*Dmitry Fuchs*

# NOTES ON LIE GROUPS AND LIE ALGEBRAS (261)

## 1. Lie groups and their Lie algebras.

### 1.1 Definition and examples of Lie groups.

**1.1.1. Definition.** A *Lie group* is a set furnished with two structures: that of a group and that of a smooth[1]) manifold. This two structures are supposed to be compatible in the sense that the maps

$$\mu: G \times G \to G,\ \mu(g,h) = gh,$$
$$\eta: G \to G,\ \eta(g) = g^{-1}$$

are smooth.

EXERCISE 1. Prove that the condition of the maps $\mu, \eta$ being smooth is equivalent to the condition of smoothness of a single map $G \times G \to G$, $(g,h) \mapsto g^{-1}h$.

The identity element of a Lie group is commonly denoted as $e$.

**1.1.2. Basic examples.** *A.* $\mathbb{R}^n$ with the usual operations of addition and additive inversion.

*B.* The circle $S^1$ with the usual operations of addition and inversions of angles.

*C.* The *general linear groups* $GL(n,\mathbb{R})$ and $GL(n,\mathbb{C})$ of invertible real or complex $n \times n$ matrices. Topologically, both groups are open sets in Euclidean spaces, whence the manifold structure.

*D.* The *special linear groups* $SL(n,\mathbb{R})$ and $SL(n,\mathbb{C})$ of real or complex $n \times n$ matrices with determinant 1. These groups are defined in the Euclidean spaces of all $n \times n$ matrices by the equation $\det X = 1$, and if $x_{ij}$ are entries of $X$, then $\dfrac{\partial \det}{\partial x_{ij}} = \pm \det X_{ij}$ where $X_{ij}$ is the $(n-1) \times (n-1)$ submatrix of $X$ complementary to $x_{ij}$. If $\det X \neq 0$, then $\det X_{ij} \neq 0$ for some $i, j$, and hence the equation $\det X = 1$ is non-degenerate. Hence $SL(n,\mathbb{R})$ and $SL(n,\mathbb{C})$ are submanifolds of Euclidean spaces.

*E.* The groups $O(n)$ and $U(n)$ of orthogonal or unitary matrices of order $n$.

EXERCISE 2. Prove that $O(n)$ and $U(n)$ are manifolds (submanifolds of Euclidean spaces) of dimensions, respectively, $\dfrac{n(n-1)}{2}$ and $n^2$.

*F.* The groups $SO(n)$ and $SU(n)$ of orthogonal or unitary matrices with determinant 1. (The dimensions are $\dfrac{n(n-1)}{2}$ and $n^2 - 1$.)

**1.1.3. Components and products.** The easiest way to obtain a new Lie group from a Lie group $G$ works in the case, when $G$ is disconnected: you can take the component

[1]) We will understand "smooth" as $\mathcal{C}^\infty$, although it is more common in the Lie theory to interprete "smooth" as real analytic.

$G_0$ of $e \in G$ with the same group operation as in $G$. Examples: $SO(n)$ is the component of $O(n)$; the component of $e \in GL(n,\mathbb{R})$ is the group $GL_+(n,\mathbb{R})$ of real $n \times n$ matrices with positive determinants.

EXERCISE 3. Prove that $G_0$ is a normal subgroup of $G$. (The quotient $G/G_0$ is called the group of components of $G$; in particular, for each of $O(n)$, $GL(n,\mathbb{R})$, the group of components id the cyclic group of order 2).

If $G_1$ and $G_2$ are Lie groups, then $G_1 \times G_2$ also has a natural structure of a Lie group. It is worth mentioning that it may happen that a Lie group $G$ is diffeomorphic to the product of two Lie groups, $G_1 \times G_2$, but not isomorphic to it. An example is contained in the next exercise.

EXERCISE 4. Prove that the following pairs of Lie groups are diffeomorphic, but not isomorphic. (a) $U(n)$ and $SU(n) \times S^1$ $(n \geq 2)$; (b) $O(n)$ and $SO(n) \times \mathbb{Z}_2$ $(n \geq 2)$; (c) $GL(n,\mathbb{R})$ and $O(n) \times \mathbb{R}^k$, $k = \dfrac{n(n+1)}{2}$ $(n \geq 2)$.

**1.1.4. Coverings.** Coverings are studied in topology. Understanding of the construction below requires some familiarity with this topological theory; to make the life easier for a reader not possessing this knowledge, I will give a brief description of the definitions and results needed in a footnote[2]).

---

[2]) A covering $(Y, p, X)$ is a triple consisting of two path connected topological spaces, $X$ and $Y$, and a continuous map $p\colon Y \to X$ such that every point $x \in X$ possesses a neighborhood $U \subset X$ such that the inverse image $p^{-1}(U) \subset Y$ is a disjoint union $\bigcup_\alpha U_\alpha$ of open sets in $Y$ such that for every $\alpha$, the restriction $p|_{U_\alpha}$ is a homeomorphism of $U_\alpha$ onto $U$.

Technically, the main result of the covering theory is the following Lemma of Lifting Paths: If $s\colon [0,1] \to X$ is a path with $s(0) = x_0$ and $y_0 \in Y$ is a point with $p(y_0) = x_0$ then there exists a unique path $\widetilde{s}\colon [0,1] \to Y$ such that $\widetilde{s}(0) = y_0$ and $p \circ \widetilde{s} = s$. The lifting operation is homotopy invariant: if paths $s_1, s_2$ are homotopic (by definition, a homotopy of paths is endpoints fixed), then the lifted paths $\widetilde{s}_1, \widetilde{s}_2$ are also homotopic; in particular, $\widetilde{s}_1(1) = \widetilde{s}_2(1)$.

If $s$ is a loop (that is, $s(1)$ is also $x_0$), then $\widetilde{s}$ may be also a loop (in which case we say that $s$ is covered by a loop starting at $y_0$) or not a loop (that is, $\widetilde{s}(1) \in p^{-1}(x_0) - y_0$); this property of the loop $s$ may depend on $y_0$. The property of being covered by a loop is also homotopy invariant: two homotopic loops possess it simultanously.

Consider the diagram

$$\begin{tikzcd}
y_0 \in Y \arrow[r, dashed, "F"] \arrow[d, "p"] & Y' \ni y_0' \arrow[d, "p'"] \\
x_0 \in X \arrow[r, "f"] & X' \ni x_0'
\end{tikzcd}$$

in which $p\colon Y \to X$ and $p'\colon Y' \to X'$ are coverings and $f\colon X \to X'$ is a continuous map. Assume that $p(y_0) = x_0$, $p'(y_0') = x_0'$, $f(x_0) = x_0'$. We want to construct a continuous map $F\colon Y \to Y'$ such that $F(y_0) = y_0'$ and $p' \circ F = f \circ p$. Claim: *if $F$ with this properties exists,* Now let $G$ be a Lie group, and let $(\widetilde{G}, p, G)$ be a covering. Fix an $\widetilde{e} \in \widetilde{G}$ such that $p(\widetilde{e}) = e$. Then $\widetilde{G}$ possesses a unique structure of a Lie group with the identity element $\widetilde{e}$ such that $p\colon \widetilde{G} \to G$ is a group homomorphism. To prove this, we apply the proposition in the footnote [2]) to the diagrams

$$\begin{tikzcd} (\widetilde{e},\widetilde{e}) \in \widetilde{G} \times \widetilde{G} \arrow[r, dashed, "\widetilde{\mu}"] \arrow[d, "p \times p"] & \widetilde{G} \ni \widetilde{e} \arrow[d, "p"] \\ (e,e) \in G \times G \arrow[r, "\mu"] & G \ni e \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} \widetilde{e} \in \widetilde{G} \arrow[r, dashed, "\widetilde{\eta}"] \arrow[d, "p"] & \widetilde{G} \ni \widetilde{e} \arrow[d, "p"] \\ e \in G \arrow[r, "\eta"] & G \ni e \end{tikzcd}$$

Let us check for these diagrams the condition concerning to covering loops by loops. A loop $s\colon [0,1] \to G \times G$ is the same as pair of loops in $G$: $s(t) = (s_1(t), s_2(t))$. This loop is covered by a loop starting at $(\widetilde{e}, \widetilde{e})$ if and only if each of the loops $s_1, s_2\colon [0,1] \to G$ is covered by a loop starting at $\widetilde{e}$. The composition $\mu \circ s$ is the loop $t \mapsto s_1(t)s_2(t)$. Apply to the last loop the homotopy which compresses the argument at $s_1$ to the interval $[0, 1/2]$ and the argument at $s_2$ to the interval $[1/2, 1]$. The resulting loop (homotopic to $\mu \circ s$) is

$$t \mapsto \begin{cases} s_1(2t)s_2(0) = s_1(2t) & \text{for } t \le 1/2 \\ s_1(1)s_2(2t-1) = s_2(2t-1) & \text{for } t \ge 1/2, \end{cases}$$

that is, the loop obtained by consecutive passing the loops $s_1$ and $s_2$. Certainly, this loop is covered by a loop starting are $\widetilde{e}$: first, we lift $s_1$, the endpoint of the lifted loop will be again $\widetilde{E}$, and then lift $s_2$. Thus, we obtain the map $\widetilde{\mu}\colon \widetilde{G} \times \widetilde{G} \to \widetilde{G}$.

For the second diagram, we assume that a loop $s\colon [0,1] \to G$ is covered by a loop starting at $\widetilde{e}$. We state that the loop $\eta \circ s$ is homotopic to a loop $t \mapsto s(1-t)$, that is, the same loop $s$ passed in the opposite direction. Indeed, our previous construction shows that the constant loop $t \mapsto s(t)\eta \circ s(t) = e$ is homotopic to a loop obtained by coinsecutive passing loops $s$ and $\eta \circ s$, which means precisely that the loop $\eta \circ s$ is homotopic to the loop homotopically inverse to $s$, that is, to the loop $t \mapsto s(1-t)$. The latter is covered by a loop simultaneously with $s$.

It is obvious that the maps $\widetilde{\mu}$ and $\widetilde{\eta}$ are smooth (since $p$ and $p \times p$ are local diffeomorphisms). The checking of the group axioms for $\widetilde{G}$ is based on the uniqueness property in the footnote [2]) proposition. For example, the diagrams

$$\begin{tikzcd} ((\widetilde{e},\widetilde{e}),\widetilde{e}) & \widetilde{e} \\ (\widetilde{G} \times \widetilde{G}) \times \widetilde{G} \arrow[u, phantom, "\Cap"] \arrow[r, "\widetilde{\mu} \circ (\widetilde{\mu} \times \mathrm{id})"] \arrow[d, "(p \times p) \times p"] & \widetilde{G} \arrow[u, phantom, "\Cap"] \arrow[d, "p"] \\ (G \times G) \times G \arrow[r, "\mu \circ (\mu \times \mathrm{id})"] & G \\ ((e,e),e) \arrow[u, phantom, "\Cup"] & e \arrow[u, phantom, "\Cup"] \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} (\widetilde{e},(\widetilde{e},\widetilde{e})) & \widetilde{e} \\ \widetilde{G} \times (\widetilde{G} \times \widetilde{G}) \arrow[u, phantom, "\Cap"] \arrow[r, "\widetilde{\mu} \circ (\mathrm{id} \times \widetilde{\mu})"] \arrow[d, "p \times (p \times p)"] & \widetilde{G} \arrow[u, phantom, "\Cap"] \arrow[d, "p"] \\ G \times (G \times G) \arrow[r, "\mu \circ (\mathrm{id} \times \mu)"] & G \\ (e,(e,e)) \arrow[u, phantom, "\Cup"] & e \arrow[u, phantom, "\Cup"] \end{tikzcd}$$

---

*it is unique; $F$ exists, if and only if $f$ takes every loop of $X$ starting at $x_0$ and covered by a loop starting at $y_0$ into a loop of $X'$ covered by a loop starting at $y'_0$.*

are commutative, and since the multiplication in $G$ is associative, that is, $\mu \circ (\mu \times \mathrm{id}) = \mu \circ (\mathrm{id} \times \mu)$, we have also $\widetilde{\mu} \circ (\widetilde{\mu} \times \mathrm{id}) = \widetilde{\mu} \circ (\mathrm{id} \times \widetilde{\mu})$, which means that the multiplication in $\widetilde{G}$ is also associative. The other group axioms for $\widetilde{G}$ are checked in a similar way.

There are some obvious examples for the covering construction. For example, there is a well known covering $(\mathbb{R}, p, S^1)$. Since $S^1$ is a Lie group, we obtain a Lie group structure for $\mathbb{R}$, which is not interesting, since it coincides with the usual Lie group structure on $\mathbb{R}$. A more interesting example is provided by the two-fold covering of $SO(n), n > 2$ ("two-fold" means that the inverse image of every point of $SO(n)$ consists of two points), whose existence follows from the classification theory of coverings in topology. The covering group is called the spinor group and is denoted as $Spin(n, \mathbb{R})$; it turns out that $Spin(3, \mathbb{R}) \cong SU(2)$, but all the other spinor groups are not among our basic examples (Section 1.1.2)[3]).

Notice in conclusion that if $(\widetilde{G}, p, G)$ is a covering of a Lie group by a Lie group, then $\operatorname{Ker} p$ is a central, in particular, Abelian, subgroup of $\widetilde{G}$. This follows from the following, more general statement.

Proposition. *Every discrete normal subgroup of a connected Lie group is central.*

*Proof.* Let $H$ and $\Gamma$ be a Lie group and a subgroup as in Proposition. Then, for every $\gamma \in \Gamma$, the set $H^{-1}\gamma H$ is contained in $\Gamma$ (because $\Gamma$ is normal) and is connected (because $H$ is connected). Since $\Gamma$ is discrete, its connected subset must be one-point, hence $H^{-1}\gamma H = \gamma$, hence $h^{-1}\gamma h = \gamma$, that is, $\gamma h = h\gamma$ for every $h \in H$.

**1.1.5. Lie homomorphisms and Lie subgroups.** A homomorphism of a Lie group $G$ into a Lie group $H$ is a map $G \to H$ which is simultaneously a smooth map and a group homomorphism. A Lie subgroup of a Lie group $G$ is a closed subset of $G$ which is simultaneously a submanifold and a subgroup of $G$. There is *Cartan's theorem* which state that every subgroup of a Lie group $G$ which is closed (in the topology of $G$) is a Lie subgroup of $G$; in particular, the kernel of any Lie group homomorphism $G \to H$ is a Lie subgroup of $G$. We will prove Cartan's theorem in Section 1.2.5.

Exercise 5. Prove that the kernel of a Lie homomorphism $G \to H$ is a Lie subgroup of $G$. (Later on, we will use this fact, so the reader has to take care of the proof of it, or, at very least, to believe that it is true.)

On the contrary, the image of a Lie homomorphism $G \to H$ is not, in general, a Lie subgroup of $H$, in particular, because it does not have to be closed (later, in Section 1.3.4.B, we will show that the image of a Lie homomorphism belongs to a wider class of "virtual Lie subgroups").

The "basic examples" of Lie groups from Section 1.1.2 provide a huge amount of examples of Lie subgroups. Here are the most important of them (we restrict ourselves to the compact case):

$$SO(n) \subset O(n),\ SU(n) \subset U(n),\ SO(n) \subset SU(n),\ U(n) \subset SO(2n);$$

for $m < n$,

$$SO(m) \subset SO(n),\ O(m) \subset O(n),\ SU(m) \subset SU(n),\ U(m) \subset U(n);$$

[3]) There are also "complex spinor groups" $Spin(n, \mathbb{C})$, but they are not defined as coverings.

also

$$O(m) \times O(n) \subset O(m+n)$$

(this subgroup of $O(m+n)$ consists of block diagonal matrices with two orthogonal diagonal blocks of the sizes $m \times m$ and $n \times n$), and similarly for $O$, $SU$, and $U$. More generally,

$$O(n_1) \times \ldots \times O(n_k) \subset O(n_1 + \ldots + n_k),$$

and similarly for $O$, $SU$, and $U$.

**1.1.6. Homogeneous spaces.** Let $H$ be an $m$-dimensional Lie subgroup of an $n$-dimensional Lie group $G$. We are going to show that the set $G/H$ of left cosets of $H$ has a natural structure of a manifold of dimension $n - m$.

Let $W$ be a transverse to $H$ $(n-m)$-dimensional submanifold of a small neighborhood of $e \in G$. Let us show that if $W$ is sufficiently small, then $\mu\colon W \times H \to G$, $(w, h) \mapsto wh$, is an embedding. First notice that the differential $d_{(e,h)}\mu$ is non-degenerate, which implies that $\mu$ is a local embedding at every point of $e \times H$. Hence, it is sufficient to prove that, for a sufficiently small $W$, $\mu$ is one-to-one. Otherwise, there exist sequences $(w_i, h_i), (w_i', h_i') \in W \times H$ such that $(w_i, h_i) \neq (w_i', h_i')$, $\lim\limits_{i\to\infty} w_i = \lim\limits_{i\to\infty} w_i' = e$, and $w_i h_i = w_i' h_i'$. But in this case, $k_i = h_i' h_i^{-1} = {w_i'}^{-1} w_i$ has the limit $e$, and the equality $w_i e = w_i' k_i$ contradicts to the fact that $\mu$ is an embedding in some neighborhood of $(e, e) \in W \times H$.

Assume that $W$ is this small. Then the map $W = W \times e \xrightarrow{\mu} G \xrightarrow{\text{proj.}} G/H$ is one to one, and so is the map $W = W \times e \xrightarrow{\mu} G \xrightarrow{g\cdot} G \xrightarrow{\text{proj.}} G/H$ for an arbitrary $g \in G$. These maps form an atlas of the set $G/H$; this atlas can be made countable, if we restrict it for $g$ from some countable dense subset of $G$.

Exercise 6. Prove that these charts are compatible and satisfy the Hausdorff axiom.

Thus, $G/H$ is a manifold; it is clear that the transitive action of $G$ in $G/H$ is a smooth action by difeomorphisms and that $H$ is the isotropy subgroup of $\{H\} \in G/H$. The manifold $G/H$ with this transitive action of $G$ is called a *homogeneous space* (of $G$).

The other approach to homogeneous spaces is possible: we consider a smooth manifold $M$ with a smooth transitive action $G \times M \to M$ by diffeomorphisms. For every fixed $x_0 \in M$, the isotropy subgroup $H = \{g \in G \mid gx_0 = x_0\}$ is a Lie subgroup of $G$, and we can identify $M$ with the manifold $G/H$ constructed above.

Exercise 7. Construct diffeomorphisms $SO(n+1)/SO(n) = S^n$, $SU(n+1)/SU(n) = S^{2n+1}$, $SO(n+1)/O(n) = \mathbb{R}P^n$, $SU(n+1)/U(n) = \mathbb{C}P^n$ (for the last two diffeomorphisms, you will need also to construct the appropriate embeddings $O(n) \to SO(n+1)$ and $U(n) \to SU(n+1)$), $O(m+n)/(O(m) \times O(n)) = G(m+n, n)$ (the Grassman manifold).

**1.2. Tangent vectors and vector fields.** The tangent space $T_eG$ of a Lie group $G$ at the identity element $e \in G$ plays a very special role in the Lie theory. As we will see in Section 1.3, it possesses a specific structure of a *Lie algebra*, and the study of Lie groups in many cases is reduced to a purely algebraic consideration of Lie algebras. Let us first discuss some relations between the properties of $G$ and $T_eG$ not related, at least explicitly, to this algebraic structure.

**1.2.1. Multiplication in $G$ and addition in $T_eG$.** PROPOSITION. *Let $\gamma_1$ and $\gamma_2$ be two parametrized curves in $G$ with $\gamma_1(0) = \gamma_2(0) = e$. Let $\gamma(t) = \gamma_1(t)\gamma_2(t)$. Then $\dot\gamma(0) = \dot\gamma_1(0) + \dot\gamma_2(0)$.*

*Proof.* Consider the curves in $G \times G$:

$$\widetilde{\gamma}_1(t) = (\gamma_1(t), e),\ \widetilde{\gamma}_2(t) = (e, \gamma_2(t)),\ \widetilde{\gamma}(t) = (\gamma_1(t), \gamma_2(t)).$$

Obviously, $\dot{\widetilde{\gamma}}(0) = \dot{\widetilde{\gamma}}_1(0) + \dot{\widetilde{\gamma}}_2(0)$, and $\mu \circ \widetilde{\gamma}_1 = \gamma_1, \mu \circ \widetilde{\gamma}_2 = \gamma_2, \mu \circ \widetilde{\gamma} = \gamma$. Since $d\mu: T_{(e,e)}(G \times G) \to T_eG$ is a linear map, Proposition follows.

**1.2.2. Left and right invariant vector fields.** Let $G$ be a Lie group, and let $g \in G$. There arise three transformations of $G$:

$$\begin{aligned} \text{left translation } \lambda_g: G \to G,&\ \ \lambda_g(h) = gh, \\ \text{right translation } \rho_g: G \to G,&\ \ \rho_g(h) = hg, \\ \text{conjugation } \alpha_g: G \to G,&\ \ \alpha_g(h) = ghg^{-1} \text{ (or } \alpha_{\mathrm{g}} = \rho_{\mathrm{g}^{-1}} \circ \lambda_{\mathrm{g}}). \end{aligned}$$

Notice that $\lambda_g \circ \rho_{g'} = \rho_{g'} \circ \lambda_g$. Also $\lambda_g \circ \lambda_{g'} = \lambda_{gg'}$, $\rho_g \circ \rho'_g = \rho_{g'g}$, $\alpha_g \circ \alpha_{g'} = \alpha_{gg'}$, and $\lambda_e = \rho_e = \alpha_e = \mathrm{id}, \lambda_{g^{-1}} = \lambda_g^{-1}$, $\rho_{g^{-1}} = \rho_g^{-1}$, $\alpha_{g^{-1}} = \alpha_g^{-1}$. In other words, $\lambda$ comprises a left action of $G$ on itself by difeomorphisms, $\rho$ determines a right action of $G$ on itself by difeomorphisms, and $\alpha$ produces a left action of $G$ on itself by Lie group automorphisms. The first two actions are free and transitive.

The transformations $\alpha_g$ are group automorphisms; they will be very important for us later, but now we will concentrate our attention on the left and right translations $\lambda_g$ and $\rho_g$. Obviously, they are diffeomorphisms.

Any diffeomorhism $\varphi: M \to M$ of any manifold $M$ acts on vector fields: if $X \in \operatorname{Vect} M$ is a vector field on $M$, then, by definition, $\varphi^*X(f) = X(f \circ \varphi)$ $f \in \mathcal{C}^\infty M$; equivalently, for a $p \in M$, $(\varphi^*X)_{\varphi(p)} = d_p\varphi(X_p)$. A vector field $X$ is called *$\varphi$-invariant* (or *invariant with respect to* $\varphi$), if $\varphi^*X = X$.

A vector field on $G$ is called left (right) invariant, if it is invariant with respect to all left (right) translations. Both left and right invariant vector fields on $G$ form subspaces of the vector space $\operatorname{Vect} G$ invariant with respect the commutator operation (that is, if vector fields $X, Y \in \operatorname{Vect} G$ are left (right) invariant, then so is the commutator $[X, Y]$). The notations:$\operatorname{Vect}_{\ell-inv} G$ and $\operatorname{Vect}_{r-inv} G$.

There are *canonical isomorphisms* $\operatorname{Vect}_{\ell-inv} G \cong T_eG$ *and* $\operatorname{Vect}_{r-inv} G \cong T_eG$. The constructions are as follows. For a $\xi \in T_eG$, define vector fields $L_\xi, R_\xi \in \operatorname{Vect} G$ by the formulas

$$(L_\xi)_g = d_e\lambda_g(\xi),\ (R_xi)_g = d_e\rho_g(\xi) \text{ or } (L_\xi f)(g) = \xi(f \circ \lambda_{g^{-1}}),\ (R_\xi f)(g) = \xi(f \circ \rho_{g^{-1}})$$

(in other words, we define the vector fields $L_\xi$ and $R_\xi$ by spreading the vector $\xi$ to the whole $G$ by left or right translations). It is clear that $L_\xi$ is left invariant and $R_\xi$ is right invariant. Indeed, to prove the invariance of $L_\xi$ with respect to $\lambda_g$, we need to check that for every $h \in G$, $d_h\lambda_g(L_\xi(h)) = L_\xi(\lambda_g(h))$; but this follows from the definitions:

$$d_h\lambda_g(L_\xi(h)) = d_h\lambda_g(d_e\lambda_h(\xi)) = d_e(\lambda_g \circ \lambda_h)(\xi) = d_e\lambda_{gh}(\xi) = L_\xi(gh) = L_\xi(\lambda_g(h));$$

Thus we have a map $T_eG \to \mathrm{Vect}_{\ell-inv}\, G$, $\xi \mapsto L_\xi$. Since $(L_\xi)_e = \xi$, this map is one-to-one. If $X$ is a left invariant vector field, then, obviously, $X = L_{X_e}$, so this map is onto. We obtain the promised isomorphism $T_eG \cong \mathrm{Vect}_{\ell-inv}\, G$. The proof for $R_\xi$ is the same.

