''\tau$. Then, as we established there, A is a selfadjoint linear transformation. Moreover, since τ'' and τ commute,
$$A^{-1} = (\tau''\tau)^{-1} = \tau\tau'' = \tau''\tau = A$$
and $A^2 = I$. Therefore, the eigenvalues of A are 1 and -1. On the other hand,
$$(Ax \mid y) = (x \mid Ay) = -B(x, \tau\tau''\tau y) = -B(x, \tau'' y)$$
and $(x, y) \longmapsto (Ax \mid y)$ is another inner product on $\mathfrak{g}$. Therefore, A has to be a positive selfadjoint operator, i.e., -1 cannot be its eigenvalue. It follows that $A = I$, i.e., $\tau'' = \tau$. □

Therefore, the compact conjugations form a single orbit of $\operatorname{Aut}(\mathfrak{g})$ in the set of all conjugations of $\mathfrak{g}$.

Finally, combining 1.2.2 and 1.3.2 we get the following result.

1.3.3. Theorem. *Let $\mathfrak{g}$ be a complex semisimple Lie algebra. Then:*

(i) *there exists a compact form of $\mathfrak{g}$;*
(ii) *all compact forms are conjugate by automorphisms of $\mathfrak{g}$.*

Therefore, every compact semisimple Lie algebra is a compact real form of a complex semisimple Lie algebra. Moreover, this complex semisimple Lie algebra is unique up to an isomorphism. Therefore, compact semisimple Lie algebras are determined up to an isomorphism by their root systems.

1.4. A basis of the Lie algebra of a connected compact semisimple Lie group. Let G be a connected compact semisimple Lie group. Let T be a maximal torus in G. Then the Lie algebra $L(T)$ of T is a maximal abelian Lie subalgebra of $L(G)$.

The Lie algebra $L(G)$ is a compact form of its complexification, the Lie algebra $\mathfrak{g} = L(G) \otimes_{\mathbb{R}} \mathbb{C}$ of $L(G)$. By 1.3.3 and the construction in Section 1.2, we can find a Cartan subalgebra $\mathfrak{h}$, the root system R in $\mathfrak{h}^*$ and the Chevalley system $(X_\alpha \in \mathfrak{g}_\alpha; \alpha \in R)$ such that $L(G)$ is spanned by iH_α, $i(X_\alpha + X_{-\alpha})$ and $X_\alpha - X_{-\alpha}$ for $\alpha \in R$. Clearly, iH_α, $\alpha \in R$, span a real abelian Lie subalgebra $\mathfrak{h}_0$ of $L(G)$. Its complexification in $\mathfrak{g}$ is $\mathfrak{h}$. Since $\mathfrak{h}$ is a maximal abelian Lie subalgebra in $\mathfrak{g}$ by **??**, $\mathfrak{h}_0$ is a maximal abelian Lie subalgebra in $L(G)$. By **??**, all maximal abelian Lie subalgebras of $L(G)$ are conjugate, so without any loss of generality we can assume that $L(T) = \mathfrak{h}_0$.

1.4.1. Lemma. (i) *The complexification of $L(T)$ is a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$.*
(ii) *There exists a Chevalley system $(X_\alpha; \alpha \in R)$ such that $L(G)$ is spanned by iH_α, $X_\alpha - X_{-\alpha}$ and $i(X_\alpha + X_{-\alpha})$ for $\alpha \in R$.*

1.5. The Weyl group of connected compact semisimple Lie group. Let G be a connected compact semisimple Lie group and T a maximal torus in G. Denote by $L(G)$ and $L(T)$ their Lie algebras respectively. Let $\mathfrak{g}$ be the complexification of $L(G)$ and $\mathfrak{h}$ the complexification of $L(T)$. Then $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$.

Let Ad be the adjoint representation of G on $L(G)$. We can extend it by linearity to a representation of G on $\mathfrak{g}$ which we denote by the same symbol. Clearly, Ad is a group morphism of G into $\mathrm{Aut}(\mathfrak{g})$. Let $N(T)$ be the normalizer of T in G. Then, for any $n \in N(T)$, we have $\mathrm{Ad}(n)(L(T)) = L(T)$, and $\mathrm{Ad}(n)(\mathfrak{h}) = \mathfrak{h}$. It follows that Ad induces a group morphism ψ of $N(T)$ into $\mathrm{Aut}(\mathfrak{g}, \mathfrak{h})$. Composing it with the group morphism $\mathrm{Aut}(\mathfrak{g}, \mathfrak{h}) \longrightarrow \mathrm{Aut}(R)$ described in 5.5.2.2, we get a group morphism $\Psi : N(T) \longrightarrow \mathrm{Aut}(R)$.

Since T is abelian, for any $t \in T$, we have $\mathrm{Ad}(t)|_{L(T)} = I_{L(T)}$, i.e., T is in the kernel of Ψ. Therefore, Ψ induces a group morphism of $N(T)/T$ into $\mathrm{Aut}(R)$.

In addition, if $n \in N(T)$ and $\Psi(n) = I$, we have $\mathrm{Ad}(n)|_{L(T)} = I_{L(T)}$. Therefore, by 2.2.2.16, n is in the centralizer of T. By 3.1.4.1, it follows that $n \in T$. Therefore, the kernel of Ψ is equal to T. Hence Ψ is an injective group morphism.

As in Ch. 3, we denote by $W(G, T) = N(T)/T$ the Weyl group of G with respect to T. Let $\Phi : W(G, T) \longrightarrow \mathrm{Aut}(R)$ be the group morphism we defined above. The aim of this section is to establish the following result.

1.5.1. Theorem. *The group morphism $\Phi : W(G, T) \longrightarrow \mathrm{Aut}(R)$ induces an isomorphism of $W(G, T)$ with $W(R)$.*

Therefore, two definitions of the Weyl group are the same.

We start the proof with some preparation. Let $t \in T$. Then $\mathfrak{h}$ is invariant for $\mathrm{Ad}(t)$. Moroever, it acts there as identity map. Therefore, the multiplicity of the eigenvalue 1 in $\mathrm{Ad}(t)$ is greater or equal to $\dim \mathfrak{h} = \dim L(T) = \mathrm{rank}(G)$. Since for any $g \in G$, the linear maps $\mathrm{Ad}(t)$ and $\mathrm{Ad}(gtg^{-1})$ have equal characteristic polynomials, by 3.1.3.2, we see that the following result holds.

1.5.2. Lemma. *Let $g \in G$. Then $\mathrm{Ad}(g)$ is semisimple and the multiplicity of the eigenvalue 1 is greater or equal to $\mathrm{rank}(G)$.*

The following result is central for our proof.

1.5.3. Lemma. *Let B be a basis of the root system R. Let n be an element of $N(T)$ such that $\Psi(n)(B) = B$. Then $n \in T$.*

Proof. Since $\Psi(n)$ permutes B, it also permutes R^+. Therefore, R^+ is a union of orbits of $\Psi(n)$. Let O be such orbit. Then it consists of positive roots $\Psi(n)^k\alpha$, $0 \le k \le l$. The direct sum U of the corresponding root subspaces $\mathfrak{g}_{\Psi(n)^k\alpha}$ is an invariant subspace for $\mathrm{Ad}(n)$ in $\mathfrak{g}$. If we pick a basis of root vectors in these root subspaces, the matrix of the restriction of $\mathrm{Ad}(n)$ to U looks like

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & a_{l-1} \\ a_0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & a_1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & a_{l-2} & 0 \end{pmatrix}.$$

Therefore, the characteristic polynomial of $\mathrm{Ad}(n)|_U$ is equal to $\lambda^l - a_0 a_1 \cdots a_{l-1} = \lambda^l - (-1)^l \det(\mathrm{Ad}(n)|_U)$.