**1.2.3. One-parameter subgroups.** A *one-parameter subgroup* of a Lie group $G$ is, by definition, a Lie homomorphism $\mathbb{R} \to G$. Obviously, for a continuous homomorphism $\gamma\colon \mathbb{R} \to G$, there are three possibilities for $\operatorname{Ker}\gamma$: it can be 0, $\mathbb{R}$, or $\mathbb{Z}a$ for some $a \in \mathbb{R}$. (Indeed, if $\operatorname{Ker}\gamma$ is not descrete, then for every $\varepsilon > 0$ there exists a $t \in \operatorname{Ker}\gamma$ with $0 < t < \varepsilon$; but in this case $nt \in \operatorname{Ker}\gamma$ for all $n \in \mathbb{Z}$, which shows that $\operatorname{Ker}\gamma$ is dense in $\mathbb{R}$; however for a continuous $\gamma$, $\operatorname{Ker}\gamma$ must be closed, so in this case $\operatorname{Ker}\gamma = \mathbb{R}$. If $\operatorname{Ker}\gamma$ is discrete, then either $\operatorname{Ker}\gamma = 0$, or $\operatorname{Ker}\gamma$ contains an $a = \inf(\operatorname{Ker}\gamma \cap \mathbb{R}_{>0})$; in the last case, $\operatorname{Ker}\gamma = \mathbb{Z}a$.) Thus, $\operatorname{Im}\gamma$ is either $e$, or a closed curve (in which case $\gamma$ is periodic), or an embedded curve. Notice that in the last two cases, $\gamma$ must be an immersion: for every $t \in R$, $\gamma(t+u) = \lambda_{\gamma(t)}\gamma(u)$ (as well as $\rho_{\gamma(t)}\gamma(u)$), so the velocity vectors of the parametrized curve $\gamma$ are taken into each other by isomorphisms $d_{\gamma(t)}\lambda_{\gamma(u)}$, and are all zeroes or non-zeroes simultaneously.

Notice that the classical term "one-parameter subgroup" does not fully correspond to the notion of a Lie subgroup (or even of a subgroup). If $\gamma\colon \mathbb{R} \to G$ is a one parameter subgroup of $G$, then the image $\gamma(\mathbb{R})$ is a subgroup of $G$ (in the algebraic sense), but the parameter change $t \to kt$, $k \neq 1$ changes the one-parameter subgroup, but not the image. On the other hand, if $\gamma$ is one-to-one, then the image of $\gamma$ can be not closed, even dense; the simplest example is $\mathbb{R} \xrightarrow{\beta} \mathbb{R}^2 \xrightarrow{\pi} S^1 \times S^1$ where $\beta(t) = (at, bt)$ with irrational $b/a$ and $\pi$ is the projection $\mathbb{R}^2 \to \mathbb{R}^2/\mathbb{Z}^2 = S^1 \times S^1$.

Theorem. *For every $\xi \in T_eG$ there exists a unique one-parameter subgroup $\gamma\colon \mathbb{R} \to G$ with $\dot\gamma(0) = \xi$.*

Thus, there arises a one-to-one correspondence between one-parameter subgroups of $G$ and tangent vectors to $G$ at $e$; we will denote the one-parameter subgroup of $G$ corresponding to a $\xi \in T_eG$ by $\gamma_\xi$.

*Proof of Theorem.* The only one-parameter subgroup with $\dot\gamma(0) = 0$ is the constant map $\gamma(\mathbb{R}) = e$. Let $\xi \in T_eG$ be non-zero. Consider the left invariant vector field $L_\xi$. Since $(L_\xi)_e \neq 0$, there exists a unique (up to the choice of $\varepsilon, \varepsilon' > 0$) integral curve $\gamma\colon (-\varepsilon', \varepsilon) \to G$ of $L_\xi$ with $\gamma(0) = e$. Since $L_\xi$ is left invariant, the curve $\lambda_g \circ \gamma\colon (-\varepsilon', \varepsilon) \to G$ for $g \in G$ is also an integral curve of $L_\xi$. Thus, for every $u \in (\varepsilon', \varepsilon)$ we have two integral curves of $L_\xi$: $L_{\gamma(u)\circ\gamma}$ and $(-\varepsilon' - u, \varepsilon - u) \to G$, $t \mapsto \gamma(u+t)$, both taking 0 to $\gamma(u)$. By uniqueness, these two curves coincide where they are both defined, which shows that $\gamma(u+t) = \gamma(u)\gamma(t)$ (if $t, u, t+u \in (-\varepsilon', \varepsilon)$). Also, these two curves merge into a curve defined on the union $(-\varepsilon', \varepsilon) \cup (-\varepsilon' - u, \varepsilon - u)$, and using this trick we can extend our the domain of our curve to an arbitrarily large interval, eventually to the whole line $\mathbb{R}$. In this way we define a one parameter subgroup of $G$ with the required property. On the other hand, if $\gamma$ is a one parameter group of $G$, then $\lambda_{\gamma(u)}\gamma(t) = \gamma(u+t)$ which implies that $d\lambda_{\gamma(u)}\dot\gamma(t) = \dot\gamma(u+t)$, that is, $\gamma$ is an integral curve of the vector field $L_\xi$ with $\xi = \dot\gamma(0)$. Hence, the uniqueness of the integral curve implies the uniqueness of a one-parameter subgroup $\gamma$ with a given $\dot\gamma(0)$.

Notice that in this proof we could use the vector fields $R_\xi$ instead of $L_\xi$; this means that the vector fields $L_\xi$ and $R_\xi$ have the same integral curves starting at $e$; in general they have different integral curves starting at other points: these are left and right cosets of the same one-parameter subgroup.

**1.2.4. Exponential maps.** The exponential map $\exp\colon T_eG \to G$ for a Lie group $G$ is defined by the formula

$$\exp(\xi) = \gamma_\xi(1)$$

(in particular, $\exp(0) = e$). Since $\gamma_\xi(t) = \gamma_{t\xi}(1)$, the exponential map takes straight lines passing through $0 \in T_eG$ to one-parameter subgroups of $G$. Here are two most obvious properties of exp. First, the differential

$$d_0 \exp\colon T_0(T_eG) = T_eG \to T_eG \text{ is } \mathrm{id}_{T_eG}\,.$$

In particular, exp is a local diffeomorphism, that is, it maps a neighborhood of 0 in $T_eG$ diffeomorphically onto a neighborhood of $e$ in $G$. Second, for a Lie homomorphism $f\colon G \to H$, the diagram

$$\begin{array}{ccc} T_eG & \xrightarrow{d_ef} & T_eH \\ \Big\downarrow{\scriptstyle\exp} & & \Big\downarrow{\scriptstyle\exp} \\ G & \xrightarrow{f} & H \end{array}$$

is commutative.

Examples. 1. $G = GL(n,\mathbb{R})$. Since $GL(n,\mathbb{R})$ is an open set in the space $\mathrm{Mat}_n(\mathbb{R})$ of real $n \times n$ matrices, we can identify $T_IGL(n,\mathbb{R})$ with $\mathrm{Mat}_n(\mathbb{R})$. There are explicitly constructed one-parameter subgroups of $GL(n,\mathbb{R})$. Namely, for $A \in \mathrm{Mat}_n(\mathbb{R})$ put

$$\exp(A) = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \frac{A^4}{4!} + \dots$$

It is obvious that the series converges for every $A$. The equality $\exp(A)\exp(B) = \exp(A+B)$ is true and is proved in the usual way under the condition that $A$ and $B$ commute. In particular, $\exp(tA)\exp(uA) = \exp((t+u)A)$, so

$$\gamma_A\colon \mathrm{Mat}_n(\mathbb{R}) \to GL(n,\mathbb{R}),\ \gamma_A(t) = \exp(tA)$$

is a one-parameter subgroup of $GL(n,\mathbb{R})$, and $\dot\gamma_A(0) = A$. Thus, $GL(n,\mathbb{R})$ has no other one parameter subgroups, and the exponential map $T_IGL(n,\mathbb{R}) \to GL(n,\mathbb{R})$ acts by the formula $A \mapsto \exp(A)$. The same formula describes the exponential map for every Lie subgroup of $G$ (because of the commutativity of the diagram above). These facts provide an explanation for the term *exponential.*

2. $G = SL(n,\mathbb{R})$. A matrix $A$ belongs to $T_ISL(n,\mathbb{R})$ if and only if the line $\{tA\}$ is tangent to $SL(n,\mathbb{R})$, that is, if $\frac{d}{dt}\det(I+tA)|_{t=0} = 0$. If $\lambda_1,\dots,\lambda_n$ are eigenvalues of $A$, then $\det(I+tA) = (1+t\lambda_1)\dots(1+t\lambda_n)$. Hence, $\frac{d}{dt}\det(I+tA)|_{t=0} = \lambda_1+\dots+\lambda_n = \mathrm{Tr}\,A$. Thus,

$$T_ISL(n,\mathbb{R}) = \{A \in \mathrm{Mat}_n(\mathbb{R}) \mid \mathrm{Tr}\,A = 0\}.$$

(This means, in particular, that $\det \exp(A) = 1$ if and only if $\operatorname{Tr} A = 0$.)

EXERCISE 8. Prove that the map $\exp\colon T_I SL(2,\mathbb{R}) \to SL(2,\mathbb{R})$ is not onto; describe its image.

3. $G = O(n)$. A matrix $A$ belongs to $T_I O(n)$ if and only if the line $\{tA\}$ is tangent to $O(n)$, that is, if $\dfrac{d}{dt}((I+tA)^{\mathrm{t}} - (I+tA)^{-1}) = 0$. But $(I+tA)^{-1} = I - tA + t^2A^2 - \dots$, so derivative we want to compute is $A^t + A$ and

$$T_I O(n) = \{A \in \mathrm{Mat}_n(\mathbb{R}) \mid A \text{ is skew symmetric}\}.$$

(In particular, a matrix $\exp A$ is orthogonal if and only if the matrix $A$ is skew-symmetric.)

4. *The complex case* is similar to the real case. Namely, $T_I GL(n,\mathbb{C}) = \mathrm{Mat}_n(\mathbb{C})$, $T_I SL(n,\mathbb{C})$ is the space of matrices with zero trace, $T_I U(n)$ is the space of *skew-Hermitian* matrices, that is, of complex matrices $A$ such that $A^{-1} = \overline{A}^{\mathrm{t}}$ (it is a *real* subspace of a complex vector space $\mathrm{Mat}_n(\mathbb{C})$), and $T_I SU(n)$ is the space of skew-Hermitian matrices with zero trace. There is also a Lie group $O(n,\mathbb{C})$ of *complex orthogonal matrices* satisfying the equation $A^{-1} = A^{\mathrm{t}}$. The space $T_I O(n,\mathbb{C})$ is the (complex) vector space of complex skew-symmetric $n \times n$ matrices.

EXERCISES 9. Prove that the map $\exp\colon T_I SL(2,\mathbb{C}) \to SL(2,\mathbb{C})$ is onto.

10. Prove that any neighborhood of $e$ in a connected Lie group generates this group (in the algebraic sense).

**1.2.5. Closed subgroups of a Lie group are Lie subgroups.** THEOREM (E. Cartan). *Let $H$ be a subgroup of a Lie group $G$ closed with respect to topology of $G$. Then $H$ is a Lie subgroup of $G$.*

*Proof.* It is sufficient to prove that $H$ is a submanifold of $G$ (the maps $\mu\colon H \times H \to H$ and $\eta\colon H \to H$ are restrictions of the similar maps for $G$, and they will be automatically smooth). This means that for every $h \in H$ there exists a chart $\varphi\colon U \to G$ such that $\varphi(U) \ni h$ and $\varphi^{-1}(H) = U \cap \mathbb{R}^m$ where $m = \dim H$. It is sufficient to prove this for $h = e$; indeed, if $\varphi\colon U \to G$ is a chart with properties required for $h = e$, then the chart $\lambda_h \circ \varphi\colon U \to G$ satisfies the conditions posted for an arbitrary $h \in H$ (where $\lambda_h$ is a left translation of $G$). And for this it is sufficient to prove that there exist a subspace $W \subset T_eG$ and a neighborhood $U$ of $e$ in $T_eG$ such that $U \cap \exp^{-1}(H) = U \cap W$. The proof consists of five steps.

*Step 1.* Fix an arbitrary Euclidean metric in $T_eG$, and let $S = \{\xi \in T_eG \mid \|\xi\| = 1\}$ be the "sphere". A vector $\xi \in S$ is called a *unit virtual tangent vector* to $H$, if there exists a sequence $\{h_i\}$ in $\exp(G) \cap (H - e)$ such that $\lim\limits_{i\to\infty} h_i = e$ and $\lim\limits_{i\to\infty} \dfrac{\exp^{-1}(h_i)}{\|\exp^{-1}(h_i)\|} = \xi$. A vector of the form $\lambda\xi$ where $\lambda \in \mathbb{R}$ and $\xi$ is a unit virtual tangent vector to $H$ is called a *virtual tangent vector* to $H$. Let $W \subset T_eG$ be the set of all virtual tangent vectors to $H$.

*Step 2.* If $\gamma\colon (-\varepsilon', \varepsilon) \to G$ is a smooth curve such that $\gamma(0) = e$ and $\gamma(-\varepsilon', \varepsilon) \subset H$, then $\dot\gamma(0) \in W$. Indeed, if $0 \in W$, so we can assume that $\dot\gamma(0) \neq 0$. In this case, we can choose a sequence $\{t_i\}$ in $(0, \varepsilon)$ such that $h_i = \gamma(t_i) \neq e$ and $\lim\limits_{i\to\infty} h_i = e$. Obviously,

$\lim\limits_{i\to\infty} \dfrac{\exp^{-1}(h_i)}{\|\exp^{-1}(h_i)\|}$ is a unit vector tangent to the curve $\gamma$. Thus this vector, as well as $\dot\gamma(0)$ is virtual tangent to $H$.

*Step 3.* Next, we will prove that if $\xi$ is a virtual tangent vector to $H$, then the whole one-parameter subgroup $\gamma_\xi$ is contained in $H$. If $\xi = 0$, then $\gamma_\xi$ is just $e$, so we need to consider the case when $\xi \neq 0$. We can assume that $\|\xi\| = 1$. Then $\xi = \lim\limits_{i\to\infty} \xi_i$, $\xi_i = \dfrac{\eta_i}{\|\eta_i\|}$, $h_i = \exp(\eta_i) \in H - e$, and $\lim\limits_{i\to\infty} h_i = e$. The one parameter subgroups $\gamma_{\xi_i}$ converge to $\gamma_\xi$ in the sense that $\lim\limits_{i\to\infty} \gamma_{\xi_i}(t_i) = \gamma_\xi(t)$, if $\lim\limits_{i\to\infty} t_i = t$. We want to prove now that $\gamma_\xi(t) \in H$ for every $t$. Let $h_i = \gamma_{\xi_i}(t_i)$. Then $\lim\limits_{i\to\infty} t_i = 0$, and we can find integers $n_i$ such that $\lim\limits_{i\to\infty} n_i t_i = t$. Hence, $\lim\limits_{i\to\infty} h_i^{n_i} = \lim\limits_{i\to\infty} \gamma_{\xi_i}(n_i t_i) = \gamma_\xi(t)$, and $\gamma_\xi(t) \in H$, because $h_i^{n_i} \in H$ and $H$ is closed.

*Step 4.* Now we can prove that $W$ is a subspace of the vector space $T_eG$. Since $W$, by construction, is closed with respect to multiplication by real numbers, we need only to prove that $W$ is closed with respect to addition, that is, if $\xi', \xi'' \in W$, then $\xi = \xi' + \xi'' \in W$. Since $\gamma_{\xi'}, \gamma_{\xi''} \subset H$ (Step 3), the curve $t \mapsto \gamma_{\xi'}(t)\gamma_{\xi''}(t)$ is contained in $H$. The velocity vector of this curve at 0 is $\xi' + \xi'' = \xi$ (Section 1.2.1) and hence $\xi \in W$ (Step 2).

*Step 5.* We already know that $\exp(W) \subset H$ (Step 3), and it remains to prove that, at least at some neighborhood of $e$, $H$ has no elements not in $\exp(W)$. Let $Z = W^\perp$. In some neighborhood of $e$, $G$ is the product $\exp(W)\exp(Z)$ (meaning that the map $\exp(W) \times \exp(Z) \to G$, $((h, j) \mapsto hj$ is a local diffeomorphism). We want to prove that the following is impossible: there exists a sequence $\{h_i\}$ such that every $h_i \in H \cap (G - \exp(W))$ and $\lim\limits_{i\to\infty} = e$. But if such a sequence exists, we can project it (maybe, starting from some term) onto $\exp(Z)$, and we get a new sequence, $\{h_i^\circ\}$ such that every $h_i^\circ \in H \cap (\exp(Z) - e)$ and $\lim\limits_{i\to\infty} h_i^\circ = e$. Since $S \cap Z$ is compact, some subsequence of the sequence $\left\{\dfrac{\exp^{-1}(h_i^\circ)}{\|\exp^{-1}(h_i^\circ)\|}\right\}$ has a limit, this limit is virtually tangent to $H$ and does not belong to $W$. This is a contradiction, which completes the proof of Theorem.

### 1.3. The Lie algebra of a Lie group.

**1.3.1. An algebraic introduction.** Let $L$ be a real or complex[4]) vector space. It is called a *Lie algebra*, if there given a bilinear multiplication $L \times L \to L$ (traditionally denoted by the symbol $[x, y]$ and called the *commutator*) satisfying the axioms

(1) $[x, y] = -[y, x]$,
(2) $[[x, y], z] + [[y, z], x] + [[z, x], y] = 0$.

The most visible example is provided by actual commutators in associative algebras: if $L$ is an associative algebra, then the same space $L$ is a Lie algebra with respect to the operation $[x, y] = xy - yx$. Another example (this time, infinite dimensional) is the space $\operatorname{Vect} M$ of vector fields on a smooth manifold $M$ with the classical commutator operation. A more elementary example is provided by the cross-product in $\mathbb{R}^3$.

---

[4]) Actually, any ground field of characteristic 0 will work. The case of finite characteristic is also possible (with some complications), but this characteristic should not be 2.

Exercise 11. Check the Lie algebra axioms for these three examples.

The usual algebraic terminology is applied to Lie algebras: homomorphisms, subalgebras, ideals, etc.

We will return to the algebraic theory of Lie algebras when we need it.

**1.3.2. Lie algebra structure in $T_eG$.** Let $G$ be a Lie group. For $\xi, \eta \in T_eG$ there arise right invariant vector fields $R_\xi, R_\eta$, and their commutator $[R_\xi, R_\eta]$ is also right invariant, and hence has the form $R_\zeta$ for a uniquely defined $\zeta \in T_eG$. This $\zeta$ is taken for the commutator $[\xi, \eta]$ of the chosen elements of $T_eG$. The space $T_eG$ with this commutator operation is a Lie algebra (the Lie algebra axioms follow from these axioms for commutators of vector fields); it is called the *Lie algebra of the Lie group $G$* and is denoted as $\mathrm{Lie}\, G$. Also, if a Lie group is denoted by an upper case Roman letter, its Lie algebra is usually denoted by the same lower case Fraktur (German) letter. So, the Lie algebras of $GL(n, \mathbb{R})$, $GL(n, \mathbb{C})$, $SL(n, \mathbb{R})$, $SL(n, \mathbb{C})$, $O(n)$, $SO(n)$, $U(n)$, $SU(n)$ are denoted as $\mathfrak{gl}(n, \mathbb{R})$, $\mathfrak{gl}(n, \mathbb{C})$, $\mathfrak{sl}(n, \mathbb{R})$, $\mathfrak{sl}(n, \mathbb{C})$, $\mathfrak{o}(n)$, $\mathfrak{so}(n)$, $\mathfrak{u}(n)$, $\mathfrak{su}(n)$.

In Analysis, there is a direct limit formula for the commutator of vector fields. Namely, a vector field $X$ on a manifold $M$ determines, locally with respect to both $M$ and parameter, a flow $\varphi_t \colon M \to M$ such that, for a $p \in M$, $X_p$ is a velocity vector at $p$ of the curve $t \mapsto \varphi_t(p)$. Then, if $X$ and $Y$ are vector fields on $M$, and $\varphi_t, \psi_t$ are corresponding flow, then $[X, Y]_p$ is the velocity vector at $p$ of the curve $t \mapsto \psi_{-\sqrt{t}} \circ \varphi_{-\sqrt{t}} \circ \psi_{\sqrt{t}} \circ \varphi_{\sqrt{t}}(p)$ (defined for $t \geq 0$).If we apply these rules to the vector fields $R_\xi, R_\eta$ and $p = e$, we will obtain the following procedure for finding $[\xi, \eta]$: we should take the one-parameter subgroups $\gamma_\xi$ and $\gamma_\eta$, then compute

$$\gamma_\xi(\sqrt{t})\gamma_\eta(\sqrt{t})\gamma_\xi(-\sqrt{t})\gamma_\eta(-\sqrt{t})$$

and take the velocity vector for this at $t = 0$. For $G = GL(n, \mathbb{R})$ or $GL(n, \mathbb{C})$, and $\xi = A, \eta = B$, the last product is

$$\begin{aligned}
&\exp(A\sqrt{t})\exp(B\sqrt{t})\exp(-A\sqrt{t})\exp(-B\sqrt{t}) \\
&= \left(I + A\sqrt{t} + \frac{A^2}{2}t + \dots\right)\left(I + B\sqrt{t} + \frac{B^2}{2}t + \dots\right) \cdot \\
&\qquad\qquad \left(I - A\sqrt{t} + \frac{A^2}{2}t - \dots\right)\left(I - B\sqrt{t} + \frac{A^2}{2}t - \dots\right) \\
&= I + (A + B - A - B)\sqrt{t} + \\
&\qquad \left(AB - A^2 - AB - BA - B^2 + AB + \frac{A^2}{2} + \frac{B^2}{2} + \frac{A^2}{2} + \frac{B^2}{2}\right)t + \dots \\
&\qquad\qquad = I + (AB - BA)t + \dots
\end{aligned}$$

Thus, the commutator in $\mathfrak{gl}(n, \mathbb{R})$ and $\mathfrak{gl}(n, \mathbb{C})$ is defined by the habitual formula $[A, B] = AB - BA$, and is defined by the same formula for the Lie algebras of all Lie groups listed above. By the way, this shows that the spaces of matrices with zero trace, as well as the spaces of skew symmetric and skew-Hermitian matrices, are closed with respect to the commutator operation (actually, the commutator of *any* two matrices has zero trace).

This fact is, certainly, obvious, but still it should be noted, that neither symmetric, nor Hermitian matrices form a space closed with respect to commutators.

Remark. In our description of the commutator in $T_eG$ we can replace one-parameter subgroups by any smooth curves: if $\gamma_1, \gamma_2\colon (-\varepsilon, \varepsilon) \to G$ are smooth curves with $\gamma_1(0) = \gamma_2(0) = e$, then the velocity vector of the curve $t \mapsto \gamma_1(\sqrt{t})\gamma_2(\sqrt{t})\gamma_1(-\sqrt{t})\gamma_2(-\sqrt{t})$ is $[\dot\gamma_1(0), \dot\gamma_2(0)]$; this follows from the fact that every parametrized smooth curve through $e$ is tangent (as a parametrized curve) to some one-parameter subgroup.

**1.3.3. Lie group homomorphisms and Lie algebra homomorphisms.** *A. Main statement.* Let $f\colon H \to G$ be a Lie homomorphism. Then there arises a linear map $d_ef\colon T_eH \to T_eG$, and it is easy to understand that it is, actually, a Lie algebra homomorphism $\operatorname{Lie} H \to \operatorname{Lie} G$. Indeed, $f$ takes one-parameter subgroups of $H$ into one-parameter subgroups of $G$, and the description of commutators in Lie algebras of Lie groups shows that $d_ef$ takes commutators into commutators. We will denote the Lie algebra homomorphism constructed simply by $df$, or, sometimes, by $\operatorname{Lie} f$. Notice that, in particular, the Lie algebra $\operatorname{Lie} H$ of a Lie subgroup $H$ of a Lie group $G$ is a Lie subalgebra of $\operatorname{Lie} H$.

The goal of this Section is to show that, essentially, the homomorphisms $f$ and $\operatorname{Lie} f$ determine each other. Thanks to this (and to the fact that every finite-dimensional Lie algebra is a Lie algebra of a certain Lie groups; we will discuss this later), the theory of Lie groups can be, essentially, reduced to the pure algebraic theory of Lie algebras, which leads to many important results concerning Lie groups.

Here is our main statement. (Some additional statements will appear in the proof.)