If we multiply n by an element t in the torus T, the determinant changes from $\det(\mathrm{Ad}(n)|_U)$ to

$$\det(\mathrm{Ad}(nt)|_U) = \det(\mathrm{Ad}(n)|_U \,\mathrm{Ad}(t)|_U) = \det(\mathrm{Ad}(n)|_U)\det(\mathrm{Ad}(t)|_U).$$

The function $t \longmapsto \det(\mathrm{Ad}(t)|_U)$ is a character of T. The matrix of $\mathrm{Ad}(t)$ is diagonal in the basis given by root vectors, and diagonal matrix coefficients are c_β, $\beta \in R$. Since all roots in the orbit O are positive, $\det(\mathrm{Ad}(t)|_U) = \prod_{\beta\in B} c_\beta^{n_\beta}$ where n_β are positive integers and not all of them are zero. Therefore, the differential of this character of T extends to a linear form on $\mathfrak{h}$ which is equal to $\sum_{\alpha\in B} n_\alpha \alpha$. Clearly, it must be nonzero, and the character $t \longmapsto \det(\mathrm{Ad}(t)|_U)$ is not equal to 1 on T.

An analogous argument applies to orbits of $\Psi(n)$ in negative roots.

We can find $t \in T$ such that $\det(\mathrm{Ad}(t)|_U) \neq (-1)^l \det(\mathrm{Ad}(n)|_U)^{-1}$ for any orbit O in the set R. This in turn implies that $\mathrm{Ad}(nt)|_U$ does not have eigenvalue 1 for any orbit O in R. Therefore, the restriction of $\mathrm{Ad}(nt)$ to $\bigoplus_{\alpha\in R}\mathfrak{g}_\alpha$ does not have eigenvalue 1. Hence, only $\mathfrak{h}$ can contain the eigenspace of $\mathrm{Ad}(nt)$ for eigenvalue 1. By 1.5.2, the eigenspace of $\mathrm{Ad}(nt)$ for the eigenvalue 1 has to be at least of dimension $\mathrm{rank}(G)$. This implies that the restriction of $\mathrm{Ad}(nt)$ to $\mathfrak{h}$ is the identity map. Hence, $\Psi(n) = \Psi(nt)$ is also the identity. Since Φ is injective, it follows that $n \in T$. □

We also need the following result. We start with the discussion of a compact form of $\mathfrak{sl}(2,\mathbb{C})$. Clearly,

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\ e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

is a Chevalley system for $\mathfrak{sl}(2,\mathbb{C})$. Therefore,

$$ih = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},\ e - f = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\ i(e+f) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

is the basis of the compact real form of $\mathfrak{sl}(2,\mathbb{C})$ consisting of skewadjoint 2×2 matrices with trace 0. This is the Lie algebra $\mathfrak{su}(2)$ of $\mathrm{SU}(2)$.

Consider the element in $SU(2)$ given by the matrix

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

Then $\mathrm{Ad}(m)(h) = -h$, $\mathrm{Ad}(m)(e) = -f$ and $\mathrm{Ad}(m)(f) = -e$. The one-dimensional torus

$$C = \left\{ \begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix} \middle|\, \varphi \in R \right\}$$

of SU(2) has the Lie algebra spanned by ih. Moreover, its normalizer is the union of C and the coset mC. Therefore, the Weyl group of SU(2) with respect to C is isomorphic to $\mathbb{Z}/2\mathbb{Z}$.

1.5.4. LEMMA. *Let α be a root in R. Then there exists an element m_α in $N(T)$ such that $\Psi(m_\alpha) = s_\alpha$.*

PROOF. We know from 5.2.1.5, that $\mathfrak{g}_\alpha$, $\mathfrak{g}_{-\alpha}$ and $\mathfrak{h}_\alpha$ span a Lie subalgebra $\mathfrak{s}_\alpha$ of $\mathfrak{g}$ and there is an isomorphism φ_α of $\mathfrak{sl}(2,\mathbb{C})$ into $\mathfrak{s}_\alpha$. The compact conjugation of $\mathfrak{g}$ defining $L(G)$ leaves $\mathfrak{s}_\alpha$ invariant. Moreover it induces a compact conjugation on $\mathfrak{s}_\alpha$ such that $\mathfrak{u}_\alpha = L(G) \cap \mathfrak{s}_\alpha$ is isomorphic to $\mathfrak{su}(2)$ under φ_α. The corresponding integral subgroup U_α of G is isomorphic to either SU(2) or SO(3). Therefore, the element m in SU(2), or its image in SO(3), define an element m_α in G.