Theorem. (a) *let $f, f'\colon H \to G$ be Lie homomorphisms, and let $H$ be connected. If* $\operatorname{Lie} f = \operatorname{Lie} f'$, *then $f = f'$.*

(b) *Let $G, H$ be Lie groups, and let $H$ be connected. Let also $\varphi\colon \operatorname{Lie} H \to \operatorname{Lie} G$ be a Lie algebra homomorphism. Then there exist a covering $\widetilde H \to H$ and a (unique by Part (a)) Lie homomorphism $f\colon \widetilde H \to G$ such that* $\operatorname{Lie} f = \varphi$. (Recall that $\widetilde H$ is a Lie group (see Section 1.1.4) and (obviously) $\operatorname{Lie} \widetilde H = \operatorname{Lie} H$.) *In particular, if $H$ is simply connected, then $\widetilde H = H$.*[5]).

*B. Proof of (a).* Let $\operatorname{Lie} f = \operatorname{Lie} f'$, and let $A = \{g \in G \mid f(g) = f'(g)\}$. Since $f$ and $f'$ are continuous, $A$ is closed. Since Lie homomorphisms take one-parameter subgroup, and every element of the Lie algebra is a velocity vector of precisely one one-parameter subgroup, $f$ and $f'$ must coincide on every one-parameter subgroup; in particular they coincide on a neighborhood $U$ of $e$ where $\exp^{-1}$ is a diffeomorphism. If $g \in A$, then $f$ and $f'$ coincide on every element of $gU$, so $gU$, which is a neighborhood of $g$, is contained in $A$. Thus, $A$ is also open, and, since $G$ is connected, $A = G$ ($A$ is non-empty, since $e \in A$).

[5]) A path connected topological space $M$ is called *simply connected*, if every two paths $s, s'\colon [0,1] \to M$ with $s(0) = s'(0)$, $s(1) = s'(1)$ are homotopic. A theorem in theory of coverings (see footnote[2])) states that every sufficiently good topological space, in particular, every manifold, possesses a unique with respect to a natural equivalence simply connected covering. This covering is also a covering over any other covering of $M$, and by this reason it is called *universal*. For example, a simply connected space is its own universal covering, and hence simply connected spaces have no other coverings.

*C. Proof of local version of (b).* There exists a *local version* of the theory of Lie groups: we will turn to it from time to time; now we will need two definitions. A *local Lie subgroup* of a Lie group $G$ is a connected submanifold $H$ of a neighborhood $U$ of $e \in G$ such that (1) if $h_1, h_2 \in H$ and $h_1 h_2 \in U$, then $h_1 h_2 \in H$; (2) if $h \in H$ and $h^{-1} \in U$, then $h^{-1} \in H$. A local Lie homomorphism of a Lie group $G$ into a Lie group $H$ is a smooth map $f$ of a neighborhood $U$ of $e \in G$ into $H$ such that (1) if $g_1, g_2, g_1 g_2 \in U$, then $f(g_1 g_2) = f(g_1) f(g_2)$; (2) if $g, g^{-1} \in U$, then $f(g^{-1}) = f(g)^{-1}$. It is clear that if $H$ is a local Lie subgroup of a Lie group $G$, then $T_e H$ is a Lie subalgebra of the Lie algebra $\mathrm{Lie}\, G$ (we can denote this subalgebra as $\mathrm{Lie}\, H$). Also, if $f$ is a local Lie homomorphism of a Lie group $G$ into a Lie group $H$, then there arises a Lie algebra homomorphism $\mathrm{Lie}\, f \colon \mathrm{Lie}\, G \to \mathrm{Lie}\, H$.

LEMMA 1. *Let $\mathfrak{h}$ be a Lie subalgebra of the Lie algebra $\mathfrak{g} = \mathrm{Lie}\, G$ of a Lie group $G$. Then there exists a local Lie subgroup $H$ of $G$ such that $\mathrm{Lie}\, H = \mathfrak{h}$.*

*Proof* is based on the following *Frobenius Integrability Criterion.* Let $M$ be an $n$-dimensional manifold, and let $0 < m < n$. Suppose that for every $p \in M$ there fixed an $m$-dimensional subspace $S_p$ of $T_p M$ in such a way that $S_p$ depends smoothly on $p$ (that is, the union $\bigcup_p S_p$ is a smooth submanifold of the manifold $TM = \bigcup_p T_p M$ of tangent vectors to $M$). Such a family $S = \{S_p\}$ is called an *$m$-dimensional distribution on $M$*. A vector field $X \in \mathrm{Vect}\, M$ is called *subordinated* to the distribution $S$, if $X_p \in S_p$ for every $p \in M$. A distribution $S$ is called *integrable*, if for every point $p \in M$ there exists an $m$-dimensional submanifold $N_p$ of a neighborhood of $p$ such that for every point $q \in N$, $T_q N_p = S_q$; such a submanifold is called an *integral submanifold.* It is obvious that an integral submanifold is *locally unique*, that is, if two ijntegral submanifolds have a common point, then they match in a neighborhood of this point.

EXERCISE 12. Prove that the following two (codimension one) distributions (which are, actually, essentially the same) are not integrable.

(a) $M = \mathbb{R}^{2n+1}$, for $p = (x_1^\circ, \ldots, x_{2n+1}^\circ)$ the equation of the space $N_p \subset T_p \mathbb{R}^{2n+1} = \mathbb{R}^{2n+1}$ is $x_{2n+1} = \sum_{i=1}^n x_{n+i}^\circ x_i$.

(b) $M = S^{2n+1} \subset \mathbb{R}^{2n+2} = \mathbb{C}^{n+1}$, for a $p \in M$ $N_p$ is the (unique) $n$-dimensional complex subspace of the real $(2n+1)$-dimensional space $T_p M$ (otherwise, $N_p = T_p M \cap iT_p M$). Another statement of the same problem: for no holomorphic function $f$ in a neighborhood of $p \in \mathbb{C}^{n+1}$ the manifold $\{f = 0\}$ is contained in $S^{2n+1}$.

The Frobenius integrability criterion says that *a distribution $S$ is integrable if and only if for every two vector fields $X, Y$ subordinated to $S$ their commutator $[X, Y]$ is also subordinated to $S$* [6]).

---

[6]) Let us briefly discuss the proof of this statement. In one direction, it is obvious: if $S$ is integrable, then a vector field $X$ is subordinated to $S$ if and only if for every integral submanifold $N_p$ the restriction of $X$ to $N_p$ is tangent to $N_p$; but if two vector fields on $M$ are tangent to some submanifold $N$ of $M$, then so is their commutator; hence, the commutator of two vector fields subordinated to $S$ is subordinated to $S$. The proof in the opposite direction is more involved, and we restrict ourselves to its schematic presentation. To begin with, let $m = 2$; let $X$ and $Y$ be two vector fields subordinated to $S$ such that,

Consider the distribution $S = \{S_g\}$, $S_g = d\lambda_g(\mathfrak{h})$ on $G$. Let us show that $S$ satisfies the condition in the integrability criterion. Let $\eta_1, \dots \eta_m$ be a basis in $\mathfrak{h}$. Then the vector fields $L_{\eta_1}, \dots, L_{\eta_m}$ are subordinated to $S$, and the commutators $[L_{\eta_j}, L_{\eta_k}]$ are linear combinations (with coefficients in $\mathbb{R}$) of the vector fields $L_{\eta_i}$. Furthermore, every vector field subordinated to $S$ is the linear combination of the vector fields $L_{\eta_i}$ with functional coefficients. If $X = \sum_i f_j L_{\eta_i}$ and $Y = \sum_k g_k L_{\eta_i}$, then

$$[X, Y] = \sum_{jk} \left( f_j g_k [L_{\eta_j}, L_{\eta_k}] + f_j L_{\eta_j} g_k L_{\eta_k} - g_k L_{\eta_k} f_j L_{\eta_j} \right),$$

which is subordinated to $S$. Thus, $S$ is integrable, and, in particular, $S$ possesses an integral $m$-dimensional manifold $H \ni e$ in a neighborhood of $e$ tangent to $S$. Since $S$ is left invariant, the images $\lambda_g(H)$ are also integral manifolds of $S$. In particular, if $h \in H$, then $\lambda_h(H)$ matches with $H$ in a neighborhood of $h$, that is, for $h' \in H$ sufficiently close to $e$, $hh' \in H$. This means that $H$ is a local Lie subgroup of $G$ with $\operatorname{Lie} H = \mathfrak{h}$.

LEMMA 2. *Let $\varphi\colon \operatorname{Lie} H \to \operatorname{Lie} G$ be a Lie algebra homomorphism. Then there exists a local Lie homomorphism $f$ of $H$ into $G$ such that* $\operatorname{Lie} f = \varphi$.

*Proof.* Consider the "graph of $\varphi$", that is, the subspace

$$\mathfrak{gr}(\varphi) = \{(\xi, \eta) \in \mathfrak{h} \times \mathfrak{g} \mid \eta = \varphi(\xi)\}$$

of the Lie algebra $\mathfrak{h} \times \mathfrak{g} = \operatorname{Lie}(H \times G)$. This is a Lie subalgebra of $\mathfrak{h} \times \mathfrak{g}$: for $\eta, \eta' \in \mathfrak{h}$,

$$[(\eta, \varphi(\eta)), (\eta', \varphi(\eta'))] = ([\eta, \eta'], [\varphi(\eta), \varphi(\eta')]) = ([\eta, \eta'], \varphi[\eta, \eta']) \in \mathfrak{gr}(\varphi).$$

---

in a neighborhood of $p$, the vectors $X_q, Y_q$ form a basis of $S_q$. Choose a small $(n-2)$-dimensional disk $D \subset M$ centered at $p$ and transverse to $S_q$ for every $q \in D$. Then, for every $q \in D$ take the trajectory $\phi_q\colon (-\varepsilon, \varepsilon) \to M$ of $X$ with $\varphi_q(0) = q$, after which, for every $q \in D$ and $t \in (-\varepsilon, \varepsilon)$, take the trajectory $\psi_{d,t}\colon (-\varepsilon, \varepsilon) \to M$ of $Y$ with $\psi_{d,t}(0) = \varphi_d(t)$. If $D$ and $\varepsilon$ are sufficiently small, then the formula $(d, t, u) \mapsto \psi_{d,t}(u)$ provides an embedding $D \times (-\varepsilon, \varepsilon) \times (-\varepsilon, \varepsilon) \to M$; we will show that the surfaces $d \times (-\varepsilon, \varepsilon) \times (-\varepsilon, \varepsilon)$ $(d \in D)$ are integral for the distribution $S$, which will mean that this distribution is integrable. We need to prove that if a function $F\colon D \times (-\varepsilon, \varepsilon) \times (-\varepsilon, \varepsilon) \to \mathbb{R}$ is constant on every $d \times (-\varepsilon, \varepsilon) \times (-\varepsilon, \varepsilon)$, then $XF = 0$ and $YF = 0$. What we already know, is that $YF = 0$ (the trajectories of $Y$ are contained in the surfaces $d \times (-\varepsilon, \varepsilon) \times (-\varepsilon, \varepsilon)$) and $XF = 0$ on $D \times (-\varepsilon, \varepsilon) \times 0$ (which is made of trajectories of $X$). On the other hand, our condition for the commutators means that, on $D \times (-\varepsilon, \varepsilon) \times (-\varepsilon, \varepsilon)$, $[X, Y] = fX + gY$ where $f$ and $g$ are two functions. Hence, $X(YF) - Y(XF) = f(XF) + g(YF)$, which means, since $YF = 0$, that $Y(XF) + f(XF) = 0$. This is a first order differential equation for $XF$, and the condition $XF|_{D \times (-\varepsilon, \varepsilon) \times 0} = 0$ makes the solution unique; since $XF = 0$ is a solution, we have $XF = 0$.

The proof in the case $m > 2$ is basically the same: the square $(-\varepsilon, \varepsilon) \times (-\varepsilon, \varepsilon)$ is replaced by the $m$-dimensional cube $(-\varepsilon, \varepsilon) \times \ldots \times (-\varepsilon, \varepsilon)$, the vector fields $X, Y$ are replaced by $m$ vector fields $X_1, \dots, X_m$, and the two-step construction of the embedding $D \times (-\varepsilon, \varepsilon) \times (-\varepsilon, \varepsilon) \to M$ is replaced by an $m$-step construction of the embedding $D \times (-\varepsilon, \varepsilon) \times \ldots \times (-\varepsilon, \varepsilon) \to M$. All the rest is the same as before (each of the equalities $X_2 F = 0, \dots, X_m F = 0$ is proved with the help of a first order differential equation).

Hence, by Lemma 1, there exists a local Lie subgroup $GR(\varphi)$ of $H\times G$ with $\operatorname{Lie} GR(\varphi)=\mathfrak{gr}(\varphi)$. This local subgroup can be regarded as the graph of a local Lie homomorphism $f$ of $H$ into $G$ with $\operatorname{Lie} f=\varphi$.

*D. A globalization: proof of Statement (b).* In this proof, we will use not just Lemmas 1 and 2, but also some details of their proofs. In the proof of Lemma 2, we constructed a Lie subalgebra $\mathfrak{gr}(\varphi)$ of $\operatorname{Lie}(H\times G)$, and, according to proof of Lemma 1, this subalgebra gives rise to an left invariant integrable distribution $S=\{S_{(h,g)}=d\lambda_{(h,g)}(\mathfrak{gr}(\varphi))\}$ on $H\times G$. There arises an integral manifold $GR\subset H\times G$ which we can assume to be diffeomorphic to the ball $D^m$ ($m=\dim H$) centered at $(e,e)\in H\times G$; this manifold has a one-to one projection on $H$. All manifolds $\lambda_{(h,g)}(GR)$ are integral manifold of $H\times G$ with a one-to-one projection onto $H$; if two such integral manifold have a common point, then they match in a neighborhood of this point. We introduce a new topology in $H\times G$ (which may look monstrous): a base of this topology is formed by open subsets of manifolds $\lambda_{(h,g)}(GR)$. This topology arises from the $m$-dimensional atlas $\{GR,\lambda_{(h,g)}|_{GR}\}$ of $H\times G$ which could make $H\times G$ into an $m$-dimensional manifold, if the topology had been second countable, which it is not. But this difficulty disappears, if we restrict ourselves to a path component (with respect to this topology) of $(e,e)$ which we denote by $\widetilde{H}$. This component, with respect to the topology described, is an $m$-dimensional manifold; indeed, the projection $\widetilde{H}\to H$ is (obviously) a covering, so $\widetilde{H}$ is not just a manifold, but also a Lie group. The inclusion map $\widetilde{H}\to H\times G$ is a one-to-one Lie homomorphism (although $\widetilde{H}$ is not a Lie subgroup of $H\times G$, since, in general, it is not closed; it may be even dense). The restriction $f\colon\widetilde{H}\to G$ of the projection $H\times G\to G$ is a Lie homomorphism whose restriction to $GR$ is the graph of a local Lie homomorphism of $H$ into $G$ constructed in the proof of Lemma 2. In particular $\operatorname{Lie} f=\varphi$, which completes the proof of Statement (b), and, hence, of Theorem [7]).

*E. The language of representation.* A (complex, $N$-dimensional) representation of a (connected) Lie group $G$ is, by (not the most popular) definition a Lie homomorphism $G\to GL(N,\mathbb{C})$; a (complex, $N$-dimensional) representation of a Lie algebra $\mathfrak{g}$ is a Lie algebra homomorphism $\mathfrak{g}\to\mathfrak{gl}(N,\mathbb{C})$. The constructions above can be applied to representations. If $R\colon G\to GL(N,\mathbb{C})$ is a representation of a Lie group $G$, then there arises a representation $\rho\colon\mathfrak{g}=\operatorname{Lie} R\colon\operatorname{Lie} G\to\mathfrak{gl}(N,\mathbb{C})$. On the other side, a representation $\rho\colon\mathfrak{g}=\operatorname{Lie} G\to\mathfrak{gl}(N,\mathbb{C})$ gives rise to a representation $R\colon\widetilde{G}\to GL(N,\mathbb{C})$ of some covering $\widetilde{G}$ of $G$. For example, a representation of $\mathfrak{sl}(n)$ (with $n\geq 2$) gives rise to a representation of not, in general, the

---

[7]) The construction in this proof has a broadly known generalization. An $m$-dimensional integrable distribution on an $n$-dimensional manifold $M$ gives rise to a new topology on $M$, with respect to which $M$ possesses an $m$-dimensional atlas without the second countability. Path component of this topology become $m$-dimensional manifolds, with one-to-one immersions into $M$. The images are not, in general, submanifolds of $M$, they may be even dense in $M$. The whole structure is called a *foliation*, the described immersed manifolds are called *leaves* of the foliation. Foliations are studied in Dofferential Topology, there are books about them (for example, I. Tamura, "*Topology of foliations: an introduction*," Translations of Math. Monographs, Vol. 97, Amer. Math, Soc, 1992). We will not consider them in these lectures any seriously.

group $SO(n)$, but rather of the group $Spin(n)$.

**1.3.4. Lie subalgebras of Lie algebras and Lie subgroups of Lie groups.** *A. A question and a negative answer.* Let $G$ be a Lie group, and let $\mathfrak{h}$ be a Lie subalgebra of its Lie algebra $\mathfrak{g} = \operatorname{Lie} G$. There arises a question: is there a Lie subgroup $H$ of $G$ with $\operatorname{Lie} H = \mathfrak{h}$? The immediate answer is *no*, because of the following, very simple example. Let $G = S^1 \times S^1$ be a torus. There is a covering $\mathbb{R}^2 \to S^1 \times S^1$, and $\operatorname{Lie}(S^1 \times S^1) = \operatorname{Lie} \mathbb{R}^2 = \mathbb{R}^2$ with zero commutator (we will usually refer to a Lie algebra with a zero commutator as to a *commutative* Lie algebra). Every subspace of a commutative Lie algebra is a Lie subalgebra. Let $L \subset \mathbb{R}^2$ be a one dimensional subspace of $^2$ (a line). Obviously, $L = \operatorname{Lie} L$ (the Lie algebra of itself). The same $L$ regarded as a subalgebra of $\operatorname{Lie}(S^1 \times S^1)$ can be a Lie algebra only of the image of $L$ with respect to the projection $\mathbb{R}^2 \to S^1 \times S^1$. This image, however, may be not closed in $S^1 \times S^1$; actually, if the slope of $L$ in $\mathbb{R}^2$ is irrational, then the image of $L$ in $S^1 \times S^1$ is dense. This construction has a generalization: a subalgebra $L$ of the Lie algebra $\operatorname{Lie}(\underbrace{S^1 \times \ldots \times S^1}_{n}) = \operatorname{Lie}(\mathbb{R}^n) = \mathbb{R}^n$ is the Lie algebra of the same $L$ in $\mathbb{R}^n$, but, in general, is not a Lie algebra of any Lie subgroup of the "torus" $S^1 \times \ldots \times S^1$: the image of $L$ in the torus is, in general, not closed; its closure may be a "subtorus" of some dimension between $n$ and $m = \dim L$, maybe, even, the whole torus $S^1 \times \ldots \times S^1$.

Although this example looks very convincing, we must say that in some (maybe, rather loose) sense *there are no other examples*. The explanation of this is the goal of this Section.

*B. Useful terminology: virtual subgroups.* A virtual Lie subgroup of a Lie group $G$ is, by definition, the image of a one-to-one Lie homomorphism $H \to G$ where $H$ is another Lie group. The examples above are examples of virtual Lie subgroups of tori. Usually we will denote the image of such a homomorphism $H \to G$ by the same letter $H$. Thus, a virtual subgroup of a Lie group is a subgroup in the algebraic sense, and it has a Lie group structure of its own. In particular, it has a Lie algebra, which is a subalgebra of a Lie algebra of the ambient Lie group. The only reason why we cannot count it as a Lie subgroup of $G$ is that it is not closed in $G$.

Proposition. *Every subalgebra of the Lie algebra of a Lie group $G$ is the Lie algebra of some (unique, if we assume it path connected) virtual Lie subgroup of $G$.*

*Proof* is basically known to us. A subalgebra $\mathfrak{h}$ of a Lie algebra $\mathfrak{g} = \operatorname{Lie} G$ is included in the left invariant distribution $\{d_e\lambda_g(\mathfrak{h})$, the integral manifolds of this distribution from a new topology in $G$ (see Proof of Lemma 2 and footnote [7])), and the path component $H$ of $e$ with respect to this topology is a virtual subgroup of $G$ whose Lie algebra is $\mathfrak{h}$.

Exercise 13. Prove that images and inverse images of virtual Lie subgroups with respect to Lie homomorphisms are virtual Lie subgroups. (Compare with the remark after Exercise 4.) Prove also that the Lie algebras of images and inverse images of a virtual Lie subgroup $H$ (of a Lie group) are images and inverse images of $\operatorname{Lie} H$.

*C. Commutator subgroups of Lie groups and commutator subalgebras of Lie algebras.* The *commutator subalgebra* (also called the *derived subalgebra*) $\mathfrak{g}'$ of a Lie algebra $\mathfrak{g}$ is the subspace of $\mathfrak{g}$ spanned by all commutators $[g, h]$, $d, h \in \mathfrak{g}$.

Exercise 14. Prove that $\mathfrak{g}'$ is a Lie subalgebra, moreover, an ideal in $\mathfrak{g}$. Prove also that the Lie algebra $\mathfrak{g}/\mathfrak{g}'$ is commutative.

The *commutator subgroup* $G'$ of a Lie group $G$ is its commutator subgroup in the algebraic sense.

PROPOSITION. *Let $G$ be a Lie group, and let $\mathfrak{g} = \operatorname{Lie} G$. If $G$ is connected, then $G'$ is a virtual Lie subgroup of $G$. If $G$ is simply connected, then $G'$ is a Lie subgroup of $G$. In both cases,* $\operatorname{Lie} G' = \mathfrak{g}'$.

*Proof.* We begin with the case when $G$ is simply connected. Since $\mathfrak{g}'$ is an ideal in $\mathfrak{g}$ and $\mathfrak{g}/\mathfrak{g}' = \operatorname{Lie} \mathbb{R}^n$ (with some $n$), the projection $\mathfrak{g} \to \operatorname{Lie} \mathbb{R}^n$ is the differential of some Lie homomorphism $p\colon G \to \mathbb{R}^n$ (Part (b) of Theorem in Section 1.3.3). Let $H = \operatorname{Ker} p$; it is a Lie subgroup of $G$ and $\operatorname{Lie} H = \mathfrak{g}'$. We will show that $H = G'$.

The inclusion $G' \subset H$ holds, since $\mathbb{R}^n$ is commutative. Moreover, since $\mathbb{R}^n$ is simply connected and $G$ is connected, $H$ is also connected: if $H_0$ is the component of $e$ in $H$, then $G/H_0$ is a covering over $\mathbb{R}^n = G/H$, hence $H_0 = H$.

To prove that $G' \supset H$, we notice that $G'$ contains the curves $\gamma_\xi(t)\gamma_\eta(t)\gamma_\xi(-t)\gamma_\eta(-t)$ (for $\xi, \eta \in \mathfrak{g}$) and also product of such curves. From this, $G'$ contains the exponential image of a neighborhood of 0 in $\mathfrak{g}' = \operatorname{Lie} \mathfrak{H}$, thus, it contains a neighborhood of $e$ in $H$, thus it contains the component of $e$ in $H$ (see Exercise 10), thus it contains $H$.

If $G$ is not simply connected, then we consider the universal covering $p\colon \widetilde{G} \to G$. Since $G' = p(\widetilde{G}')$ (because $\widetilde{\mathfrak{g}} = \mathfrak{g}$ and hence $\widetilde{\mathfrak{g}}' = \mathfrak{g}'$), $G'$ is a virtual Lie subgroup of $G$ (the image of a Lie homomorphism $(\widetilde{G})' \to G$.

*D. The structure of virtual Lie subgroups.* Let $G$ be a connected Lie group, $\mathfrak{g} = \operatorname{Lie} G$ be the corresponding Lie algebra, $\mathfrak{h} \subset \mathfrak{g}$ be a Lie subalgebra, $H$ be a virtual subgroup of $G$ with $\operatorname{Lie} H = \mathfrak{h}$. Our goal is ti investigate, how much $H$ is different from Lie subgroups of $G$.

Let $H^M$ be the closure of $H$; it is a Lie subgroup of $G$, the intersection of all Lie subgroups of $G$ which contain $H$. Let $\operatorname{Lie} H^M = \mathfrak{h}^M \supset \mathfrak{h}$; it is a Lie subalgebra of $\mathfrak{g}$, it is called the *Malcev closure* of $\mathfrak{h}$.

LEMMA 3. $(\mathfrak{h}^M)' = \mathfrak{h}' \subseteq \mathfrak{h}$.

*Proof* is contained in subsection $F$ below.

Thus $\mathfrak{h}$ is squeezed between $\mathfrak{h}^M$ and $(\mathfrak{h}^M)'$. There arises a diagram

$$\begin{array}{ccccc} \mathfrak{h}^M & \to & \mathfrak{h}^M/(\mathfrak{h}^M)' & = & \mathbb{R}^n \\ & & & & \cup \\ \mathfrak{h} & \to & \mathfrak{h}/\mathfrak{h}' & = & \mathbb{R}^m \end{array}$$

The upper row gives rise to a Lie homomorphism $f\colon \widetilde{H}^M \to \mathbb{R}^n$ and, obviously, $H = p(f^{-1}(\mathbb{R}^m))$ (where $p\colon \widetilde{H}^M \to H^M$ is the universal covering). This shows that an arbitrary virtual subgroup of $G$ is obtained by a construction which is a direct generalization of construction in Subsection $A$.

*E. Ad and ad.* There are two important representations, one of a Lie group $G$ and another one of its Lie algebra $\mathfrak{g}$. They are called *adjoint* and denoted accordingly.

For a $g \in G$, we define a linear automorphism $\operatorname{Ad}_G g = \operatorname{Ad} g\colon \mathfrak{g} \to \mathfrak{g}$ as $d_e\alpha_g$ (where $\alpha_g\colon G \to G$, $\alpha_g(h) = ghg^{-1}$, see Section 1.2.2). The correspondence $\operatorname{Ad}_G\colon G \to GL(\mathfrak{g})$ is a (real) representation of $G$ in the space $\mathfrak{g}$. There arises a representation of $\operatorname{Lie} G = \mathfrak{g}$ which we denote as $\mathrm{ad}_{\mathfrak{g}}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$. (For a vector space $V$, $GL(V)$ denotes the Lie group of linear automorphisms of $V$, and $\mathfrak{gl}(V) = \operatorname{Lie} GL(V)$ is the Lie algebra of linear endomorphisms of $V$.

Lemma 4 (a major fact). *For $\xi, \eta \in \mathfrak{g}$, $(\operatorname{ad}\xi)\eta = [\xi, \eta]$.*

*Proof.*

$$\begin{aligned}
(\operatorname{ad}\xi)\eta &= \lim_{t\to 0} \frac{(\operatorname{Ad}\gamma_\xi(t))\eta - \eta}{t} \\
&= \lim_{\substack{t\to 0\\ u\to 0}} \frac{\gamma_\xi(t)\gamma_\eta(u)\gamma_\xi(-t)\gamma_\eta(-u)}{tu} = \lim_{t\to 0} \frac{\gamma_\xi(t)\gamma_\eta(t)\gamma_\xi(-t)\gamma_\eta(-t)}{t^2} = [\xi, \eta].
\end{aligned}$$

(We identify neighborhoods of $e \in G$ and $0 \in \mathfrak{g}$ by means of the map exp.)

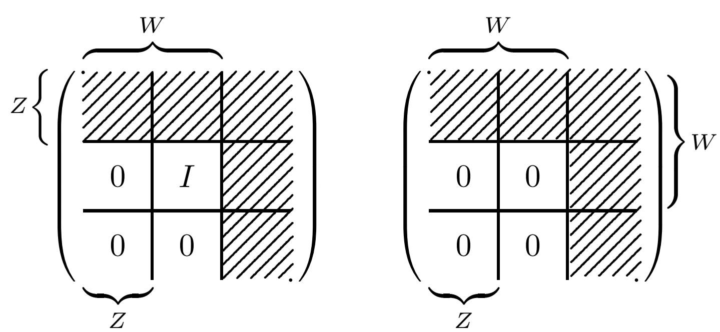
*F. Proof of Lemma 3.* We will use (twice) the following construction. Let $V \supset W \supset Z$ be a triple of vector spaces. Put

$$GL(V; W, Z) = \{A \in GL(V) \mid (A - I)W \subset Z\};$$

then

$$\mathfrak{gl}(V; W, Z) = \operatorname{Lie} GL(V; W, Z) = \{A \in \mathfrak{gl}(V) \mid AW \subset Z\};$$

thus the matrices from $GL(V; W, Z)$ and $\mathfrak{gl}(V; W, Z)$ have the block form




$GL(V; W, Z)$ $\qquad$ $\mathfrak{gl}(V; W, Z)$

(the diagonal blocks of a matrix from $GL(V; W, Z)$ must be non-degenerate). For a representation (a Lie homomorphism) $\varphi\colon G \to GL(V)$,

$$\begin{gathered}
H = \varphi^{-1} GL(V; W, Z) = \{g \in G \mid (\varphi(g) - \mathrm{id})W \subset Z\}, \\
\mathfrak{h} = \operatorname{Lie} H = \{\xi \in \mathfrak{g} \mid d\varphi(\xi)W \subset Z\}
\end{gathered}$$

(where $\mathfrak{g} = \operatorname{Lie} G$).

First, let $\varphi = \mathrm{Ad}$, $V = \mathfrak{g}$, $W = \mathfrak{h}$, $Z = \mathfrak{h}'$. Then

$$\begin{aligned}
H_1 &= \{g \in G \mid (\operatorname{Ad} g - \mathrm{id})\mathfrak{h} \subset \mathfrak{h}'\}, \\
\mathfrak{h}_1 &= \{\xi \in \mathfrak{g} \mid [\xi, \mathfrak{h}] \subset \mathfrak{h}'\}.
\end{aligned}$$

Thus, $\mathfrak{h}_1 \supset \mathfrak{h}$, hence $\mathfrak{h}_1 \supset \mathfrak{h}^M$, and hence $[\mathfrak{h}^M, \mathfrak{h}] \subset \mathfrak{h}'$.

Second, let $\varphi = \mathrm{Ad}$, $V = \mathfrak{g}$, $W = \mathfrak{h}^M$, $Z = \mathfrak{h}'$. Then

$$H_2 = \{g \in G \mid (\mathrm{Ad}\, g - \mathrm{id})\mathfrak{h}^M \subset \mathfrak{h}'\},$$
$$\mathfrak{h}_2 = \{\xi \in \mathfrak{g} \mid [\xi, \mathfrak{h}^M] \subset \mathfrak{h}'\}.$$

By the previous remark, $\mathfrak{h}_2 \supset \mathfrak{h}$, hence, $\mathfrak{h}_2 \supset \mathfrak{h}^M$ and hence $(\mathfrak{h}^M)' = [\mathfrak{h}^M, \mathfrak{h}^M] \supset \mathfrak{h}'$.

Since $\mathfrak{h} \subset \mathfrak{h}^M \Rightarrow \mathfrak{h}' \subset (\mathfrak{h}^M)'$, this proves Lemma 3.

### 1.4 Additional facts.

Below, we formulate three classical theorem from the Lie theory with short comments concerning their proofs. The reader can consider this section as a sequence of challenging exercises.

**1.4.1. The Ado theorem.** *Every finite dimensional real or complex Lie algebra is isomorphic to a Lie subalgebra of the Lie algebra* $\mathfrak{gl}(N, \mathbb{R}$ or $\mathbb{C})$.

Differently speaking, this means that every finite-dimensional Lie algebra $\mathfrak{g}$ has an *exact* finite dimensional representation, that is, such a representation $\rho: \mathfrak{g} \to \mathfrak{gl}(V)$ that for every non-zero $\xi \in \mathfrak{g}$ there exists a $v \in V$ such that $(\rho(\xi))(v) \neq 0$. In many important cases it is obvious. For example, if $\mathfrak{g}$ has no center, then the adjoint representation $\mathrm{ad}: \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ has this property. Also, it is not hard to construct an exact *infinite dimensional* representation: this is the canonical representation of $\mathfrak{g}$ in the so called *universal enveloping algebra* of $\mathfrak{g}$ (which we will have to consider later). The proof in the general case is more involved, but the reader can try to find it (if nowhere else, then in the literature).

We must notice that there is no similar result for Lie groups. Namely, there are Lie groups not embeddable into groups $GL(N, \mathbb{R})$ for any $N$. The first example arises in dimension 3: the universal covering over the Lie group $SL(2, \mathbb{R})$ is not Lie isomorphic to a Lie subgroup of a group $GL(N, \mathbb{R})$, whatever $N$ is.

**1.4.2. The third Lie theorem.** I prefer to avoid answering the question what the First and the Second Lie theorems are; certainly, both are covered by this course.

THEOREM. *Every finite-dimensional (real) Lie algebra is a Lie algebra of some Lie group.*

COMMENTS. (1) Certainly, this follows from the Ado theorem: a Lie subalgebra of $\mathfrak{gl}(N, \mathbb{R})$ is a Lie algebra of some virtual Lie subgroup of $GL(N, \mathbb{R})$, hence, it is a Lie algebra of some Lie group.

(2) For a Lie algebra $\mathfrak{g}$ without center, there is a simple proof. Let $\mathrm{Aut}\,\mathfrak{g}$ be the group of automorphisms of the Lie algebra $\mathfrak{g}$; this is a closed subgroup of $GL(\mathfrak{g})$ and hence a Lie group. It is easy to show that $\mathrm{Lie}\,\mathrm{Aut}\,\mathfrak{g}$ is $\mathrm{Der}\,\mathfrak{g}$, the Lie algebra of derivations of $\mathfrak{g}$ (a *derivation* of $\mathfrak{g}$ is a linear map $d: \mathfrak{g} \to \mathfrak{g}$ such that $d[\xi, \eta] = [d\xi, \eta] + [\xi, d\eta]$; it is easy to check that the commutator of two derivations is a derivation which makes $\mathrm{Der}\,\mathfrak{g}$ a Lie algebra: a Lie subalgebra of $\mathfrak{gl}(\mathfrak{g})$). For any $\xi \in \mathfrak{g}$, there is the *inner derivation* of $\mathfrak{g}$: $\eta \mapsto [\xi, \eta]$ (it is a derivation by the Jacobi identity). Thee arises a homomorphism $\mathfrak{g} \to \mathrm{Der}\,\mathfrak{g}$ which is one-to-one, if $\mathfrak{g}$ has no center. Thus, in this case, $\mathfrak{g} \subset \mathrm{Der}\,\mathfrak{g} = \mathrm{Lie}\,\mathrm{Aut}\,\mathfrak{g}$ which makes $\mathfrak{g}$ a Lie algebra of a virtual Lie subgroup of $\mathrm{Aut}\,\mathfrak{g}$.

(3) A proof of the Third Lie theorem in the general case is based on the notion of an universal enveloping algebra $U(\mathfrak{g})$ of $\mathfrak{g}$. There is a canonical Hopf algebra structure in $U(\mathfrak{g})$, and the set $\{x \in U(\mathfrak{g}) \mid \Delta(x) = x \otimes x\}$ (where $\Delta$ is the "comultiplication") has a structure of a Lie group with the Lie algebra $\mathfrak{g}$.

**1.4.3. The Baker-Campbell-Hausdorff formula.** The exponential map identification of a neighborhood of $0 \in \mathfrak{g}$ with a neighborhood of $e \in G$ makes the multiplication in $G$ a locally defined operation in $\mathfrak{g}$:

$$\xi \circ \eta = \exp^{-1}(\exp \xi \exp \eta).$$

This operation is determined by the commutator operation in $\mathfrak{g}$, and, no wonder, can be presented by a series involving the commutators. There exists an explicit formula for this series:

$$\xi \circ \eta = \sum_{m>0} \sum_{\substack{(p_1,\dots,p_m;q_1,\dots,q_m) \\ p_i \geq 0, q_i \geq 0, p_i+q_i>0}} \frac{(-1)^{m-1}\,[\xi^{p_1}, [\eta^{q_1}, \dots [\xi^{p_m}, \eta^{q_m}]\dots]]}{m \sum_i (p_i + q_i) \prod_i (p_i! q_i!)}$$

where $[x^k, y] = (\operatorname{ad} x)^k y$ and it is supposed that either $q_m = 1$ or $p_m = 1, q_m = 0$ (in which case $[\xi^{p_m}, \eta^{q_m}] = \xi$). This formula is called the Baker-Campbell-Hausdorff formula (in this form it was first written by E. B. Dynkin ("A computation of the coefficients in the Campbell-Hausdorff formula" (Russian), *Doklady AN SSSR*, **57** (1947), 323–326.)

The main inconvenience arising when one tries to use the formula above is that it contains a lot of like terms. For example, the part of this sum consisting of terms with $\sum_i (p_i + q_i)$ is

$$\begin{aligned} \eta + \xi + \frac{1}{2}[\xi, \eta] - \frac{1}{4}[\xi, \eta] - \frac{1}{4}[\eta, \xi] + \frac{1}{6}[\xi, [\xi, \eta]] - \frac{1}{6}[\eta, [\xi, \eta]] - \frac{1}{6}[\xi, [\xi, \eta]] - \frac{1}{12}[\xi, [\xi, \eta]] \\ + \frac{1}{9}[\eta, [\xi, \eta]] + \frac{1}{9}[\xi, [\xi, \eta]] - \frac{1}{12}[\eta, [\eta, \xi]] - \frac{1}{6}[[\xi, [\eta, \xi]] + \frac{1}{9}[\eta, [\eta, \xi]] + \frac{1}{9}[\xi, [\eta, \xi]] \end{aligned}$$

(we omitted terms with $[\xi, \xi]$ and $[\eta, \eta]$) which gives, after combining like terms,

$$\xi \circ \eta = \xi + \eta + \frac{1}{2}[\xi, \eta] + \frac{1}{12}[\xi, [\xi, \eta]] - \frac{1}{12}[\eta, [\xi, \eta]] + \dots$$

(in which form the formula is contained in many textbooks). By the way, to prove the last formula, we need only to check, up to the terms of degree 3, the equality

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

for non-commuting variables $X, Y$ using the usual series for exp. The left hand side (up to degree 3) is

$$\begin{aligned} \exp(X) \exp(Y) &= \left(1 + X + \frac{X^2}{2} + \frac{X^3}{6}\right)\left(1 + Y + \frac{Y^2}{2} + \frac{Y^3}{6}\right) \\ &= 1 + X + Y + \frac{X^2 + 2XY + Y^2}{2} + \frac{X^3 + 3X^2Y + 3XY^2 + Y^3}{6}; \end{aligned}$$

the right hand side (again, up to degree 3) is

$$1 + \left(X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]]\right)$$
$$+ \frac{1}{2}\left(X^2 + XY + \frac{1}{2}X[X,Y] + YX + Y^2 + \frac{1}{2}Y[X,Y] + \frac{1}{2}[X,Y]X + \frac{1}{2}[X,Y]Y\right)$$
$$+ \frac{1}{6}(X^3 + X^2Y + XYX + YX^2 + XY^2 + YXY + Y^2X + Y^3) =$$

$$\begin{aligned}1 &+ (X+Y) + \frac{XY - YX + X^2 + XY + YX + Y^2}{2}\\ &+ \frac{X^2Y - XYX - XYX + YX^2}{12} - \frac{YXY - Y^2X - XY^2 + YXY}{12}\\ &+ \frac{X^2Y - XYX + YXY - Y^2X + XYX - YX^2 + XY^2 - YXY}{4}\\ &+ \frac{X^3 + X^2Y + XYX + YX^2 + XY^2 + YXY + Y^2X + Y^3}{6},\end{aligned}$$

and, after combining the like terms, the two sides become the same.

## 2. Nilpotent, solvable, and semisimple Lie algebras.

It is seen from the results of the first part (and will be still more clear later) that in many cases the study of Lie groups, which may involve deep results from topology and analysis, can be reduced to a pure algebraic study of Lie algebras. This part is fully devoted to Lie algebras, Lie groups will not even mentioned. All vector spaces (in particular, all Lie algebras and their representations) considered will be finite dimensional, the ground field will be $\mathbb{R}$ or $\mathbb{C}$; we can use a generic notation $\mathbb{K}$ for this ground field.

Speaking of a representation $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ of a Lie algebra, we will often omit the notation of the representation: for $\xi \in \mathfrak{g}$ and $v \in V$, we may use a brief notation $\xi v$ for $(\rho(\xi))(v)$.

### 2.1. Nilpotent Lie algebras.

**2.1.1. Definitions.** A linear operator $A\colon V \to V$ is called nilpotent, if $A^k = 0$ for some $k$. In other words, $A$ is nilpotent, if and only if all eigenvalues of $A$ are zeroes. In still other words, $A$ is nilpotent, if with respect to some basis of $V$ the matrix of $A$ is strictly upper triangular,

$$\begin{pmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & 0 & a_{23} & \cdots & a_{2n} \\ 0 & 0 & 0 & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

If $W \subset V$ is a subspace invariant with respect to a nilpotent operator $A$, then the restriction $W \to W$ and the quotient $V/W \to V/W$ of $A$ are also nilpotent.

Exercise 1. Prove that if the operator $A\colon V \to V$ is nilpotent, then $A^k = 0$ for every $k \geq \dim V$.

A representation $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ of a Lie algebra $\mathfrak{g}$ is called nilpotent, if $\rho(\xi)$ is a nilpotent operator for every $\xi \in \mathfrak{g}$. It is clear that any subrepresentation or quotientrepresentation of a nilpotent representation of a Lie algebra, as well as a restriction of a nilpotent representation to a Lie subalgebra, is nilpotent.

A Lie algebra $\mathfrak{g}$ is called nilpotent, if the adjoint representation $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is nilpotent. In other words, $\mathfrak{g}$ is nilpotent, if there exists a $k$ such that for every $\xi, \eta \in \mathfrak{g}$,

$$[\underbrace{\xi, \dots [\xi}_{k}, \eta] \dots] = 0.$$

Exercise 2. Show that (a) not every representation of a nilpotent Lie algebra must be nilpotent; (b) a non-nilpotent Lie algebra may have nilpotent representations.

Still, the following is true.

Proposition 1. *If a Lie algebra possesses a faithful*[8]) *nilpotent representation, then it is nilpotent.*

*Proof.* Let $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a faithful nilpotent representation, let $\xi, \eta \in \mathfrak{g}$, and let $\rho(\xi) = A$, $\rho(\eta) = B$. Then $A = \rho(\xi)$ is nilpotent, so $A^k = 0$ for some $k$. For $\rho([\underbrace{\xi, \dots [\xi}_{m}, \eta] \dots]) = [\underbrace{A, \dots [A}_{m}, B] \dots]$, there is a formula

$$[\underbrace{A, \dots [A}_{m}, B] \dots] = \sum_{p=0}^{m} (-1)^p \binom{m}{p} A^{m-p} B A^p$$

(proved by an obvious induction; we leave this to the reader). The last expression is 0, if $m \geq 2k$. Thus, for such $m$, $[\underbrace{A, \dots [A}_{m}, B] \dots] = 0$ and hence $([\underbrace{\xi, \dots [\xi}_{m}, \eta] \dots]) = 0$ (since $\rho$ is faithful).

It is obvious that a Lie subalgebra of a nilpotent Lie algebra and a quotient of a nilpotent Lie algebra over an ideal are nilpotent Lie algebras. The following is also true.

Proposition 2. *If a quotient of a Lie algebra over a central ideal is nilpotent, then it is nilpotent itself.*

*Proof.* Let $\mathfrak{i}$ be a central ideal of a Lie algebra $\mathfrak{g}$, and let $\mathfrak{g}/\mathfrak{i}$ be nilpotent. Let $\xi, \eta \in \mathfrak{g}$ and let $\overline{\xi}$ and $\overline{\eta}$ be projections of $\xi$ and $\eta$ onto $\mathfrak{g}/\mathfrak{i}$. Then, for some $k$, $(\mathrm{ad}\,\overline{\xi})^k \overline{\eta} = 0$. This means that $(\mathrm{ad}\,\xi)^k \eta \in \mathfrak{i}$ and hence $(\mathrm{ad}\,\xi)^{k+1} \eta = [\xi, (\mathrm{ad}\,\xi)^k \eta] = 0$.

**2.1.2. The Engel Theorem.** *Let $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a nilpotent representation. Then there exists a chain of subspaces*

$$0 \underset{\neq}{\subset} V_0 \underset{\neq}{\subset} V_1 \underset{\neq}{\subset} V_2 \dots \underset{\neq}{\subset} V_k = V$$

[8]) A representation $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is called *faithful*, if $\operatorname{Ker} \rho = 0$, that is, if for every non-zero $\xi \in \mathfrak{g}$, there exists a $v \in V$ such that $(\rho(\xi))(v) \neq 0$.

*such that* $\mathfrak{g}V_i \subset V_{i-1}$. *In other words, there exists a* $k \leq \dim V + 1$ *such that for every* $\xi_1, \dots, \xi_k \in \mathfrak{g}$ *and* $v \in V$,

$$\xi_1 \dots \xi_k v = 0.$$

*In still other words, a basis in* $V$ *with respect to which the operator* $\rho(\xi)$ *has a upper strictly triangular matrix may be chosen simultaneously for all* $\xi \in \mathfrak{g}$.

Applying the Engel Theorem to the representation ad, we get the following

COROLLARY 1. *For a nilpotent Lie algebra* $\mathfrak{g}$, *there exists a chain of ideals*

$$0 \underset{\neq}{\subset} \mathfrak{g}_0 \underset{\neq}{\subset} \mathfrak{g}_1 \underset{\neq}{\subset} \mathfrak{g}_2 \dots \underset{\neq}{\subset} \mathfrak{g}_k = \mathfrak{g}$$

*such that* $[\mathfrak{g}, \mathfrak{g}_i] \subset \mathfrak{g}_{i-1}$. *In other words, there exists a* $k \leq \dim \mathfrak{g} + 1$ *such that for every* $\xi_1, \dots, \xi_k \in \mathfrak{g}$,

$$[\xi_1, [\xi_2, \dots [\xi_{k-1}, \xi_k] \dots]] = 0.$$

Notice that the Engel Theorem and Corollary 1 provide *criteria* of nilpotency of a Lie algebra or of a representation: they can be formulated in the form: a Lie algebra, or a representation of a Lie algebra, is nilpotent if and only if, etc.

A proof of the Engel Theorem is based on the following lemma (in some books, the term “Engel Theorem” is attributed to this lemma).

LEMMA. *For every nilpotent representation of a Lie algebra* $\mathfrak{g}$ *in a space* $V$, *there exists a non-zero vector* $v \in V$ *such that* $\xi v = 0$ *for all* $\xi \in \mathfrak{g}$.

*Proof of Lemma.* We can assume that the representation is faithful: if $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is not faithful, we can consider, instead of $\rho$ the representation of $\mathfrak{g}/\operatorname{Ker}\rho$. Thus, the Lie algebra $\mathfrak{g}$ is nilpotent by Proposition 1.

We use the induction with respect to $\dim \mathfrak{g}$. If $\dim \mathfrak{g} = 1$, then a representation of $\mathfrak{g}$ is the same as one nilpotent operator, and there is a non-zero vector which is annihilated by this operator. Assume that Lemma is true for representations of nilpotent Lie algebras of dimensions $< \dim \mathfrak{g}$.

We begin with proving the following *auxiliary statement*: the Lie algebra $\mathfrak{g}$ has a codimension 1 ideal. Take an arbitrary Lie subalgebra $\mathfrak{b} \underset{\neq}{\subset} \mathfrak{g}$; then there exists a Lie subalgebra $\mathfrak{c}$ of $\mathfrak{g}$ such that $\mathfrak{b} \subset \mathfrak{c} \subset \mathfrak{g}$, $\operatorname{codim}_{\mathfrak{c}} \mathfrak{b} = 1$, $\mathfrak{b}$ is an ideal in $\mathfrak{c}$. Indeed, the representation of $\mathfrak{b}$ in $\mathfrak{g}/\mathfrak{b}$ is nilpotent; thus, there is a non-zero vector $v \in \mathfrak{g}/\mathfrak{b}$ which is annihilated by $\mathfrak{b}$. Then, for a representative $\widetilde{v} \in \mathfrak{g} - \mathfrak{b}$ of $v$, $[\mathfrak{b}, \widetilde{v}] \in \mathfrak{b}$ and we can put $\mathfrak{c} = \mathfrak{b} + \mathbb{K}\widetilde{v}$. Now, the prove our auxiliary statement, we take an arbitrary $\mathfrak{b}$ (for example, 0) and apply this construction sufficiently many times: $\mathfrak{b} \subset \mathfrak{c}_1 \subset \mathfrak{c}_2 \subset \dots \subset \mathfrak{c}_{m-1} \subset \mathfrak{c}_m = \mathfrak{g}$ $(m = \dim \mathfrak{g} - \dim \mathfrak{b})$. Then $\mathfrak{h} = \mathfrak{c}_{m-1}$ is a codimension 1 ideal in $\mathfrak{g} = \mathfrak{c}_m$.

Now, return to Lemma. Let $\mathfrak{h}$ be a codimension 1 ideal in $\mathfrak{g}$. Let then $W \subset V$ be the space of all vectors which are annihilated by $\mathfrak{h}$; then $\dim W > 0$. For a $\xi \in \mathfrak{g} - \mathfrak{h}$, $\xi(W) \subset W$; indeed, for $w \in W$, $\xi \in \mathfrak{g}$, and $\eta \in \mathfrak{h}$, $\eta(\xi w) = \xi(\eta w) + [\eta, \xi]w = 0$ since $\eta, [\eta, \xi] \in \mathfrak{h}$; thus, $\xi w \in W$. So, $\xi\colon W \to W$ is a nilpotent operator, and hence there is a non-zero vector $w \in W$ which is annihilated by $\xi$. Then $\mathfrak{g}w = 0$.