Since we have
$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$
we see that $m_\alpha(H_\alpha) = -H_\alpha$. On the other hand, elements of $\mathrm{ad}(\mathfrak{s}_\alpha)$ act trivially on $\ker\alpha \subset \mathfrak{h}$. Therefore, m_α acts trivially on $\ker\alpha$. It follows that m_α acts as s_{H_α} on $\mathfrak{h}$. Hence, we have $\Psi(m_\alpha) = s_\alpha$. □

Since the Weyl group $W(R)$ is generated by reflections s_α, $\alpha \in R$, the above lemma implies the following result.

1.5.5. COROLLARY. *The image Ψ contains $W(R)$.*

Now we can complete the proof of 1.5.1.

Let n be in $N(T)$. Then $\Psi(n)$ is an automorphism of R which maps a basis B of R into another basis B' of R. By 5.3.4.5, $W(R)$ acts transitively on all bases of R. Hence, there exists a Weyl group element $w \in W(R)$ such that $w(B') = B$. By 1.5.5, there exists an element n' in $N(T)$ such that $\Psi(n') = w$. Hence, we have
$$\Psi(n'n) = \Psi(n')\Psi(n) = w\Psi(n)$$
and $\Psi(n'n)(B) = B$. Since $n'n \in N(T)$, by 1.5.3 we see that $n'n \in T$. Therefore $\Psi(n'n) = I$, and $\Psi(n) = w^{-1} \in W(R)$. Hence, the image of Ψ is contained in $W(R)$. By 1.5.5, it follows that the image of Φ is equal to $W(R)$. This completes the proof of 1.5.1.

1.6. Adjoint groups. Let $\mathfrak{g}$ be a compact semisimple Lie algebra. Let $B : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathbb{R}$ be its Killing form. Then $(x \mid y) = -B(x,y)$ is a positive definite inner product on $\mathfrak{g}$. Denote by $\mathrm{Aut}(\mathfrak{g})$ the automorphism group of $\mathfrak{g}$. Since the Killing form B is invariant for $\mathrm{Aut}(\mathfrak{g})$ by 3.1.8.8, the inner product on $\mathfrak{g}$ is invariant for $\mathrm{Aut}(\mathfrak{g})$, i.e., $\mathrm{Aut}(\mathfrak{g})$ is a closed subgroup of the orthogonal group $\mathrm{O}(\mathfrak{g})$. Hence, by Cartan's theorem, $\mathrm{Aut}(\mathfrak{g})$ is a compact Lie group. The Lie algebra of $\mathrm{Aut}(\mathfrak{g})$ is the Lie subalgebra $\mathrm{Der}(\mathfrak{g})$ of all derivations of the Lie algebra $\mathfrak{g}$. Let $\mathrm{Int}(\mathfrak{g})$ be the identity component of $\mathrm{Aut}(\mathfrak{g})$. Then it is a connected compact Lie group. Its Lie algebra is $\mathrm{Der}(\mathfrak{g})$. Since the center $\mathfrak{z}$ of $\mathfrak{g}$ is $\{0\}$, the adjoint representation $\mathrm{ad} : \mathfrak{g} \longrightarrow \mathrm{Der}(\mathfrak{g})$ is injective. On the other hand, since every derivation of $\mathfrak{g}$ is inner by **??**, it is also, surjective. Therefore, ad is an isomorphism of the Lie algebra $\mathfrak{g}$ onto the Lie algebra of $\mathrm{Int}(\mathfrak{g})$. It follows that $\mathrm{Int}(\mathfrak{g})$ is a connected compact Lie group with Lie algebra isomorphic to $\mathfrak{g}$.