*Proof of the Engel Theorem.* Let $V_0$ is the space of 0-vectors of the representation $\rho$. By lemma, $V_0 \neq 0$. There arises a representation of $\mathfrak{g}$ in $V/V_0$, also nilpotent. Take the space of zero vectors of this representation, and let $V_1$ be the inverse image of this space in $V$. Then $\mathfrak{g}(V_1) \subset V_0$. In particular, $V_1$ is an invariant subspace of $V$. There arises a representation of $\mathfrak{g}$ in $V/V_1$, take for $V_2$ the inverse image of this space in $V$. And so on.

**2.1.3. Further corollaries.** COROLLARY 2. *Every nilpotent Lie algebra has a non-zero center.*

This is $\mathfrak{g}_0$ from Corollary 1.

If we iterate this statement, we arrive at the following result.

COROLLARY 3. *A Lie algebra is nilpotent if and only if it is obtained from* 0 *by a sequence of central extensions*[9]).

And the last one.

COROLLARY 4. *A nilpotent Lie algebra* $\mathfrak{g}$ *contains ideals of all dimensions between* 1 *and* $\dim \mathfrak{g} - 1$.

Indeed, every subspace of $\mathfrak{g}$ between $\mathfrak{g}_i$ and $\mathfrak{g}_{i-1}$ (where $\mathfrak{g}_{-1} = 0$) is an ideal of $\mathfrak{g}$.

**2.2. A deviation: some generalities concerning Lie algebra representations.**

**2.2.1. Irreducible representations.** A representation $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ of real or complex Lie algebra is called *irreducible*, if there are no proper subrepresentations, that is, proper subspaces of $V$ invariant with respect to $\mathfrak{g}$. For example, the adjoint representation $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is irreducible if and only if $\mathfrak{g}$ is *simple*, that is, does not contain proper ideals.

PROPOSITION. *Let* $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ *be a (real or complex) representation. Then there exists a chain*

$$0 = V_0 \subset V_1 \subset \ldots \subset V_{r-1} \subset V_r = V$$

*of subrepresentations such that every quotient* $V_i/V_{i-1}$ *is irreducible.*

*Proof*: induction by $\dim V$. If $\dim V = 1$, then $\rho$ itself is irreducible, and we have nothing to do. Assume that Proposition is true for all representations of $\mathfrak{g}$ of dimensions $< \dim V$. Let $\dim V > 1$. Again, if $\rho$ is irreducible, then we have nothing to do. If $\rho$ is reducible, then there are proper invariant subspaces in $V$. Choose a proper invariant subspace $W \subset V$ of maximal possible dimension. On one hand, the quotient $V/W$ is irreducible: otherwise, there is a proper invariant subspace $Z \subset V/W$, and the inverse image of $Z$ with respect to the projection $V \to V/W$ is a proper invariant subspace of $V$ with $\dim Z > \dim W$, in contradiction of our choice of $W$. On the other hand, be the induction hypothesis, there exists a chain

$$0 = W_0 \subset W_1 \subset \ldots \subset W_{s-1} \subset W_s = W$$

with irreducible quotients $W_i/W_{i-1}$, and we can take $r = s+1$, $V_i = W_i$ for $i = 0, \ldots, r-1$, $V_r = V$.

**2.2.2. Weights and weight spaces.** Let $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a complex representation of a complex Lie algebra $\mathfrak{g}$. For a linear function $\lambda\colon \mathfrak{g} \to \mathbb{C}$, put

$$V_\lambda = \{v \in V \mid \rho(\xi)v = \lambda(\xi) \cdot v \text{ for all } \xi \in \mathfrak{g}\}.$$

[9]) A Lie algebra $\widetilde{\mathfrak{g}}$ is called a *central extension* of a Lie algebra $\mathfrak{g}$, if there is a fixed isomorphism $\mathfrak{g} \cong \widetilde{\mathfrak{g}}/\mathfrak{c}$ where $\mathfrak{c}$ is a central ideal of $\widetilde{\mathfrak{g}}$ (that is, $[\mathfrak{c}, \widetilde{\mathfrak{g}}] = 0$).

If $V_\lambda \neq 0$, then $\lambda$ is called a *weight* for $\rho$, and $V_\lambda$ is called a *weight space*.

PROPOSITION. *Let $\mathfrak{g}, \rho$, and $V$ be as above and let $\mathfrak{h}$ be an ideal in $\mathfrak{g}$. If $\lambda\colon \mathfrak{h} \to \mathbb{C}$ is a weight of the representation $\rho|_{\mathfrak{h}}$, then the space $V_\lambda$ is invariant with respect to $\mathfrak{g}$, that is, $\rho(\xi)V_\lambda \subset V_\lambda$ for every $\xi \in \mathfrak{g}$.*

*Proof.* Let $v \in V$ be a non-zero vector, and let $\xi \in \mathfrak{g}, \eta \in \mathfrak{h}$. Then

$$\rho(\eta)(\rho(\xi)v) = \rho(\xi)(\rho(\eta)v) + \rho([\xi,\eta])v = \lambda(\eta) \cdot \rho(\xi)v + \lambda([\xi,\eta]) \cdot v$$

(we observe that $[\xi,\eta] \in \mathfrak{h}$, since $\mathfrak{h}$ is an ideal in $\mathfrak{g}$). To prove Proposition, we need to show that $\lambda([\xi,\eta]) = 0$ for all $\xi \in \mathfrak{g}, \eta \in \mathfrak{h}$.

Let $m$ be the biggest integer for which the vectors $v, \rho(\xi)v, \rho(\xi)^2v, \ldots, \rho(\xi)^m v$ are linearly independent, and let $W = \operatorname{span}(v, \rho(\xi)v, \rho(\xi)^2v, \ldots, \rho(\xi)^m v)$. Let us prove that $W$ is invariant with respect to $\mathfrak{h}$ and, moreover, for every $\eta \in \mathfrak{h}$ the matrix of the transformation $W \xrightarrow{\rho(\eta)} W$ with respect to the basis $v, \rho(\xi)v, \rho(\xi)^2v, \ldots, \rho(\xi)^m v$ is an upper triangular matrix with all the diagonal entries being $\lambda(\eta)$.

First of all, $\rho(\eta)v = \lambda(\eta)\cdot v$, so the first column of the matrix is $(\lambda(\eta), 0, \ldots, 0)$. Then, by induction on the columns,

$$\rho(\eta)(\rho(\xi)^{i+1}v) = \rho([\xi,\eta])(\rho(\xi)^i v + \rho(\xi)(\rho(\eta)(\rho(\xi)^i v)) = \lambda(\eta) \cdot \rho(\xi)^{i+1}v + u$$

where $u \in \operatorname{span}(v, \rho(\xi)v, \ldots, \rho(\xi)^i v$ because $[\xi,\eta] \in \mathfrak{h}$ and the induction hypothesis.

Thus, the space $W$ is invariant with respect to $\mathfrak{h}$ and $\xi$; so, it is invariant with respect to $\mathfrak{h}+\operatorname{span}(\xi)$. For every $\eta \in \mathfrak{h}$, the commutator $[\xi,\eta]$ is contained in $\mathfrak{h}$, so the matrix of its action in $W$ is an upper triangular matrix with the diagonal entries $\lambda([\xi,\eta]$. On the other hand, this matrix is the commutator of the matrices of transformations $\rho(\xi)$ and $\rho(\eta)$, hence its trace is zero. Therefore, $\lambda([\xi,\eta]) = 0$, which completes the proof of Proposition.

**2.3. Solvable Lie algebras.** A Lie algebra $\mathfrak{g}$ is called solvable, if the sequence

$$\mathfrak{g}, \mathfrak{g}', (\mathfrak{g}')', ((\mathfrak{g}')')', \ldots$$

ends with 0. Equivalently: $\mathfrak{g}$ is solvable, if there exists a chain of subalgebras

$$0 = \mathfrak{g}_0 \subset \mathfrak{g}_1 \subset \ldots \subset \mathfrak{g}_r = \mathfrak{g}$$

such that for every $i = 1, \ldots, r$, $\mathfrak{g}_{i-1}$ is an ideal in $\mathfrak{g}_i$ and the quotient Lie algebra $\mathfrak{g}_i/\mathfrak{g}_{i-1}$ is commutative. Example: the Lie algebra $\mathfrak{b} \subset \mathfrak{gl}(n, \mathbb{R}$ or $\mathbb{C})$ of all upper triangular matrices.

In particular, every nilpotent Lie algebra is solvable. However, not every solvable Lie algebra is nilpotent, as the following exercise shows.

EXERCISES 3. Prove that the Lie algebra $\mathfrak{b}$ from the example above is not nilpotent.

4. Let $\mathfrak{h}$ be an ideal in a Lie algebra $\mathfrak{g}$. Prove that if $\mathfrak{g}$ is solvable, then $\mathfrak{h}$ and $\mathfrak{g}/\mathfrak{h}$ are solvable. Prove also that if $\mathfrak{h}$ and $\mathfrak{g}/\mathfrak{h}$ are solvable, then $\mathfrak{g}$ is solvable.

**2.3.1. Lie's theorem.** *Let $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of a* **complex** *solvable Lie algebra $\mathfrak{g}$ in a complex vector space $V$. Then there exists a basis $e_1, \ldots, e_n$ in $V$ such* *that for every $\xi \in \mathfrak{g}$, the matrix of the transformation $\rho(\xi): V \to V$ with respect to this basis is upper triangular, that is, $(\rho(\xi))(e_k) \in \operatorname{span}(e_1, \ldots, e_k)$.*

REMARK. For *real* Lie algebras a similar statement does not hold. For example, if $\dim \mathfrak{g} = 1$, then $\mathfrak{g}$ is commutative, and hence, solvable. But a representation $\rho: \mathfrak{g} \to \mathfrak{gl}(V)$ is reduced to one operator $\rho(\xi)$, $\xi \neq 0$, which is arbitrary. If not all the eigenvalues of this operator are real, then it can be represented by an upper triangular matrix in any basis.

Like the Engel Theorem, the Lie Theorem is almost equivalent to a lemma similar to the lemma in the proof of the Engel Theorem. Sometimes, the term "the Lie Theorem" is attributed to this lemma.

LEMMA. *Let $\rho: \mathfrak{g} \to \mathfrak{gl}(V)$ be a complex representation of a complex solvable Lie algebra. Then there exists a linear function $\lambda: \mathfrak{g} \to \mathbb{C}$ and a non-zero vector $v \in V$ such that for any $\xi \in \mathfrak{g}$, $(\rho(\xi))(v) = \lambda(\xi) \cdot v$. In other word, $\rho$ possesses at least one weight.*

*Proof of Lemma*: induction by $\dim \mathfrak{g}$. If $\dim \mathfrak{g} = 1$, then, for a fixed $\xi \in \mathfrak{g}$, there exists a non-zero $v \in V$ and a $\lambda \in \mathbb{C}$, such that $\rho(\xi)v = \lambda v$. We put $\lambda(c\xi) = c\lambda$ and have $\rho(c\xi)v = c\rho(\xi)v = c\lambda v = \lambda(c\xi)v$. Assume that $\dim \mathfrak{g} > 1$ and that Lemma holds for all nilpotent Lie algebras of dimensions $< \dim \mathfrak{g}$.

Since $[\mathfrak{g}, \mathfrak{g}] \underset{\neq}{\subset} \mathfrak{g}$, there exists a subspace $\mathfrak{h}$ of $\mathfrak{g}$ such that $\mathfrak{h} \supset [\mathfrak{g}, \mathfrak{g}]$ and $\dim \mathfrak{h} = \dim \mathfrak{g} - 1$. Then $\mathfrak{h}$ is an ideal, in particular, a Lie subalgebra, in $\mathfrak{g}$: $[\mathfrak{g}, \mathfrak{h}] \subset [\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{h}$. Thus, $\mathfrak{h}$ is a solvable Lie algebra, and, by the induction hypothesis, there exists a weight $\mu: \mathfrak{h} \to \mathbb{C}$ with a non-zero weight space $W_\mu = \{v \in V \mid \rho(\eta)v = \mu(\eta) \cdot v \text{ for all } \eta \in \mathfrak{h}\}$.

Let $\zeta \in \mathfrak{g} - \mathfrak{h}$. By Proposition in Section 2.2.2, $W_\mu$ is $\rho(\xi)$-invariant. The transformation $W_\mu \xrightarrow{\rho(\zeta)} W_\mu$ has an eigenvector $w \in W_\mu$, $\rho(\zeta)w = \nu \cdot w$, $\nu \in \mathbb{C}$. Define $\lambda: \mathfrak{g} = \mathfrak{h} \oplus \mathbb{C}\zeta \to \mathbb{C}$ by the formula $\lambda(\eta + c\zeta) = \mu(\eta) + c\nu$ ($\eta \in \mathfrak{h}$, $c \in \mathbb{C}$). Then $\rho(\eta + c\zeta)w = \rho(\eta)w + c\rho(\zeta)w = (\mu(\eta) + c\nu)w = \lambda(\eta + c\zeta) \cdot w$ which completes the proof of Lemma.

*Proof of Lie's theorem:* induction with respect to $\dim V$. If $\dim V = 1$, then we have nothing to prove, since every $1 \times 1$ matrix is upper triangular. Assume now that Theorem holds for all representations of $\mathfrak{g}$ of dimensions $< n = \dim V$. By Lemma, there exists a vector $e_1 \in V$ such that $\rho(\xi)e_1$ is a multiple of $e_1$, and, in particular, the space $\mathbb{C}e_1 = V_1 \subset V$ is invariant with respect to the action of $\mathfrak{g}$. In other words, $V_1$ is a subrepresentation of $V$, and we can form the quotient representation $\sigma$ of $\mathfrak{g}$ in $W = V/V_1$ with a $\mathfrak{g}$-projection $p: V \to V/W$. By the induction hypothesis, there exists a basis $f_1, \ldots, f_{n-1}$ such that $\sigma(\xi)f_\ell \in \operatorname{span}(f_1, \ldots, f_\ell)$. Choose $e_2, \ldots, e_n \in V$ such that $p(e_i) = f_{i-1}$ for $i = 2, \ldots n$. Obviously, $e_1, \ldots, e_n$ is a basis of $V$ with the properties required.

**2.3.2. Relations between solvability and nilpotency.** We already know that every nilpotent Lie algebra is solvable. Now we can say more.

PROPOSITION 1. *Let $\rho$ be a representation of a solvable Lie algebra $\mathfrak{g}$ over $\mathbb{C}$ or $\mathbb{R}$. Then $\mathfrak{n} = \{\xi \in \mathfrak{g} \mid \rho(\xi) \text{ is nilpotent}\}$ is an ideal in $\mathfrak{g}$ containing $\mathfrak{g}'$.*

*Proof: the case of* $\mathbb{C}$. By the Lie Theorem, there exists a basis in the space $V$ of the representation, such that

$$\text{the matrix of } \rho(\xi) = \begin{pmatrix} \lambda_1(\xi) & & * \\ & \ddots & \\ 0 & & \lambda_n(\xi) \end{pmatrix}$$

This matrix is nilpotent, that is, $\xi \in \mathfrak{n}$, *if and only if* all $\lambda_i(\xi) = 0$; *if* $\xi \in \mathfrak{g}'$, then our matrix is the commutator of matrices of the same form, and hence all $\lambda_i(\xi) = 0$. Hence, $\mathfrak{n} \supset \mathfrak{g}'$, and hence $\mathfrak{n}$ is an ideal in $\mathfrak{g}$.

*The case of* $\mathbb{R}$. Apply the complexification. Let $\mathfrak{g}_{\mathbb{C}} = \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$; then $(\mathfrak{g}_{\mathbb{C}})' = (\mathfrak{g}')_{\mathbb{C}}$ and $\mathfrak{g}' = \mathfrak{g}'_{\mathbb{C}} \cap \mathfrak{g}$. Let also $\rho_{\mathbb{C}} = \rho \otimes_{\mathbb{R}} \mathbb{C}$. We have already proved that

$$\{\xi \in \mathfrak{g}_{\mathbb{C}} \mid \rho_{\mathbb{C}}(\xi) \text{ is nilpotent}\} \supset (\mathfrak{g}_{\mathbb{C}})'.$$

Hence,

$$\{\xi \in \mathfrak{g} \mid \rho(\xi) \text{ is nilpotent}\} = \{\xi \in \mathfrak{g}_{\mathbb{C}} \mid \rho_{\mathbb{C}}(\xi) \text{ is nilpotent}\} \cap \mathfrak{g} \supset (\mathfrak{g}_{\mathbb{C}})' \cap \mathfrak{g} = \mathfrak{g}'.$$

PROPOSITION 2. *A (real or complex) Lie algebra $\mathfrak{g}$ is solvable, if and only if $\mathfrak{g}'$ is nilpotent.*

*Proof.* If $\mathfrak{g}'$ is nilpotent, then it is solvable and the sequence $\mathfrak{g}', (\mathfrak{g}')', \dots$ ends with zero; hence, $\mathfrak{g}$ is solvable. If $\mathfrak{g}$ is solvable, then $\mathfrak{g}'$ is contained in $\{\xi \in \mathfrak{g} \mid \operatorname{ad}(\xi) \text{ is nilpotent}\}$. Thus, $\mathfrak{g}'$ is nilpotent.

**2.3.3. Nilpotent radicals.** PROPOSITION. *Let $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of a (complex or real) Lie algebra. Let $\mathcal{N}$ be the set of all ideals $\mathfrak{n} \subset \mathfrak{g}$ such that $\rho(\xi)$ is nilpotent for every $\xi \in \mathfrak{n}$. Then there exists an $\mathfrak{n}_\rho \in \mathcal{N}$ such that $n \in \mathcal{N}$ if and only if $\mathfrak{n} \subset \mathfrak{n}_\rho$.*

*Proof. Case 1: $\rho$ is irreducible.* In this case $\mathfrak{n} \in \mathcal{N}$ if and only if $\rho_{\mathfrak{n}} = 0$. Indeed, since the representation $\rho_{\mathfrak{n}}$ is nilpotent, the space $W = \{v \in V \mid \rho(\eta)v = 0 \text{ for all } \eta \in \mathfrak{n}\}$ is not zero by Engel's theorem. This space is $\mathfrak{g}$-invariant: if $\xi \in \mathfrak{g}, \eta \in \mathfrak{n}$, and $w \in W$, then $\rho(\eta)(\rho(\xi)w) = \rho(\xi)(\rho(\eta)w) + \rho([\eta, \xi])w = 0$ ($[\eta, \xi] \in \mathfrak{n}$, because $\mathfrak{n}$ is an ideal; hence, $W = V$ and $\rho|_{\mathfrak{n}} = 0$. This shows that we can put $\mathfrak{n}_\rho = \operatorname{Ker} \rho$.

*Case 2: general.* According to Proposition of Section 2.2.1, there is a chain of subrepresentations

$$0 = V_0 \subset V_1 \subset \dots \subset V_{r-1} \subset V_r = V$$

with irreducible subsequent quotients; we denote the representation in $V_i/V_{i-1}$ by $\rho_i$. Our claim is $\mathfrak{n}_\rho = \bigcap_i \operatorname{Ker} \rho_i$. Indeed, every $\operatorname{Ker} \rho_i$ is an ideal, hence $\bigcap_i \operatorname{Ker} \rho_i$ is an ideal. If $\xi \in \bigcap \rho_i$, then $\rho(\xi)(V_i) \subset V_{i-1}$, and hence $\rho(\xi)^r{=}0$. Hence, $\bigcap_i \operatorname{Ker} \rho_i \in \mathcal{N}$. On the other hand, as follows from the first part of the proof, $\rho_i|_{\mathfrak{n}} = 0$ for every $\mathfrak{n} \in \mathcal{N}$, which shows that $\mathfrak{n} \subset \bigcap_i \operatorname{Ker} \rho_i$.

DEFINITION. The ideal $\mathfrak{n}_{\mathrm{ad}}$ is called the *nilpotent radical* of $\mathfrak{g}$ and is denoted as $\operatorname{nil}\operatorname{rad}(\mathfrak{g})$.

REMARK. In general, it is *not* the set of all $\xi \in \mathfrak{g}$ with nilpotent $\operatorname{ad} \xi$.

EXERCISE 5. Prove that $\operatorname{nil}\operatorname{rad} \mathfrak{g}$ is a nilpotent ideal in $\mathfrak{g}$ which contains all other nilpotent ideals.

**2.3.4. Radicals.** PROPOSITION. *Let $\mathfrak{g}$ be a (real or complex) Lie algebra. If $\mathfrak{h}$ and $\mathfrak{k}$ are solvable ideals in $\mathfrak{g}$, then $\mathfrak{h}+\mathfrak{k}$ is also a solvable ideal.*

*Proof.* It is obvious that $\mathfrak{h}+\mathfrak{k}$ is an ideal. It is solvable, because $\mathfrak{h}\cap\mathfrak{k}$ and $(\mathfrak{h}+\mathfrak{k})/\mathfrak{h} = \mathfrak{k}/(\mathfrak{h}\cap\mathfrak{h})$ are solvable.

This proposition shows that every Lie algebra contains a unique maximal solvable ideal (which can be described, for example, as the sum of all solvable ideals). The maximal solvable ideal of a Lie algebra $\mathfrak{g}$ is called the *radical* of $\mathfrak{g}$ and is denoted as $\operatorname{rad}\mathfrak{g}$.

Obviously, $(\operatorname{rad}\mathfrak{g})' \subset \operatorname{nil}\operatorname{rad}\mathfrak{g} \subset \operatorname{rad}\mathfrak{g}$.

**2.3.5. Introducing semisimple Lie algebras.** A Lie algebra $\mathfrak{g}$ is called *semisimple* if it does not have non-zero solvable ideals. In other words, $\mathfrak{g}$ is semisimple, if $\operatorname{rad}\mathfrak{g} = 0$.

A simple Lie algebra does not have proper ideals at all; so, it is semisimple, if it is not solvable itself. But a solvable Lie algebra is simple, only if it is one-dimensional. Thus, a simple Lie algebra of dimension $> 1$ is semisimple. Also, the following holds.

PROPOSITION. *Let $\mathfrak{h}$ be an ideal in a Lie algebra $\mathfrak{g}$. If $\mathfrak{h}$ and $\mathfrak{g}/\mathfrak{h}$ are both semisimple, then $\mathfrak{g}$ is semisimple.*

*Proof.* Let $\mathfrak{k}$ be a solvable ideal in $\mathfrak{g}$. Then $\mathfrak{k}\cap\mathfrak{h}$ is an ideal in $\mathfrak{g}$, hence in both $\mathfrak{k}$ and $\mathfrak{h}$. The first implies that $\mathfrak{k}\cap\mathfrak{h}$ is solvable, after which the second implies that $\mathfrak{k}\cap\mathfrak{h} = 0$ (because $\mathfrak{h}$ is semisimple). Hence the image of $\mathfrak{k}$ in $\mathfrak{g}/\mathfrak{h}$ is an ideal isomorphic to $\mathfrak{k}$, which shows that $\mathfrak{k} = 0$ (because $\mathfrak{g}/\mathfrak{h}$ is semisimple).

COROLLARY. *A direct sum of semi-simple Lie algebras is semisimple.*

We will study semisimple Lie algebras in details, starting with Section 2.4. We will see, in particular, that a semisimple Lie algebra is always a direct sum of simple Lie algebras of dimensions $> 1$.

**2.4 Killing forms.**

**2.4.1. Definition and basic properties.** Let $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of a (real or complex) Lie algebra $\mathfrak{g}$. Put

$$B^\rho(\xi,\eta) = \operatorname{Tr}(\rho(\xi)\circ\rho(\eta)).$$

Obviously, $B^\rho$ is a symmetric bilinear form (notice that even in the complex case the form $B^\rho$ is *symmetric*, not *Hermitian*). It is called the *Killing form* of the representation $\rho$.

PROPOSITION 1. *The form $B^\rho$ possesses the following invariance property:*

$$B^\rho([\xi,\eta],\zeta) = B^\rho(\xi,[\eta,\zeta]) \text{ for every } \xi,\eta,\zeta \in \mathfrak{g}.$$

*Proof.* $B^\rho([\xi,\eta],\zeta) = \operatorname{Tr}(\rho(\xi)\circ\rho(\eta)\circ\rho(\zeta)) - \operatorname{Tr}(\rho(\eta)\circ\rho(\xi)\circ\rho(\zeta)) = \operatorname{Tr}(\rho(\xi)\circ\rho(\eta)\circ\rho(\zeta)) - \operatorname{Tr}(\rho(\xi)\circ\rho(\zeta)\circ\rho(\eta)) = B^\rho(\xi,[\eta,\zeta])$.

PROPOSITION 2. *The orthogonal complement $\mathfrak{h}^\perp = \mathfrak{h}^{\perp_\rho}$ of an ideal $\mathfrak{h}\subset\mathfrak{g}$ with respect to the Killing form $B^\rho$ is also an ideal.*

*Proof.* $\xi \in \mathfrak{h}^\perp$, if $B^\rho(\xi,\eta) = 0$ for all $\eta\in\mathfrak{h}$. But then, for a $\zeta\in\mathfrak{g}$, $B^\rho([\xi,\zeta],\eta) = B^\rho(\xi,[\zeta,\eta]) = 0$ (since $[\zeta,\eta]\in\mathfrak{h}$). Thus, $[\xi,\zeta]\in\mathfrak{h}^\perp$.

COROLLARY. *The kernel* $\operatorname{Ker} B^\rho$ *of the form $B^\rho$ is an ideal.*

Indeed, $\operatorname{Ker} B^\rho = 0^{\perp_\rho}$.

The form $B = B^{\mathrm{ad}}$ is called the *Killing form of the Lie algebra* $\mathfrak{g}$. We will use for it the notation $B_\mathfrak{g}$ and also notations $\langle\,,\,\rangle$ and $\langle\,,\,\rangle_\mathfrak{g}$.

The Killing form $B_\mathfrak{g}$ possesses a *strong invariance property* which generalizes, for this form, Proposition 1. Recall that a *derivation* of a Lie algebra $\mathfrak{g}$ is a linear map $D\colon \mathfrak{g} \to \mathfrak{g}$ such that $D[\xi,\eta] = [D\xi,\eta] + [\xi, D\eta]$ for all $\xi,\eta \in \mathfrak{g}$. For every $\zeta \in \mathfrak{g}$, $\operatorname{ad}\zeta\colon \mathfrak{g} \to \mathfrak{g}$ is a derivation (this is equivalent to the Jacobi identity); such derivations are called *inner*. In principle, Lie algebras may possess derivations which are not inner; for example, if $\mathfrak{g}$ is isomorphic an ideal in some Lie algebra $\mathfrak{f}$, then every $\zeta \in \mathfrak{f}$ gives rise to a derivation $\operatorname{ad}_\mathfrak{f}\zeta\colon \mathfrak{g} \to \mathfrak{g}$ which does not need to be inner. The following is a generalization of Proposition 1 from inner derivations to all derivations (only for the forms $B_\mathfrak{g}$).

Proposition 3. *For every derivation $D\colon \mathfrak{g} \to \mathfrak{g}$ and every $\xi,\eta \in \mathfrak{g}$,*

$$B_\mathfrak{g}(D\xi,\eta) = -B_\mathfrak{g}(\xi, D\eta).$$

*Proof.* By definition of a derivation, $\operatorname{ad}(D\xi) = D \circ \operatorname{ad}\xi - \operatorname{ad}\xi \circ D$. Hence,

$$\begin{aligned} B_\mathfrak{g}(D\xi,\eta) &= \operatorname{Tr}(\operatorname{ad}(D\xi)\circ\operatorname{ad}\eta) = \operatorname{Tr}(D\circ\operatorname{ad}\xi\circ\operatorname{ad}\eta) - \operatorname{Tr}(\operatorname{ad}\xi\circ D\circ\operatorname{ad}\eta)\\ &= \operatorname{Tr}(\operatorname{ad}\xi\circ\operatorname{ad}\eta\circ D) - \operatorname{Tr}(\operatorname{ad}\xi\circ D\circ\operatorname{ad}\eta) = -\operatorname{Tr}(\operatorname{ad}\xi\circ\operatorname{ad}(D\eta)) = -B_\mathfrak{g}(\xi,D\eta). \end{aligned}$$

Proposition 4. *If $\mathfrak{h}$ is an ideal of $\mathfrak{g}$, then $B_\mathfrak{h}$ is the restriction of $B_\mathfrak{g}$ to $\mathfrak{h}$.*

*Proof.* Choose a basis $e_1,\dots,e_m$ in $\mathfrak{h}$ and supplement it by $e_{m+1},\dots,e_n$ to a basis in $\mathfrak{g}$. Let $\xi \in \mathfrak{h}$. With respect to this basis, the transformation $\operatorname{ad}_\mathfrak{g}\xi\colon \mathfrak{g}\to\mathfrak{g}$ has a matrix $\|a_{ij}\|_{1\le i,j\le n}$ with $a_{ij} = 0$ for $i > m$ (since the image of this $\operatorname{ad}_\xi$ is contained in $\mathfrak{h}$). For the same $\xi$, the transformation $\operatorname{ad}_\mathfrak{h}\xi\colon \mathfrak{h}\to\mathfrak{h}$ has the matrix $\|a_{ij}\|_{1\le i,j\le m}$. For an $\eta \in \mathfrak{h}$, we use the same notations with $b$ instead of $a$. Then

$$B_\mathfrak{g}(\xi,\eta) = \operatorname{Tr}(\operatorname{ad}\mathfrak{g}\xi\circ\operatorname{ad}_\mathfrak{g}\eta) = \sum_{i=1}^n\sum_{j=1}^n a_{ij}b_{ji} = \sum_{i=1}^m\sum_{j=1}^m a_{ij}b_{ji} = \operatorname{Tr}(\operatorname{ad}_\mathfrak{h}\xi\circ\operatorname{ad}_\mathfrak{h}\eta) = B_\mathfrak{h}(\xi,\eta).$$

**2.4.2. Main results.** Theorem 1. *Let $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of a Lie algebra $\mathfrak{g}$, and let $\mathfrak{n}$ be an ideal in $\mathfrak{g}$ such that the representation $\rho|_\mathfrak{n}$ is nilpotent. Then $\mathfrak{n} \subset \operatorname{Ker} B^\rho$.*

Proof. Engel's Theorem provides a chain

$$0 = V_0 \subset V_1 \subset \ldots \subset V_{r-1} \subset V_r = V$$

such that $\rho(\eta)V_i \subset V_{i-1}$ for every $i$. We can assume that for every $i$, $V_{i-1} = \operatorname{span}\{\rho(\eta)V_i \mid \eta \in \mathfrak{n}\}$. Then an induction shows that every $V_i$ is invariant with respect to $\mathfrak{g}$. Indeed, it is true for $V_r$, and if it is true for $V_i$, then for every $\eta \in \mathfrak{n}, \xi \in \mathfrak{g}$, and $v \in V_i$, $\rho(\xi)(\rho(\eta)v) = \rho(\eta)(\rho(\xi)v) + \rho([\xi,\eta])v \in V_{i-1}$; thus, $\rho(\xi)V_{i-1} \subset V_{i-1}$. Consequently, for every $\xi \in \mathfrak{g}$, $\eta \in \mathfrak{n}$, $\rho(\xi)\circ\rho(\eta)V_i \subset V_{i-1}$, hence, the operator $\rho(\xi)\circ\rho(\eta)$ is nilpotent and $B_\mathfrak{g}(\xi,\eta) = \operatorname{Tr}(\rho(\xi)\circ\rho(\eta)) = 0$.

Corollary: $\operatorname{nil\,rad}\mathfrak{g} \subset \operatorname{Ker} B_{\mathfrak{g}}$*; in particular, if the Lie algebra* $\mathfrak{g}$ *is nilpotent, then* $B_{\mathfrak{g}} = 0$.

Theorem 2. *Let* $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ *be a representation. Then the algebra* $\rho(\mathfrak{g})$ *is solvable if and only if* $\mathfrak{g}' \subset \operatorname{Ker} B^{\rho}$.

This statement is called *Cartan's criterion of solvability*. Its proof is longer and will be presented in Section 2.4.3.

Corollary. *The Lie algebra* $\mathfrak{g}$ *is solvable if and only if* $\mathfrak{g}' \subset \operatorname{Ker} B_{\mathfrak{g}}$*. In particular, if* $B_{\mathfrak{g}} = 0$*, then* $\mathfrak{g}$ *is solvable.*

Theorem 3: $\operatorname{rad}\mathfrak{g} = (\mathfrak{g}')^{\perp}$.

*Proof.* Let $(\mathfrak{g}')^{\perp} = \mathfrak{h}$; this is an ideal in $\mathfrak{g}$ (Proposition 2 of 2.4.1), thus $B_{\mathfrak{h}}$ is a restriction of $B_{\mathfrak{g}}$ to $\mathfrak{h}$ (Proposition 4 of 2.4.1). Since every element of $\mathfrak{h}' \subset \mathfrak{g}'$ is $B_{\mathfrak{g}}$-orthogonal to $\mathfrak{h}$, it is also $B_{\mathfrak{h}}$-orthogonal to $\mathfrak{h}$. Thus, $(\mathfrak{h})' \subset \operatorname{Ker} B_{\mathfrak{h}}$, hence $\mathfrak{h}$ is solvable (Theorem 2), and hence $(\mathfrak{g}')^{\perp} = \mathfrak{h} \subset \operatorname{rad}\mathfrak{g}$. We will not need the opposite inclusion, and we leave it to a reader as an exercise:

Exercise 6. Prove that $\operatorname{rad}\mathfrak{g} \subset (\mathfrak{g}')^{\perp}$. Moreover, prove that for every representation $\rho$ of $\mathfrak{g}$, $B^{\rho}(\operatorname{rad}\mathfrak{g}, \mathfrak{g}') = 0$.

Theorem 4. *A Lie algebra* $\mathfrak{g}$ *is semisimple if and only if the Killing form* $B_{\mathfrak{g}}$ *is non-degenerate.*

*Proof.* If $\mathfrak{g}$ is semisimple, that is, $\operatorname{rad}\mathfrak{g} = 0$, then, by Theorem 3, $(\mathfrak{g}')^{\perp} = 0$. Therefore, $\operatorname{Ker} B_{\mathfrak{g}} = 0$ (it is contained in the $B_{\mathfrak{g}}$-orthogonal complement of anything), so $B_{\mathfrak{g}}$ is non-degenerate. If $\operatorname{Ker} B_{\mathfrak{g}} = 0$, then $\operatorname{nil\,rad}\mathfrak{g} = 0$ (Corollary to Theorem 1). Hence $(\operatorname{rad}\mathfrak{g})' = 0$ (see Proposition 2 in Section 2.3.2 or a remark in Section 2.3.4), hence $\operatorname{rad}\mathfrak{g}$ is commutative, hence it is nilpotent, hence it is contained in $\operatorname{nil\,rad}\mathfrak{g}$, hence it is 0.

**2.4.3. Proof of Cartan's criterion.** It is sufficient to prove Theorem 2 in the complex case. Indeed, real Lie algebra may be complexified to become a complex Lie algebra $\mathfrak{g}_{\mathbb{C}} = \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$. Representations also may be complexified: $\rho_{\mathbb{C}} = \rho \otimes \mathbb{C}\colon \mathfrak{g}_C \to \mathfrak{gl}(V \otimes \mathbb{C})$. It is true also that $(\mathfrak{g}_{\mathbb{C}})' = (\mathfrak{g}')_{\mathbb{C}}$, in particular, the Lie algebra $\mathfrak{g}$ is solvable if and only if $\mathfrak{g}_{\mathbb{C}}$ is solvable. Finally, the Killing form $B^{\rho_{\mathbb{C}}}$ is the complexification of $B^{\rho}$; the same is true for the kernels of these forms. Thus, the Cartan criterion for real Lie algebras follows from the Cartan criterion for the complex Lie algebas.

**2.4.3.1. A linear algebra preparation.** *A. Semisimple and nilpotent parts of a linear endomorphism.* Lemma 1. *Let* $f\colon V \to V$ *be a linear endomorphism of a complex vector space. Then there exists a unique pair of linear endomorphisms* $s, n\colon V \to V$ *such that:* (1) $f = s + n$; (2) $s$ *is semisimple, that is, diagonalizable,* $n$ *is nilpotent;* (3) $[s, n] = 0$.

*Proof.* Let $V = \bigoplus_i V_i$ be the decomposition of $V$ into the sum of invariant subspaces (all eigenvalues of the restriction $f_i\colon V_i \to V_i$ of $f$ are equal to some eigenvalue $\lambda_i$ of $f$). Then we define $s$ as $\cdot\lambda_i$ on $V_i$ and put $n = f - s$. All the properties of $s$ and $n$, as well as the uniqueness, are obvious.

The endomorphisms $s$ and $n$ are called semisimple and nilpotent parts of $f$.

*B: $s$ as a function of $f$.* Lemma 2. *For every* $f$*, there exists a polynomial* $p$ *of one variable without a constant term such that* $s = p(f)$.

*Proof.* Obviously, for $f, s$, and $n$ as in Lemma 1 and for any polynomial $p$,

$$p(f) = p(s) + p'(s)n + \frac{1}{2}p''(s)n^2 + \ldots + \frac{1}{k!}p^{(k)}(s)n^k$$

where $k$ is such that $n^{k+1} = 0$ (it is sufficient to check this for a monomial, $p(f) = f^m$). Thus, all we need is to find a polynomial $p$ without a constant term such that $p(s) = s$ and $p'(s) = p''(s) = \ldots = p^{(k)}(s) = 0$. For this, we can take a polynomial $p$ such that $p(0) = 0$, and $p(\lambda_i) = \lambda_i, p'(\lambda_i) = p''(\lambda_i) = \ldots = p^{(k)}(\lambda_i) = 0$ for all eigenvalues $\lambda_i$ of $s$ (that is, of $f$): if we apply a polynomial to a diagonal matrix $D$, then we get a diagonal matrix whose diagonal entries are obtained by application of the same polynomial to the diagonal entries of $D$.

*C. Semisimple and nilpotent parts of* $\operatorname{ad} f$. For an endomorphism $f : V \to V$ of a complex vector space, consider $\operatorname{ad} f : \operatorname{End} V \to \operatorname{End} V$, $\operatorname{ad} f(g) = [f, g]$.

LEMMA 3. *The semisimple and nilpotent parts of* $\operatorname{ad} f$ *are* $\operatorname{ad} s$ *and* $\operatorname{ad} n$ *where* $s$ *and* $n$ *are semisimple and nilpotent parts of* $f$.

*Proof.* (1) $\operatorname{ad} s$ *is semisimple*: choose a basis $e_1, \ldots, e_n$ in $V$ with $s(e_i) = \lambda_i e_i$. Let $E_{ij} \in \operatorname{End} V$ is defined by the formulas $E_{ij}(e_i) = e_j$, $E_{ij}(e_k) = 0$ for $k \neq i$. These $E_{ij}$ form a basis in $\operatorname{End} V$. Then $[s, E_{ij}] = (\lambda_i - \lambda_j)E_{ij}$, thus the matrix of $\operatorname{ad} s$ with respect to the basis $\{E_{ij}\}$ is diagonal, thus $\operatorname{ad} s$ is semisimple.

(2) $\operatorname{ad} n$ *is nilpotent*:

$$(\operatorname{ad} n)^m h = \sum_{\ell=0}^{m} (-1)^\ell \binom{m}{\ell} n^\ell h n^{m-\ell}$$

which is zero, if $n^{k+1} = 0$ and $m > 2k$.

(3) $\operatorname{ad} s$ *and* $\operatorname{ad} n$ *commute*:

$$\operatorname{ad} s \circ \operatorname{ad} n(h) = [s, [n, h]] = -[n, [h, s]] - [h, [s, n]] = [n, [h, s]] = \operatorname{ad} n \circ \operatorname{ad} s(h).$$

**2.4.3.2. Main Lemma.** *Let* $V$ *be a complex vector space, and let* $B \subset A \subset \operatorname{End} V$ *be subspaces. Let*

$$T = \{X \in \operatorname{End} V \mid [X, Y] \in B \text{ for every } Y \in A\}.$$

*If* $Z \in T$ *and* $\operatorname{Tr}(ZU) = 0$ *for all* $U \in T$, *then* $Z$ *is nilpotent.*

*Proof.* Let $Z = s + n$ be the decomposition of $Z$ as in Lemma 1; thus there is a basis $\{e_i\}$ of $V$ such that $s(e_i) = \lambda_i e_i$ ($\lambda_i \in \mathbb{C}$). Let $L \subset \mathbb{C}$ be the set of all rational linear combinations $\sum r_i \lambda_i$, $r_i \in \mathbb{Q}$; that is, $L$ is a rational vector subspace of $\mathbb{C}$. Let $f : L \to \mathbb{Q}$ be an arbitrary $\mathbb{Q}$-linear form. Our goal is to prove that $f = 0$ (this would imply $L = 0$, hence $s = 0$, hence $Z = n$).

Let $t \in \operatorname{End} V$, $t(e_i) = f(\lambda_i)e_i$. Then

$$(\operatorname{ad} s)E_{ij} = (\lambda_i - \lambda_j)E_{ij} \Rightarrow (\operatorname{ad} t)E_{ij} = (f(\lambda_i) - f(\lambda_j))E_{ij}.$$

Choose a polynomial $q$ without a constant term such that $q(\lambda_i - \lambda_j) = f(\lambda_i) - f(\lambda_j)$ (such a polynomial exists, because if $\lambda_i - \lambda_j = \lambda_k - \lambda_\ell$, then, since $f$ is $\mathbb{Q}$-linear, $f(\lambda_i) - f(\lambda_j) = f(\lambda_k) - f(\lambda_\ell)$). Then $\operatorname{ad} t = q(\operatorname{ad} s)$. On the other side, it follows from Lemmas 2 and 3 that $\operatorname{ad} s = p(\operatorname{ad} Z)$ for some polynomial $p$ without a constant term. Thus,

$$(\operatorname{ad} Z)(A) \subset B \Rightarrow (\operatorname{ad} s)(A) \subset B \Rightarrow (\operatorname{ad} t)(A) \subset B \Rightarrow t \in T.$$

Hence, by our assumption, $\operatorname{Tr}(Zt) = 0$. But the matrix of $Z$ (with respect to the basis $\{e_i\}$) is an upper triangular matrix with the diagonal entries $\lambda_i$, and the matrix of $t$ is the diagonal matrix with entries $f(\lambda_i)$. Hence $0 = \operatorname{Tr}(Zt) = \sum \lambda_i f(\lambda_i)$ and $0 = f(\operatorname{Tr}(Zt)) = f\left(\sum \lambda_i f(\lambda_i)\right) = \sum f(\lambda_i f(\lambda_i)) = \sum f(\lambda_i)^2$, so $f(\lambda_i) = 0$ for all $i$ and $f = 0$.

**2.4.3.3. The end of the proof.** We can assume that the representation $\rho$ is faithful, that is, $\rho\colon \mathfrak{g} \xrightarrow{\subset} \operatorname{End} V$. Indeed, $\rho(\mathfrak{g})' = \rho(\mathfrak{g}')$ and $\operatorname{Ker} \rho \subset \operatorname{Ker} B^\rho$. Hence, $\mathfrak{g}' \subset \operatorname{Ker} B^\rho$ if and only if $\rho(\mathfrak{g})' = \rho(\mathfrak{g}') \subset \rho(\operatorname{Ker} B^\rho) = B_{\rho(\mathfrak{g})}$.

*The only if part.* If $\mathfrak{g}$ is solvable, then it follows from the Lie theorem that $\rho|_{\mathfrak{g}'}$ is nilpotent. Hence $\mathfrak{g}' \subset \operatorname{Ker} B^\rho$ by Theorem 1.

*The if part.* Apply Main Lemma to $A = \mathfrak{g}, B = \mathfrak{g}'$. Thus,

$$T = \{X \in \operatorname{End} V \mid [X, \mathfrak{g}] \subset \mathfrak{g}'\}.$$

Let $U \in T$ and $\xi, \eta \in \mathfrak{g}$; let also $Z = [\xi, \eta]$. Then $[U, \xi] \in \mathfrak{g}'$ and hence

$$\operatorname{Tr}(ZU) = \operatorname{Tr}(UZ) = \operatorname{Tr}(U[\xi, \eta]) = \operatorname{Tr}([U, \xi]\eta) = B^\rho([U, \xi], \eta) = 0$$

and since $U \in T$ is arbitrary, Main Lemma implies that $Z$ is nilpotent. In the same way, we prove that an arbitrary sum of commutators in $\mathfrak{g}$ is nilpotent. Thus, $\mathfrak{g}'$ consists of nilpotent operators, that is, $\mathfrak{g}'$ is nilpotent, and hence $\mathfrak{g}$ is solvable (Proposition 2 of Section 2.3.2). This completes the proof of the Cartan criterion.

**2.4.4. Final comments.** Technically, the Cartan criterion is the main result of the theory of Lie algebras, and, consequently, of the theory of Lie groups. Geometrically better justified results, in particular, the whole theory of semisimple Lie algebras, can be derived from it without much efforts (a good example is provided by Theorem 4 of Section 2.4.2, but it will be more clear in Section 2.5 below). On the other hand, the proof of Cartan criterion is based by some specific linear algebra (mostly the Main Lemma of Section 2.4.3.2) which was mainly created by E. Cartan and is usually reffered to as the *theory of replicas*.

We did not mention replicas above, but we often showed that a linear operator can be replaced by another linear operator with very similar properties (like $f$ and $s$ in Section 2.4.2, $t$ and $s$ in Section 2.4.3.2, etc.). The formal definition of a replica is as follows. An endomorhism $f\colon V \to V$ induces, for all $p$ and $q$, endomorphisms of spaces of tensors:

$$f_{p,q} = \underbrace{f \otimes \ldots \otimes f}_{p} \otimes \underbrace{f^* \otimes \ldots \otimes f^*}_{q}\colon \underbrace{V \otimes \ldots \otimes V}_{p} \otimes \underbrace{V^* \otimes \ldots \otimes V^*}_{q} \to \underbrace{V \otimes \ldots \otimes V}_{p} \otimes \underbrace{V^* \otimes \ldots \otimes V^*}_{q}.$$

An endomorphism $g$ is called a replica of an endomorphism $f$, if $\operatorname{Ker} f_{p,q} \subset \operatorname{Ker} g_{p,q}$ for all $p, q$. An example (which was, actually used above): $g$ is a replica of $f$, if $g = p(f)$ where $p$ is a polynomial without a constant term.

Exercise 7. Prove that the relation "$g$ is a replica of $f$" is transitive, but not symmetric.

Once the Cartan criterion has been proved, we will not need replicas any seriously (although we can mention them for a couple of times).

**2.5. Semisimple Lie algebras.**

**2.5.1. Ideals of semisimple Lie algebras.** Let $\mathfrak{g}$ be a real or complex semisimple Lie algebra. Then the Killing form $B = B_{\mathfrak{g}}$ is non-degenerate; the symbol $\perp$ in this Section denotes the orthogonality with respect to this form (in particular, if $A$ is a subspace of $\mathfrak{g}$, then $\dim A + \dim A^{\perp} = \dim \mathfrak{g}$).

Proposition 1. *If $\mathfrak{h}$ is an ideal of $\mathfrak{g}$, then $[\mathfrak{h}, \mathfrak{h}^{\perp}] = 0$, $\mathfrak{h} \cap \mathfrak{h}^{\perp} = 0$, $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^{\perp}$ (as a Lie algebra).*

*Proof.* We already know that $\mathfrak{h}^{\perp}$ is an ideal of $\mathfrak{g}$ (Proposition 2 of Section 2.4.1). Let $\xi \in \mathfrak{g}, \eta \in \mathfrak{h}^{\perp}, \zeta \in \mathfrak{h}$. Then $B([\eta, \zeta], \xi) = B(\eta, [\zeta, \xi]) = 0$ (since $\eta \in \mathfrak{h}^{\perp}$ and $[\zeta, \xi] \in \mathfrak{h}$). Hence, $[\eta, \zeta] \in \operatorname{Ker} B$, which shows that $[\eta, \zeta] = 0$.

Since $[\mathfrak{h}, \mathfrak{h}^{\perp}] = 0$, the intersection $\mathfrak{h} \cap \mathfrak{h}^{\perp}$ is a commutative, hence solvable, ideal in $\mathfrak{g}$, hence $\mathfrak{h} \cap \mathfrak{h}^{\perp} = 0$.

Since $\dim \mathfrak{h} + \dim \mathfrak{h}^{\perp} = \dim \mathfrak{g}$ and $\mathfrak{h} \cap \mathfrak{h}^{\perp} = 0$, there is a vector space decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^{\perp}$. Also, since $[\mathfrak{h}, \mathfrak{h}^{\perp}] = 0$, for $\eta, \eta' \in \mathfrak{h}^{\perp}$ and $\zeta, \zeta' \in \mathfrak{h}$, $[\zeta + \eta, \zeta' + \eta'] = [\zeta, \zeta'] + [\zeta, \eta'] + [\eta, \zeta'] + [\eta, \eta'] = [\zeta, \zeta'] + [\eta, \eta']$, hence the decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^{\perp}$ is a Lie algebra decomposition.

Corollary 1. *An ideal of a semisimple Lie algebra is semisimple. An ideal of an ideal of a semisimple Lie algebra $\mathfrak{g}$ is an ideal of $\mathfrak{g}$.*

*Proof.* If $\mathfrak{h}$ is an ideal of $\mathfrak{g}$, then $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^{\perp}$. Hence, and ideal of $\mathfrak{h}$ is an ideal of $\mathfrak{g}$, in particular, no non-zero ideal of $\mathfrak{h}$ is solvable.