Let G be a connected compact Lie group with Lie algebra $\mathfrak{g}$. Then $Ad : G \longrightarrow \mathrm{Int}(\mathfrak{g})$ is a Lie group morphism, and its differential is $\mathrm{ad} : \mathfrak{g} \longrightarrow \mathrm{Der}(\mathfrak{g})$.

As we remarked above, ad is an isomorphism of Lie algebras. Therefore, by **??**, $\mathrm{Ad} : G \longrightarrow \mathrm{Int}(\mathfrak{g})$ is a covering map. By **??**, the kernel of Ad is the center of G. In particular, the center of $\mathrm{Int}(\mathfrak{g})$ is trivial, and $\mathrm{Int}(\mathfrak{g})$ is characterized up to an isomorphism by this property. The group $\mathrm{Int}(\mathfrak{g})$ is called the *adjoint group* of $\mathfrak{g}$.

Every other connected Lie group with Lie algebra $\mathfrak{g}$ is a covering group of the adjoint group $G_0 = \mathrm{Int}(\mathfrak{g})$. By Weyl's theorem, the universal covering group $\tilde{G}$ of the adjoint group is also compact. Therefore, it is a finite covering of G_0.

1.7. Root lattice and weight lattice. Let R be a root system in vector space V over k. Let R^* be the dual root system in V^*. We denote by $Q(R)$ the *root lattice* in V generated by R. If B is a basis of R, $Q(R)$ is the lattice of all integral linear combinations of vectors in B. Let $Q(R^*)$ be the root lattice of R^* in V^*. We denote by $P(R)$ the *dual lattice* of $Q(R^*)$, i.e., the set of all vectors in $v \in V$ such that $\lambda(v) \in \mathbb{Z}$ for all $\lambda \in Q(R^*)$. Clearly, v is in $P(R)$ if and only if $\alpha^*(v) \in \mathbb{Z}$ for all $\alpha^* \in R^*$. In particular, if B^* is a basis of R^*, v is in $P(R)$ if and only if $\alpha^*(v) \in \mathbb{Z}$ for all $\alpha^* \in B^*$.

The lattice $P(R)$ is called the *weight lattice* of R. Clearly, the root lattice $Q(R)$ is contained in the weight lattice $P(R)$.

1.7.1. Lemma. *Let R be a reduced root system. Let B be a basis of R. Denote by B^* the set of all roots α^* for $\alpha \in B$. Then B^* is a basis of R^*.*

Proof. We can assume that R is V is a real vector space. If we introduce on V an inner product invariant for $W(R)$, it defines a natural isomorphism of V^* with V. Under this isomorphism the dual root α^* of α corresponds to $\frac{2}{(\alpha|\alpha)}\alpha$. Therefore, the vectors α^*, $\alpha \in B$, form a basis of V^*.

Let $u \in V$ be a vector such that $(\alpha \mid v) > 0$ for all $\alpha \in B$. Then $(\beta \mid v) \neq 0$ for all $\beta \in R$; and $R^+ = \{\beta \in R \mid (\beta \mid v) > 0\}$ is the set of the positive roots corresponding to B.

Under our identification, we also have $(\beta^* \mid v) \neq 0$ for all $\beta^* \in R^*$. Therefore, $(R^*)^+ = \{\beta^* \in R^* \mid (\beta^* \mid v) > 0\} = \{\beta^* \in R^* \mid \beta \in R^*\}$ is a set of positive roots in R^* for a basis B' of R^*. Let $\alpha \in B$. Then α^* is in $(R^*)^+$ and we have $\alpha^* = \sum_{\beta^* \in B'} m_{\beta^*}\beta^*$. On the other hand, β^* is a positive multiple of β, hence $\beta^* = \sum_{\gamma \in B} c_{\beta,\gamma}\gamma^*$. It follows that

$$\alpha^* = \sum_{\beta^* \in B'} m_{\beta^*}\beta^* = \sum_{\beta^* \in B'} \sum_{\gamma \in B} m_{\beta^*} c_{\beta,\gamma}\gamma^* = \sum_{\gamma \in B} \left(\sum_{\beta^* \in B'} m_{\beta^*} c_{\beta,\gamma} \right) \gamma^*.$$

□

Let B be a basis of R. Denote by $(\omega_\alpha; \alpha^* \in B^*)$ the dual basis of the basis B^* of V^*, i.e., the weights ω_α such that $\beta^*(\omega_\alpha) = \delta_{\alpha,\beta}$ for all $\alpha, \beta \in R$. The weights ω_α are called the *fundamental weights* corresponding to B. By the above discussion, the fundamental weights of R generate the lattice $P(R)$.