Corollary 2. *If $\mathfrak{g}$ is semisimple, then $\mathfrak{g}' = \mathfrak{g}$.*

*Proof:* $(\mathfrak{g}')^{\perp} \cong \mathfrak{g}/\mathfrak{g}'$ is a commutative ideal of $\mathfrak{g}$, hence $\mathfrak{g}/\mathfrak{g}' = 0$, hence $\mathfrak{g}' = \mathfrak{g}$.

Proposition 2. *A semisimple Lie algebra is a direct sum of non-commutative simple Lie algebras. moreover, this decomposition is unique (that is, if $\mathfrak{g} = \mathfrak{h}_1 \oplus \ldots \oplus \mathfrak{h}_n = \mathfrak{k}_1 \oplus \ldots \oplus \mathfrak{k}_m$ are two decompositions of a semisimple Lie algebra into the sum of simple Lie algebras, then $m = n$ and $(\mathfrak{k}_1, \ldots, \mathfrak{k}_m)$ is a permutation of $(\mathfrak{h}_1, \ldots, \mathfrak{h}_m)$).*

*Proof.* If a semisimple Lie algebra $\mathfrak{g}$ is not simple, then it has an ideal $\mathfrak{h}$, and $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^{\perp}$ is a decomposition of $\mathfrak{g}$ into the sum of two semisimple Lie algebras. Then we apply the same arguments to $\mathfrak{h}$ and $\mathfrak{h}^{\perp}$, until $\mathfrak{g}$ is decomposed into the sum of simple Lie algebras. To prove the uniqueness of such decomposition, we use the following obvious statement: if $\mathfrak{h}$ and $\mathfrak{k}$ are ideals of a Lie algebra $\mathfrak{g}$ (semisimple or not) and $\mathfrak{h} \cap \mathfrak{k} = 0$, then $[\mathfrak{h}, \mathfrak{k}] = 0$ (indeed, $[\mathfrak{h}, \mathfrak{k}]$ is contained in both $\mathfrak{h}$ and $\mathfrak{k}$, hence is contained in $\mathfrak{h} \cap \mathfrak{k} = 0$). If there are two decomposition as above, then $\mathfrak{k}_1$ cannot have zero intersection with all $\mathfrak{h}_i$: otherwise, $[\mathfrak{k}_1, \mathfrak{h}_i] = 0$ for all $i$, hence $[\mathfrak{k}_1, \mathfrak{g}] = 0$, that is, $\mathfrak{k}_1$ is contained in the center of $\mathfrak{g}$ which is zero, since $\mathfrak{g}$ is semisimple. Hence, there is a non-zero intersection $\mathfrak{k}_1 \cap \mathfrak{h}_i$. But this intersection is an ideal in both $\mathfrak{k}_1$ and $\mathfrak{h}_i$, and hence it coincides with both, since $\mathfrak{k}_1$ and $\mathfrak{h}_i$ are both simple. In the same way, we see that each of $\mathfrak{k}_j$ coincides with one of $\mathfrak{h}_i$ and each of $\mathfrak{h}_i$ coincides with some of $\mathfrak{k}_j$.

Exercises. 8. Prove that every ideal of a semisimple Lie algebra is a subsum of a decomposition of this Lie algebra into the sum of simple Lie algebras.

9. Deduce from these that if two ideals of a semisimple Lie algebra have zero intersection, then they are orthogonal to each other.

**2.5.2. Uniqueness of the invariant form.** Proposition 3. *Let $\mathfrak{g}$ be a simple Lie algebra, and let $A, B$ be two invariant symmetric bilinear forms on $\mathfrak{g}$, and $B \neq 0$. Then $A = \lambda B$ for some constant $\lambda$.*

*Proof.* The forms $A$ and $B$ are described by symmetric matrices with respect to an arbitrary basis in $\mathfrak{g}$, and we will denote these matrices also by $A$ and $B$. Consider $\det(A - \lambda B)$; this is a polynomial of degree $\dim \mathfrak{g}$ in $\lambda$, and it is well known that in the real case (that is, when $\mathfrak{g}$ is a real Lie algebra), all its roots are real. Let $\lambda_0$ be one of these roots. Then $C = A - \lambda_0 B$ is a degenerate invariant form, $\operatorname{Ker} C \neq 0$, and it follows from the invariance that it is an ideal of $\mathfrak{g}$ (if $\xi, \eta \in \mathfrak{g}$, $\zeta \in \operatorname{Ker} C$, then $C([\zeta, \xi], \eta) = C(\zeta, [\xi, \eta]) = 0$, so $[\zeta, \xi] \in \operatorname{Ker} C$). But $\mathfrak{g}$ is simple, so $\operatorname{Ker} C = \mathfrak{g}$, hence $A = \lambda_0 B$.

In other words, every invariant symmetric bilinear form on $\mathfrak{g}$ is proportional to the Killing form. In particular, it is true for the forms $B^\rho$. Notice that if $\rho \neq 0$, then $B^\rho \neq 0$ (indeed, since $\operatorname{Ker} \rho$ is an ideal of $\mathfrak{g}$, if it is not $\mathfrak{g}$, is has to be 0; by the Cartan criterion, if $B^\rho = 0$, then $\rho(\mathfrak{g}) \cong \mathfrak{g}$ is solvable, a contradiction).

If the Lie algebra $\mathfrak{g}$ is semisimple, so, by Proposition 2, it is a direct sum of simple Lie algebras $\mathfrak{h}_i$, then an invariant symmetric bilinear form on $\mathfrak{g}$ is proportional to the Killing form on every $\mathfrak{h}_i$, but the proportianlity coefficients may be different for different $\mathfrak{h}_i$. In other words, the dimension of the space of invariant forms on $\mathfrak{g}$ equals the number of summands $\mathfrak{h}_i$.

For a representation $\rho$ of $\mathfrak{g}$, $\operatorname{Ker} \rho$ is the sum of those $\mathfrak{h}_i$, for which $B^\rho|_{\mathfrak{h}_i} = 0$. In other words, the form $B^\rho$ is non-degenerate on $(\operatorname{Ker} \rho)^\perp$.

**2.5.3. Casimir operators.**

**2.5.3.1. Definition and main properties.** Let $\mathfrak{g}$ be a (real or complex) semisimple Lie algebra. Let $\{g_1, \dots, g_m\}$ be a basis in $\mathfrak{g}$ and $\{g^1, \dots, g^m\}$ be the dual basis with respect to the Killing form (meaning that $B(g_i, g^j) = \delta_i^j$). This means, in particular, that for any $g \in G$, $g = \sum_i B(g^i, g) g_i = \sum_i B(g_i, g) g^i$. The element $C$ of the tensor product $\mathfrak{g} \otimes \mathfrak{g}$ defined by the formula $C = \sum_i g_i \otimes g^i$ is called the *Casimir element*.

Let us show that $C$ does not depend on the choice of the basis $\{g_i\}$. If $\{g'_j\}$ is a different basis and $g_i = \sum_j u_{ji} g'_j$, then for the dual basis $\{'g^k\}$ we have $g^k = \sum_\ell v_{k\ell} {'g}^\ell$ where $\sum_s u_{is} v_{sj} = \delta_{ij}$. Hence

$$\sum_i g_i \otimes g^i = \sum_{i,j,k} u_{ji} v_{ik} g'_j \otimes {'g}^k = \sum_{j,k} \delta_{jk} g'_j \otimes {'g}^k = \sum_j g'_j \otimes {'g}^j$$

For a representation $\rho \colon \mathfrak{g} \to \mathfrak{gl}(V)$, there arises a *Casimir operator*

$$\rho(C) = \sum_i \rho(g_i) \circ \rho(g^i) \colon V \to V$$

and the previous arguments show that $\rho(C)$ also does not depend on the choice of the basis.

THEOREM. *The Casimir operator commutes with the operators of the representation:*

$$\rho(C) \circ \rho(g) = \rho(g) \circ \rho(C)$$

*for every $g \in \mathfrak{g}$.*

*Proof.* Obviously,

$$\sum_i \rho(g_i) \circ \rho(g^i) \circ \rho(g) = \sum_i \rho(g_i) \circ \rho([g^i, g]) + \sum_i \rho([g_i, g]) \circ \rho(g^i) + \sum_i \rho(g) \circ \rho(g_i) \circ \rho(g^i),$$

so we need to check that $\sum_i \rho(g_i) \circ \rho([g^i, g]) + \sum_i \rho([g_i, g]) \circ \rho(g^i) = 0$ which is the same as $\sum_i g_i \otimes [g^i, g] + \sum_i [g_i, g] \otimes g^i = 0$. We have:

$$\sum_i g_i \otimes [g^i, g] = \sum_{i,k} g_i \otimes B(g_k, [g^i, g]) g^k = \sum_{i,k} B(g_k, [g^i, g]) g_i \otimes g^k,$$
$$\sum_k [g_k, g] \otimes g^k = \sum_{k,i} B(g^i, [g_k, g]) g_i \otimes g^k,$$

and we need only to check the equality $B(g_k, [g^i, g]) + B(g^i, [g_k, g]) = 0$ which follows from the invariance of the Killing form.

COROLLARY 1. *If the representation $\rho$ is irreducible, then $\rho(C)$ is a multiplication by some number.*

*Proof.* In the complex case, we find a eigenvector $v \in V$ of some $\rho(C)$ with some eigenvalue. But then every $\rho(g_{i_1}) \circ \ldots \circ \rho(g_{i_k}) v$ is an eigenvector of $\rho(C)$ with the same eigenvalue. Such vectors (with all $k, i_1, \ldots, i_k$) span a subrepresentation of $\rho$ which is the whole space $V$, if $\rho$ is irreducible. The transition to the real cases is done by means of the complexification of a real representation.

There is a small generalization of the previous construction. If $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is a representation of a semisimple Lie algebra, then, as was mentioned above, the form $B^\rho$ is non-degenerate on $\operatorname{Ker} \rho^\perp$. For a basis $g_1, \ldots, g_m$ of $\operatorname{Ker} \rho^\perp$ we can form the dual basis $g^1, \ldots, g^m$ with $B^\rho(g_i, g^j) = \delta_i^j$ and then define a $\rho$-Casimir operator $C^\rho\colon V \to V$ as $\sum_i \rho(g_i) \circ \rho(g^i)$. The independence of $C^\rho$ of the choice of the basis and its commuting with all the operators $\rho(g)$ are proved as before. One more property:

PROPOSITION: $\operatorname{Tr} C^\rho = \dim(\mathfrak{g}/\operatorname{Ker} \rho)$.

*Proof.* For every $i$, $\operatorname{Tr}(\rho(g_i) \circ \rho(g^i) = B^\rho(g_i, g^i) = 1$.

COROLLARY 2. *If $\rho \neq 0$, then the operator $C^\rho$ is not nilpotent.*

**2.5.3.2. An application to invariants of representations.** Let again $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of a semisimple Lie algebra. A vector $v \in V$ is called an *invariant* of $\rho$, if $\rho(g)v = 0$ for every $g \in \mathfrak{g}$. The space of all invariants of $\rho$ is denoted as $V^\rho$. The notation $V_\rho$ is used for the space $\sum_{g \in \mathfrak{g}} gV$ (elements of the space $V/V_\rho$ are called sometimes *coinvariants* of $\rho$).

THEOREM. *The spaces $V^\rho$ and $V_\rho$ are subrepresentations of $\rho$, and $V = V^\rho \oplus V_\rho$.*

Lemma (from Linear Algebra). *Let $f: V \to V$ be an endomorphism of a (real or complex) vector space $V$. Let*

$$W_1 = \bigcup_{k\geq 1} \operatorname{Ker} f^k,\ \ W_2 = \bigcap_{k\geq 1} \operatorname{Im} f^k.$$

*Then $V$ is a direct sum of $W_1$ and $W_2$.*

*Proof.* We begin with the complex case. For a For a $\lambda \in \mathbb{C}$, put

$$V(f,\lambda) = V^\lambda = \{v \in V \mid (f - \lambda \cdot \mathrm{id})^n v = 0 \text{ for some } n\}$$

(one can take $n = \dim V$). Obviously, $(f - \lambda \cdot \mathrm{id})V^\lambda \subset V^\lambda$.

It is known that $V^\lambda \neq 0$ if and only if $\lambda$ is an eigenvalue of $f$, and $V = \bigoplus_\lambda V^\lambda$. It is also obvious that $W_1 = V^0$ and $W_2 = \bigoplus_{\lambda\neq 0} V^\lambda$, which proves Lemma.

In the real case we apply the previous result to $f \otimes \mathbb{C}: V \otimes \mathbb{C} \to V \otimes \mathbb{C}$ and observe that the spaces $W_1$ and $W_2$ related to $f \otimes \mathbb{C}$ are $W_1 \otimes \mathbb{C}$ and $W_2 \otimes \mathbb{C}$.

*Proof of Theorem.* The invariance of $V^\rho$ and $V_\rho$ with respect to the action of $\mathfrak{g}$ is obvious, let us prove the direct sum decomposition. We use the induction with respect to $\dim V$. If $\dim V = 0$, then we have nothing to prove; let $\dim V > 0$. Consider the decomposition of $V$ from Lemma with respect to the operator $C^\rho$:

$$W_1 = \bigcup_{k\geq 1} \operatorname{Ker}(C^\rho)^k,\ \ W_2 = \bigcap_{k\geq 1} \operatorname{Im}(C^\rho)^k.$$

Since $C^\rho$ commutes with the representation operators, both spaces are subrepresentations of $\rho$, and, by Lemma, $V = W_1 \oplus W_2$. If both $W_1$ and $W_2$ are non-zero, the statement holds for them by the induction hypothesis, hence it holds for $V$. If $\rho = 0$, then $V^\rho = V, V_\rho = 0$, and the statements holds. If $\rho \neq 0$, then $C^\rho$ is not nilpotent (see Corollary 2 in Section 2.5.3.1), so $W_1 \neq V$, $W_2 \neq 0$, and it remains to consider the case when $W_1 = 0$. In this case, $C^\rho$ must be invertible. Then for every $v \in V$,

$$v = C^\rho (C^\rho)^{-1} v = \sum_i \rho(g_i) \circ \rho(g^i) \circ (C^\rho)^{-1} v \in V_\rho,$$

so $V_\rho = V$, and if $v \in V^\rho$, then $C^\rho v = \sum_i \rho(g_i) \circ \rho(g^i) v = 0$, so $v \in W_1 = 0$, so $V^\rho = 0$. Thus $V = V^\rho \oplus V_\rho$, which completes the proof.

# 3. Cartan algebras, weights, and roots.

## 3.1. General theory.

**3.1.1. Weight space decomposition.** Let $\rho: \mathfrak{h} \to \mathfrak{gl}(V)$ be a representation of a nilpotent complex Lie algebra. (Several first definitions do not require either nilpotency of $\mathfrak{h}$, or the ground field being $\mathbb{C}$.) A linear form $\lambda: \mathfrak{h} \to \mathbb{C}$ is called a *weight* of $\rho$, if there exists a non-zero vector $v \in V$ such that $\rho(h)v = \lambda(h) \cdot v$ for every $h \in \mathfrak{h}$. (This definition appeared in Section 2.2.2 above; the Lie theorem asserts that a weight exists for every representation of a complex solvable Lie algebra.)

For a representation $\rho$ and a $\lambda \in V^*$, put

$$V(\mathfrak{h}, \lambda) = V^\lambda = \bigcap_{h \in \mathfrak{h}} V(\rho(h), \lambda(h))$$

(we use the notations from Section 2.5.3.2).

Proposition. *Let $\mathfrak{h}$ be a nilpotent algebra over $\mathbb{C}$. Then*

(1) $\rho(\mathfrak{h}) V^\lambda \subset V^\lambda$;

(2) *if $V^\lambda \neq 0$, then $\lambda$ is a weight;*

(3) $V = \bigoplus_\lambda V^\lambda$.

*Proof.* (1) We will prove a little bit more (and it will be needed for Part (3)): for every $g, h \in \mathfrak{h}$,

$$\rho(g) V(\rho(h), \lambda(h)) \subset V(\rho(h), \lambda(h)).$$

Since $\mathfrak{h}$ is nilpotent, $(\operatorname{ad} h)^k g = 0$ for some $k$. We use the induction with respect to $k$. If $k = 0$, then $g = 0$. Assume that $k > 0$ and the statement is true for the less values of $k$. It is obvious that

$$(\rho(h) - \lambda(h))^n \circ \rho(g) = \rho(g) \circ (\rho(h) - \lambda(h))^n + \sum_{s=0}^{n-1} (\rho(h) - \lambda(h))^{n-s-1} \rho([h, g]) (\rho(h) - \lambda(h))^s.$$

Take $n > 2 \dim V(\rho(h), \lambda(h))$ and apply the both sides of the last equality to a vector $v \in V(\rho(h), \lambda(h))$. The first summand gives 0, the second summand gives 0 for $s > \dim V(\rho(h), \lambda(h))$. If $0 \leq s \leq \dim V(\rho(h), \lambda(h))$, then $(\rho(h) - \lambda(h))^s v \in V(\rho(h), \lambda(h))$, and $V(\rho(h), \lambda(h))$ is invariant with respect to $\rho([h, g])$ by the induction hypothesis (since $(\operatorname{ad} h)^{k-1}[h, g] = (\operatorname{ad} h)^k g = 0$) and $V(\rho(h), \lambda(h))$ is annihilated by $(\rho(h) - \lambda(h))^{n-s-1}$. Hence $\rho(g) v \in \operatorname{Ker}(\rho(h) - \lambda(h))^n = V(\rho(h), \lambda(h))$.

(2): Lie's theorem.

(3) First of all, the sum of $V^\lambda$ is direct (take $h \in \mathfrak{h}$ such that all $\lambda_i(h)$ are different, and we see that if vectors $v_i \in V^{\lambda_i}$ are linearly dependent, then they all are zeroes). Then we apply the induction with respect to $\dim V$. If for every $h$ the operator $\rho(h)$ has only one eigenvalue, $\lambda(h)$, then $V = V^\lambda$; if for at least one $h$, there are more than one eigenvalue, then the representation is reducible, and we can apply the induction hypothesis to a subrepresentation and the quotient.

**3.1.2. Roots.** Let $\mathfrak{g}$ be a Lie algebra, and let $\mathfrak{h}$ be a nilpotent Lie subalgebra of $\mathfrak{h}$. The weights of the representation $\operatorname{ad}_{\mathfrak{g}}|_{\mathfrak{h}} \colon \mathfrak{h} \to \mathfrak{gl}(\mathfrak{g})$ are called *roots* of the Lie algebra $\mathfrak{g}$ with respect to $\mathfrak{h}$. Thus, $\mathfrak{g} = \bigoplus_{\alpha \in \text{roots}} \mathfrak{g}^\alpha$; the spaces $\mathfrak{g}^\alpha$ are called *root spaces.*

Example. Let $\mathfrak{g} = \mathfrak{gl}(n, \mathbb{C})$, $\mathfrak{h} = \{\text{diagonal matrices}\}$. Then

$$\left[ \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, E_{ij} \right] = (\lambda_i - \lambda_j) E_{ij}.$$

Thus, roots are $\lambda_i - \lambda_j$ $(i \neq j)$ and 0. Obviously, $\mathfrak{g}^0 = \mathfrak{h}$ and $\dim \mathfrak{g}^\alpha = 1$ for all non-zero roots.

It is also obvious that $\mathfrak{h} \subset \mathfrak{g}^0$; moreover, $\mathfrak{g}^0$ contains any nilpotent subalgebra $\mathfrak{k}$ of $\mathfrak{g}$ such that $\mathfrak{k} \supset \mathfrak{h}$ (indeed, $\mathfrak{g}^0$ is the set of $g \in \mathfrak{g}$ such that $(\operatorname{ad} h)^k g = 0$ for all $h \in \mathfrak{h}$ and $k \geq \dim \mathfrak{g}$).

Proposition 1: $[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] = \mathfrak{g}^{\alpha+\beta}$.

(If $\gamma$ is not a root, then $\mathfrak{g}^\gamma$ is 0.)

*Proof of Proposition 1.* Let $g_1 \in \mathfrak{g}^\alpha$, $g_2 \in \mathfrak{g}^\beta$. Then, for every $h \in \mathfrak{h}$,

$$\begin{aligned}(\operatorname{ad} h - (\alpha+\beta)(h))[g_1, g_2] &= [[h, g_1], g_2] + [g_1, [h, g_2]] - \alpha(h)[g_1, g_2] - \beta(h)[g_1, g_2] \\ &= [[h, g_1], g_2] + [g_1, [h, g_2]] - [\alpha(h) g_1, g_2] - [g_1, \beta(h) g_2] \\ &= [(\operatorname{ad} h - \alpha(h)) g_1, g_2] + [g_1, (\operatorname{ad} h - \beta(h)) g_2].\end{aligned}$$

An $n$-fold application of this formula gives:

$$(\operatorname{ad} h - (\alpha+\beta)(h))^n [g_1, g_2] = \sum_{p=0}^{n} \binom{n}{p} \left[(\operatorname{ad} h - \alpha(h))^p g_1, (\operatorname{ad} h - \beta(h))^{n-p} g_2\right]$$

which is zero for a sufficiently large $n$.

Corollary: *$\mathfrak{g}^0$ is a Lie subalgebra of $\mathfrak{g}$.*

In conclusion, let us mention two obvious properties of roots.

Proposition 2. *If $\alpha\colon \mathfrak{h} \to \mathbb{C}$ is a root, then $\alpha|_{\mathfrak{h}'} = 0$.*

*Proof.* For $h_1, h_2 \in \mathfrak{h}$ and $g \in \mathfrak{g}^\alpha$, $\alpha([h_1, h_2])g = \operatorname{ad}[h_1, h_2]g = \operatorname{ad} h_1(\operatorname{ad}_2 g) - \operatorname{ad} h_2(\operatorname{ad} h_1 g) = (\alpha(h_1)\alpha(h_2) - \alpha(h_2)\alpha(h_1))g = 0$.

For a root $\alpha$, let $d(\alpha) = \dim \mathfrak{g}^\alpha$.

Proposition 3. *For $h, h' \in \mathfrak{h}$,*

$$\langle h, h' \rangle = \sum_{\alpha \in \{\text{roots}\}} d(\alpha)\alpha(h)\alpha(h').$$

*Proof.* Since the both sides of this equality are symmetric bilinear forms, it is sufficient to prove it for $h' = h$. Let us do this: $\langle h, h \rangle = \operatorname{Tr}(\operatorname{ad} h)^2 = \sum_\alpha \operatorname{Tr}\left((\operatorname{ad} h)^2|_{\mathfrak{g}^\alpha}\right) = \sum_\alpha d(\alpha)\alpha(h)^2$.

**3.1.3. Cartan algebras.** Definition 1. If $\mathfrak{g}^0 = \mathfrak{h}$, then $\mathfrak{h}$ is called a *Cartan algebra* of $\mathfrak{g}$.

In particular, a Cartan algebra is a maximal nilpotent subalgebra of $\mathfrak{g}$.

Definition 2. An element $g$ of $\mathfrak{g}$ is called *regular*, if $\dim \mathfrak{g}(\operatorname{ad} g, 0)$ is minimal (that is, does not exceed $\dim \mathfrak{g}(\operatorname{ad} g', 0)$ for any $g' \in \mathfrak{g}$).

The following description of regular elements of $\mathfrak{g}$ shows that the set of regular elements is a dense open subset (actually, a Zariski open subset) of $\mathfrak{g}$. Let $p(\lambda) =$

$\lambda^m + p_{m-1}(g)\lambda^{m-1} + \ldots + p_1(g)\lambda + p_0(g)$ (where $m = \dim g$) be the characteristic polynomial of $\operatorname{ad} g$. Obviously, $p_0(g) \equiv 0$ (since $g \in \operatorname{Ker} \operatorname{ad} g$). Let $\ell$ be the minimal number, for which $p_\ell(g) \not\equiv 0$. Then $g$ is regular, if and only if $p_\ell(g) \neq 0$. (Indeed, for a linear operator $f\colon V \to V$ the dimension of $V(f, 0)$ is the multiplicity of the root 0 of the characteristic polynomial of $f$.)

Exercise 1. Prove that $A \in \mathfrak{gl}(n, \mathbb{C})$ is regular if and only if all eigenvalues of $A$ are pairwise different.

Theorem. *Let $g$ be a regular element of $\mathfrak{g}$. Then $\mathfrak{g}(\operatorname{ad} g, 0)$ is a Cartan algebra of $\mathfrak{g}$.*

*Proof.* Let $\mathfrak{h} = \operatorname{span}(g)$. With respect to $\mathfrak{h}$,

$$\mathfrak{g} = \mathfrak{g}^0 \oplus \underbrace{\bigoplus\nolimits_{\alpha \neq 0} \mathfrak{g}^\alpha}_{\widetilde{\mathfrak{g}}}; \ \mathfrak{g}^0 = \mathfrak{g}(\operatorname{ad} g, 0).$$

Since $[\mathfrak{g}^0, \mathfrak{g}^\alpha] \subset \mathfrak{g}^\alpha$, then $[\mathfrak{g}, \widetilde{\mathfrak{g}}] \subset \widetilde{\mathfrak{g}}$.