1.7.2. Lemma. *Let R be a reduced root system and B its basis. Let R^+ be the corresponding set of positive roots and ρ its half-sum. Then we have*

$$\rho = \sum_{\alpha \in B} \omega_\alpha$$

for every root $\alpha \in B$.

Proof. By **??**, we know that $\alpha^*(\rho) = 1$. □

1.8. Classification of connected compact semisimple Lie groups. Let $\mathfrak{g}$ be a compact semisimple Lie algebra and $\mathfrak{t}$ a maximal abelian Lie subalgebra of $\mathfrak{g}$. Let $\mathfrak{g}_{\mathbb{C}}$ be the complexification of $\mathfrak{g}$. Denote by $\mathfrak{t}_{\mathbb{C}}$ the complex Lie subalgebra of $\mathfrak{g}_{\mathbb{C}}$ spanned by $\mathfrak{t}$. Then $\mathfrak{t}_{\mathbb{C}}$ is a Cartan subalgebra of $\mathfrak{g}_{\mathbb{C}}$. Let R be the root system of $(\mathfrak{g}_{\mathbb{C}}, \mathfrak{t}_{\mathbb{C}})$ in $\mathfrak{t}_{\mathbb{C}}^*$. By **??**, $\mathfrak{t}$ is spanned by iH_α, $\alpha \in R$, over $\mathbb{R}$.

Let G_0 be the adjoint group and T_0 the maximal torus in G_0 with Lie algebra $\mathfrak{t}$. Then $\exp : \mathfrak{t} \longrightarrow T$ is a Lie group morphism and $\operatorname{Ad}(\exp H) = e^{\operatorname{ad} H}$ for $H \in \mathfrak{t}$. Clearly, this action is diagonal in the basis consisting of root vectors, and $e^{\alpha(H)}$ is the action on the root subspace $\mathfrak{g}_{\mathbb{C},\alpha}$. It follows that H is in the kernel of exp if $e^{\alpha(H)} = 1$ for all $\alpha \in R$. This in turn implies that $\alpha(H) \in 2\pi i\mathbb{Z}$ for all $\alpha \in R$. Hence $H \in 2\pi i P(R^*)$.

Therefore, the lattice L_0 from **??** is equal to $2\pi i P(R^*)$.

We claim that the lattice $\tilde{L}$ is $2\pi i Q(R^*)$.

The proof consists of several steps. First we observe that by Weyl's theorem, the universal covering group $\tilde{G}$ of the adjoint group G_0 is compact. Therefore, by Peter-Weyl therem, for $g \in \tilde{G}$, such that $g \neq 1$, there exists an irreducible finite-dimensional representation π such that $\pi(g) \neq 1$.

Let $t \in \tilde{T}$. Then $t = \exp(H)$ for some $H \in i\mathfrak{h}_0$. By the above discussion, $t = 1$ if and only if $\pi(t) = 1$ for every irreducible finite-dimensional representation π of $\tilde{G}$. Since by **??** irreducible finite dimensional representations of $\tilde{G}$ correspond to irreducible finite-dimensional representations of $\mathfrak{g}$, $\pi(t) = 1$ if and only if $\exp \lambda(H) = 1$ for every weight $\lambda \in P(R)$. This in turn implies that $\lambda(H) \in 2\pi i\mathbb{Z}$. Hence, $H \in 2\pi i Q(R^*)$.

This implies that $L_0/\tilde{L}$ is isomorphic to $P(R^*)/Q(R^*)$.