For a $\xi \in \mathfrak{g}^0$, put $d(\xi) = \det(\operatorname{ad} \xi\colon \widetilde{\mathfrak{g}} \to \widetilde{\mathfrak{g}})$. This is a polynomial function of $\mathfrak{g}^0$, and $d(g) \neq 0$ (indeed, the eigenvalues of $\operatorname{ad} g$ on $\widetilde{\mathfrak{g}}$ are $\alpha(g)$, all different from 0), so $\{\xi \in \mathfrak{g}^0, d(\xi) \neq 0\}$ is dense in $\mathfrak{g}^0$. Let $\eta \in \mathfrak{g}^0$, $d(\eta) \neq 0$. Then $\operatorname{ad} \eta\colon \widetilde{\mathfrak{g}} \to \widetilde{\mathfrak{g}}$ is non-degenerate, so $\mathfrak{g}(\operatorname{ad} \eta, 0) \subset \mathfrak{g}^0 = \mathfrak{g}(\operatorname{ad} g, 0)$. But $\dim \mathfrak{g}(\operatorname{ad} g, 0) \leq \dim \mathfrak{g}(\operatorname{ad} \eta, 0)$, because $g$ is regular (see the definition of regularity). Hence $\mathfrak{g}(\operatorname{ad} \eta, 0) = \mathfrak{g}^0$. This shows that $\operatorname{ad} \xi\colon \mathfrak{g}^0 \to \mathfrak{g}^0$ is nilpotent whenever $d(\xi) \neq 0$, but then it is nilpotent for every $\xi \in \mathfrak{g}^0$, since the set $\{d(\xi) \neq 0\}$ is dense in $\mathfrak{g}^0$.

Thus, $\mathfrak{g}^0$ is nilpotent and $\mathfrak{g}^0 \subset \mathfrak{g}(\mathfrak{g}^0, 0)$. But also $\mathfrak{g}^0 \supset \mathfrak{h}$, hence $\mathfrak{g}(\mathfrak{g}^0, 0) \subset \mathfrak{g}(\mathfrak{h}, 0) = \mathfrak{g}^0$. So, $\mathfrak{g}^0 = \mathfrak{g}(\mathfrak{g}^0, 0)$, and $\mathfrak{g}^0$ is a Cartan algebra.

Exercise 2. Prove that if a Cartan algebra $\mathfrak{h} \subset \mathfrak{g}$ contains a regular element $g \in \mathfrak{g}$, then it is $\mathfrak{g}(\operatorname{ad} g, 0)$. In particular, if two Cartan algebras share a regular element, then they coincide.

Actually, it is true that every Cartan algebra contains a regular element. and hence every Cartan algebra has the form $\mathfrak{g}(\operatorname{ad} g, 0)$ for a regular $g$. In addition to this, if $\mathfrak{g} = \operatorname{Lie} G$, then all Cartan algebras are obtained from each other by transformations $\operatorname{Ad} g, g \in G$. We will not need this, and we will not prove this (at least, in this generality).

**3.1.4. Root series.** From now on, speaking of roots of a Lie algebra $\mathfrak{g}$ we always mean roots with respect to some fixed Cartan algebra $\mathfrak{h} \subset \mathfrak{g}$.

Let $\alpha, \beta$ be roots and $\alpha \neq 0$. Then there a a maximal $p \neq 0$ and a minimal $q \neq 0$ such that $[\mathfrak{g}^{-\alpha}, \mathfrak{g}^{\beta + p\alpha}] \neq 0, [\mathfrak{g}^\alpha, \mathfrak{g}^{\beta + q\alpha}] \neq 0$. The set of roots $\{\beta + k\alpha \mid p \leq k \leq q\}$ is called a *root series*.

Proposition. *For an $h \in [\mathfrak{g}^\alpha, \mathfrak{g}^{-\alpha}] \subset \mathfrak{g}^0 = \mathfrak{h}$,*

$$\beta(h) = -r\alpha(h) \text{ where } r = \frac{\sum_{k=p}^q k d(\beta + k\alpha)}{\sum_{k=p}^q d(\beta + k\alpha)}.$$

*Proof.* The space $V = \bigoplus_{k=p}^{q} \mathfrak{g}^{\beta+k\alpha}$ is both $\operatorname{ad} \mathfrak{g}^{\alpha}$-invariant and $\operatorname{ad} \mathfrak{g}^{-\alpha}$-invariant. Let $g_1 \in \mathfrak{g}^{\alpha}$ and $g_2 \in \mathfrak{g}^{-\alpha}$, and let $h = [g_1, g_2]$. Then

$$0 = \operatorname{Tr}(\operatorname{ad} h|_V) = \sum_{k=p}^{q} d(\beta + k\alpha)(\beta(h) + k\alpha(h))$$

which proves our proposition.

**3.2. Lie algebra cohomology.** Below, $\mathbb{K}$ denotes $\mathbb{R}$ or $\mathbb{C}$ (actually, any field of characteristic zero will do).

**3.2.1. Basic definitions.** Let $\rho: \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of a Lie algebra $\mathfrak{g}$ (equivalently, $V$ is a $\mathfrak{g}$-module). We put

$$C^q(\mathfrak{g}; V) \text{ or } C^q(\mathfrak{g}; \rho) = \operatorname{Hom}_{\mathbb{K}}(\Lambda^q \mathfrak{g}, V),$$

the space of skew-symmetric $q$-linear forms on $\mathfrak{g}$ with values in $V$; these forms are called *cochains* of $\mathfrak{g}$ with values in $V$ of degree (or dimension) $q$. In particular,

$$\begin{aligned} &C^0(\mathfrak{g}; V) = V, \\ &C^1(\mathfrak{g}; V) = \operatorname{Hom}_{\mathbb{K}}(\mathfrak{g}, V). \end{aligned}$$

The *differential*

$$\delta = \delta^q: C^q(\mathfrak{g}; V) \to C^{q+1}(\mathfrak{g}; V)$$

is defined by the formula

$$\begin{aligned} \delta^q c(g_1, \dots, g_{q+1}) = &\sum_{1 \le s < t \le q+1} (-1)^{s+t-1} c([g_s, g_t], g_1, \dots \widehat{g}_s \dots \widehat{g}_t \dots, g_{q+1}) \\ &+ \sum_{1 \le u \le q+1} (-1)^u g_u c(g_1, \dots \widehat{g}_u \dots, g_{q+1}). \end{aligned}$$

EXAMPLES. (1) If $c = v \in C^0(\mathfrak{g}; V) = V$, then $\delta^0 c(g) = -gv$.

(2) $\delta^1 c(g_1, g_2) = c([g_1, g_2]) - g_1 c(g_2) + g_2 c(g_1)$.

(3) $\delta^2 c(g_1, g_2, g_3) = c([g_1, g_2], g_3) + c([g_2, g_3], g_1) + c([g_3, g_1], g_2)$
$\qquad - g_1 c(g_2, g_3) - g_2 c(g_3, g_1) - g_3 c(g_1, g_2)$.

PROPOSITION. *For all $q$, $\delta^{q+1} \circ \delta^q = 0$.*

*Proof*: direct verification.

EXERCISE 3. Do this verification (do it at least for $q = 0, 1$).

DEFINITION of cohomology[10]) (of degree or dimension $q$):

$$H^q(\mathfrak{g}; V) = \operatorname{Ker} \delta^q / \operatorname{Im} \delta^{q-1}.$$

---

[10]) The terminology used for the Lie algebra cohomology has, mostly, a topological origin. The words *cochain, cohomology, dimension* are the most obvious examples; but also cochains from $\operatorname{Ker} \delta$ are called *cocycles*, the cochains from $\operatorname{Im} \delta$ are called *coboundaries*, cocycles are called *cohomologous*, if their difference is a coboundary, and the differential $\delta$ itself is often called the *coboundary operator*.

We will not use the cohomology of dimensions $> 2$ any seriously.

**3.2.2. Cohomologies of dimensions 0,1,2.** The cohomology of dimensions $\leq 2$ (and a little bit of dimension 3) have direct algebraic interpretations. We list some of them below (mostly as exercises). We must warn the reader that almost nothing of the material of this section will be used later. It may be of some interest to a reader who wants to get some additional understanding of the beautiful theory of cohomology of Lie algebras.

EXERCISES. 4. Construct a natural[11]) isomorphism

$$H^0(\mathfrak{g}; V) \cong V^{\mathfrak{g}} = \{v \in V \mid gv = 0 \text{ for all } g \in \mathfrak{g}\}$$

(elements of $V^{\mathfrak{g}}$ are called *invariants* of the representation).

5. Construct a natural isomorphism

$$H^1(\mathfrak{g}; \mathbb{K}) \cong (\mathfrak{g}/\mathfrak{g})^*.$$

(In particular, $H^1(\mathfrak{g}; \mathbb{K}) = 0$ if and only if $\mathfrak{g}' = \mathfrak{g}$; for example, $H^1(\mathfrak{g}; \mathbb{K}) = 0$, if the Lie algebra $\mathfrak{g}$ is semisimple: see Corollary 2 in Section 2.5.1).

Remind (see Comment (2) in Section 1.4.2) that a linear endomorphism $D: \mathfrak{g} \to \mathfrak{g}$ of a Lie algebra is called a *derivation*, if $D[\xi, \eta] = [D\xi, \eta] + [\xi, D\eta]$ for all $\xi, \eta \in \mathfrak{g}$. The space $\operatorname{Der} \mathfrak{g}$ of all derivations of $\mathfrak{g}$ possesses a natural structure of a Lie algebra. There are *inner derivations* $\operatorname{ad} \xi$. Elements of the cokernel of ad, $\operatorname{Der} \mathfrak{g} / \operatorname{ad}(\mathfrak{g})$ are called *outer derivations*, and the space of outer derivations of $\mathfrak{g}$ is denoted as $\operatorname{Out} \mathfrak{g}$.

EXERCISES. 6. Construct a natural isomorphism

$$H^1(\mathfrak{g}; \operatorname{ad}) \cong \operatorname{Out} \mathfrak{g}.$$

(*Hint*: cochains in $C^1(\mathfrak{g}; \operatorname{ad})$ are the same as linear endomorphisms of $\mathfrak{g}$; cocycles are the same as derivations of $\mathfrak{g}$; coboundaries are the same as inner derivations of $\mathfrak{g}$.)

7. For an arbitrary $\mathfrak{g}$-module $V$ construct a natural bijection between $H^1(\mathfrak{g}; V)$ and the set of equivalence classes of $\mathfrak{g}$-modules $W$ containing $V$ as a codimension 1 submodule such that $gW \subset V$ for every $g \in \mathfrak{g}$. (Equivalence relation: $W' \supset V$ is equivalent to $W'' \supset V$ if there exists an isomorphism $W' \to W''$ which is the identity on $V$).

DEFINITION (compare footnote[9]) in Section 2.1.3.) A one-dimensional central extension of a Lie algebra $\mathfrak{g}$ is a triple $(\widetilde{\mathfrak{g}}, z \in \widetilde{\mathfrak{g}}, \pi: \widetilde{\mathfrak{g}} \to \mathfrak{g})$ where: $\widetilde{\mathfrak{g}}$ is a Lie algebra; $z$ is a non-zero central element of $\widetilde{\mathfrak{g}}$ (that is, $[z, \widetilde{g}] = 0$ for every $\widetilde{g} \in \widetilde{\mathfrak{g}}$); $\pi$ is a Lie algebra homomorphism such that $\operatorname{Im} \pi = \mathfrak{g}$ and $\operatorname{Ker} \pi = \operatorname{span}(z)$. An equivalence $(\widetilde{\mathfrak{g}}_1, z_1, \pi_1) \sim (\widetilde{\mathfrak{g}}_2, z_2, \pi_2)$ is defined as a Lie algebra isomorphism $f: \widetilde{\mathfrak{g}}_1 \to \widetilde{\mathfrak{g}}_2$ such that $f(z_1) = z_2$ and $\pi_1 = \pi_2 \circ f$.

PROPOSITION-CONSTRUCTION. *There is a natural bijection between* $H^2(\mathfrak{g}; \mathbb{K})$ *and the set* $\operatorname{CExt}(\mathfrak{g})$ *of equivalence classes of one-dimensional central extensions of* $\mathfrak{g}$.

---

[11]) An isomorphisms is *natural*, if it is defined for all representations, and a homomorphism between representations gives rise to a commutative diagram involving these isomorphisms.

*Construction-proof.* Let $(\widetilde{\mathfrak{g}}, z, \pi)$ be a one-dimensional central extension of $\mathfrak{g}$. Fix a linear map (not required to be a Lie algebra homomorphism!) $\sigma: \mathfrak{g} \to \widetilde{\mathfrak{g}}$ such that $\pi \circ \sigma = \mathrm{id}_{\mathfrak{g}}$. (It obviously exists.) For every $g, h \in \mathfrak{g}$,

$$\pi([\sigma(g), \sigma(h)] - \sigma([g, h])) = 0$$

(indeed,$\pi([\sigma(g), \sigma(h)]) = [\pi \circ \sigma(g), \pi \circ \sigma(h)] = [g, h] = \pi \circ \sigma([g, h]))$. Thus,

$$[\sigma(g), \sigma(h)] - \sigma([g, h]) \in \operatorname{Ker} \pi, \text{ that is, } [\sigma(g), \sigma(h)] - \sigma([g, h]) = c(g, h) z$$

for a unique $c(g, h) \in \mathbb{K}$. Obviously, the function $c: \mathfrak{g} \times \mathfrak{g} \to \mathbb{K}$ is bilinear and skew-symmetric: $c(h, g) = -c(g, h)$; thus, $c \in C^2(\mathfrak{g}; \mathbb{K})$. Also,

$$\begin{aligned} & [[\sigma(g), \sigma(h)], \sigma(k)] - \sigma([[g, h], k]) \\ =& [[\sigma(g), \sigma(h)], \sigma(k)] - [\sigma([g, h]), \sigma(k)] + [\sigma([g, h]), \sigma(k)] - \sigma([[g, h], k]) \\ =& [c(g, h) z, \sigma(k)] + c([g, h], k) z = c([g, h], k) z \end{aligned}$$

(the last equality holds, because $z$ is central). The Jacobi identity shows that

$$c([g, h], k) + c([h, k], g) + c([k, g], h) = 0,$$

that is, $c$ is a cocycle. In this construction, there was an arbitrary choice of $\sigma$. What happens, if we make a different choice of $\sigma$, that is, replace it by a $\sigma'$ with $\pi \circ \sigma' = \mathrm{id}$? We will have $\pi \circ (\sigma' - \sigma) = 0$, that is $(\sigma' - \sigma)\mathfrak{g} \subset \operatorname{Ker} \pi = \operatorname{span}(z)$, that is, $\sigma'(g) = \sigma(g) + b(g) z$ for some function $b: \mathfrak{g} \to \mathbb{K}$ which is obviously linear. Consider the function $c': \mathfrak{g} \times \mathfrak{g} \to \mathbb{K}$ which is constructed from $\sigma'$ as $c$ was constructed from $\sigma$. We have:

$$\begin{aligned} c'(g, h) z &= [\sigma'(g), \sigma'(h)] - \sigma'([g, h]) = [\sigma(g) + b(g) z, \sigma(h) + b(h) z] - \sigma([g, h]) - b([g, h]) z \\ &= [\sigma(g), \sigma(h)] - \sigma([g, h]) - b([g, h]) z = (c(g, h) - b([g, h])) z, \end{aligned}$$

which shows that $c' = c - \delta b$. We see that the cohomology class of the cocycle $c$ is determined by the central extension, and, obviously, by its equivalence class. Thus, we have a well defined natural map $\mathrm{CExt}(\mathfrak{g}) \to H^2(\mathfrak{g}; \mathbb{K})$.

To show that it is a bijection, we will construct an inverse map $H^2(\mathfrak{g}; \mathbb{K}) \to \mathrm{CExt}(\mathfrak{g})$. Let $\gamma \in H^2(\mathfrak{g}; \mathbb{K}$ be a cohomology class, and let $c \in C^2(\mathfrak{g}; \mathbb{K})$ be a cocycle from the class $\gamma$. Let $\widetilde{\mathfrak{g}}$, as a vector space, be $\mathfrak{g} \times \mathbb{K}$, let $z = (0, 1)$ and let $\pi: \mathfrak{g} \times \mathbb{K} \to \mathfrak{g}$ be the usual projection. We define the commutator in $\widetilde{\mathfrak{g}}$ by the formula $[(g, x), (h, y)] = ([g, h], c(g, h))$. Bilinearity and skew-symmetricity are obvious, the Jacobi identity follows from the fact that $c$ is a cocycle:

$$[[(g, x), (h, y)], (k, w)] = [([g, h], c(g, h)), (k, w)] = ([[g, h], k], c([g, h], k)),$$

and

$$\begin{aligned} & [[(g, x), (h, y)], (k, w)] + [[(h, y), (k, w)], (g, x)] + [[(k, w), (g, x)], (h, y)] \\ =& ([[g, h], k] + [[h, k], g] + [[k, g], h], c([g, h], k) + c([h, k], g) + c([k, g], h)) = (0, 0). \end{aligned}$$

It remains to check that the equivalence class of this central extension does not depend on the choice of a cocycle $c$ within $\gamma$. Let $c' = c - \delta b, b \in C^1(\mathfrak{g}; \mathbb{K})$ be a different choice. The new extension differs from the old extension only be the commutator; we denote the new commutator by the symbol $[\,,\,]'$. Define $f\colon \mathfrak{g} \times \mathbb{K} \to \mathfrak{g} \times \mathbb{K}$ by the formula $f(g,x) = (g, x + b(g))$. Obviously, this is a linear isomorphism ($f^{-1}(g,x) = (g, x - b(g))$), it takes $(0,1)$ into $(0,1)$ and is compatible with the projection $\mathfrak{g} \times \mathbb{K} \to \mathfrak{g}$. Let us check that it is compatible with the commutators.

$$\begin{aligned}
[f(g,x), f(h,y)]' &= [(g, x + b(g)), (h, y + b(h))]' = ([g,h], c'(g,h)) \\
&= ([g,h], c(g,h) + \delta b(g,h)) = ([g,h], c(g,h) + b([g,h]) \\
&= f([g,h], c(g,h)) = f([g,x],[h,y]).
\end{aligned}$$

Thus, we have a well defined map $H^2(\mathfrak{g}; \mathbb{C}) \to \mathrm{CExt}(\mathfrak{g})$, and it is obvious that the maps $\mathrm{CExt}(\mathfrak{g}) \longleftrightarrow H^2(\mathfrak{g}; \mathbb{C})$ are inverse to each other.

Exercise 8. Let $V$ be a $\mathfrak{g}$-module. Define a $V$-extension of $\mathfrak{g}$ as a Lie algebra $\widetilde{\mathfrak{g}}$ with a commutative ideal $W$ and isomorphisms $\widetilde{\mathfrak{g}}/W \to \mathfrak{g}$ and $V \to W$ (the latter is a $\mathfrak{g}$-isomorphism with respect to the structure $gw = [g,w]$ of a $\mathfrak{g}$-module in $W$). Construct a natural bijection between the set of equivalence classes of $V$-extensions of $\mathfrak{g}$ and $H^2(\mathfrak{g}; V)$.

A *deformation* of a Lie algebra $\mathfrak{g}$ is a family (continuous, smooth, algebraic, formal, etc.) of commutators $[\,,\,]_t$ in the space of $\mathfrak{g}$ such that $[\,,\,]_0$ is the commutator in $\mathfrak{g}$. Two deformations, $[\,,\,]_t$ and $[\,,\,]'_t$, are called equivalent, if there exists a family $f_t\colon \mathfrak{g} \to \mathfrak{g}$ of linear isomorphisms (belonging to the same class with respect to $t$, as the families of commutators) such that $[g,h]'_t = f_t^{-1}[f_t h, f_t h]_t$. An *infinitesimal* deformation is the family $[\,,\,]_t$ which satisfies the Jacobi identity up to the terms of order $t^2$. A more algebraic (and more general) definition: if $A$ is an associative algebra over $\mathbb{K}$ with an augmentation $\varepsilon\colon A \to \mathbb{K}$, then the deformation of $\mathfrak{g}$ with the base $A$ (or $\operatorname{spec} A$) is a $A$-Lie algebra structure on $A \otimes \mathfrak{g}$ such that $\varepsilon \otimes \mathrm{id}\colon A \otimes \mathfrak{g} \to \mathfrak{g}$ is a Lie algebra homomorphism. The definition of an $A$-Lie algebra structure is as follows: $\mathfrak{g}$ must be a left $A$-module, and $a[g,h] = [ag,h] + [g,ah]$ for all $a \in A$, $g,h \in \mathfrak{g}$. An equivalence between deformations becomes an $A$-Lie isomorphism. A continuous, smooth, algebraic, formal deformation is a deformation with the base, respectively, $\mathcal{C}(\mathbb{K},\mathbb{K}), \mathcal{C}^\infty(\mathbb{K},\mathbb{K}), \mathbb{K}[t], \mathbb{K}[[t]]$. An infinitesimal deformation is a deformation with the base $\mathbb{K}[t]/(t^2)$.

A more direct description of an infinitesimal deformation of a Lie algebra $\mathfrak{g}$ is a skew-symmetric bilinear function $\alpha\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ such that the "commutator" $[g,h]_t = [g,h] + t\alpha(g,h)$ satisfies the Jacobi identity up to terms of degree $\geq 2$. Two infinitesimal deformations, $\alpha$ and $\beta$ are (infinitesimally) equivalent, if there exists a linear function $\varphi\colon \mathfrak{g} \to \mathfrak{g}$ such that, modulo terms of degree $\geq 2$, the equality $[g + t\varphi(g), h + t\varphi(h)] + t\alpha(g,h) = [g,h] + t(\varphi([g,h]) + \beta(g,h))$ holds.

Exercises. 9. Construct a natural bijection between equivalence classes of infinitesimal deformations of $\mathfrak{g}$ and $H^2(\mathfrak{g}; \mathrm{ad})$.

10. For a cochain $c \in C^2(\mathfrak{g}; \mathrm{ad})$, we can form a cochain $d = [c,c] \in C^3(\mathfrak{g}, ad)$ by the following formula: $d(g,h,k) = c(c(g,h),k) + c(c(h,k),g) + c(c(k,g),h)$.

(a) Prove that if $c$ is a cocycle, then $[c,c]$ is a cocycle.

(b) Prove that if $c, c'$ are cohomologous, then $[c, c], [c', c']$ are cohomologous.

Thus, for a cohomology class $\gamma \in H^2(\mathfrak{g}; \mathrm{ad})$ there arises a cohomology class $[\gamma, \gamma] \in H^3(\mathfrak{g}; \mathrm{ad})$.

(c) Prove that an infinitesimal deformation corresponding (in the sense of Exercise 9) to $\gamma \in H^2(\mathfrak{g}; \mathrm{ad})$ can be extended to a deformation over $\mathbb{K}[t]/(t^3)$ is and only if $[\gamma, \gamma] = 0$.

This theory has an infinite continuation, but we have to stop here.

**3.2.3. The operators $\iota(\mathbf{g})$ and $\pi(\mathbf{g})$.** Let $g \in \mathfrak{g}$. We define homomorphisms

$$\iota(g)\colon C^q(\mathfrak{g}; V) \to C^{q-1}(\mathfrak{g}; V) \text{ and } \pi(g)\colon C^q(\mathfrak{g}; V) \to C^q(\mathfrak{g}; V)$$

by the formulas

$$(\iota(g)c)(g_1, \dots, g_{q-1}) = c(g, g_1, \dots, g_{q-1}),$$
$$(\pi(g)c)(g_1, \dots, g_q) = \sum_{j=1}^{q} c(g_1, \dots, [g, g_j], \dots, g_q) - gc(g_1, \dots, g_q).$$

Proposition 1. (1) *$\pi$ is a Lie algebra representation (that is, $\pi([g, h]) = \pi(g) \circ \pi(h) - \pi(h) \circ \pi(g)$).*

(2) $\pi(g) = \iota(g) \circ \delta^q + \delta^{q-1} \circ \iota(g)$.

(3) *$\pi(g)$ commutes with $\delta$.*

*Proof of* (1). $(\pi(g) \circ \pi(h)c)(g_1, \dots, g_q) =$

$$\sum_{j \neq k} (c(g_1, \dots [g, g_j] \dots [h, g_k], \dots, g_q) + \sum_{i=1}^{q} c(g_1, \dots, [g, [h, g_i]], \dots, g_q)$$
$$- \sum_{i=1}^{q} \big(gc(g_1, \dots [h, g_i] \dots g_q) + hc(g_1, \dots [g, g_i] \dots g_q)\big) + ghc(g_1, \dots, g_q);$$

$$(\pi(h) \circ \pi(g)c)(g_1, \dots, g_q) =$$

$$\sum_{j \neq k} (c(g_1, \dots [g, g_j] \dots [h, g_k], \dots, g_q) + \sum_{i=1}^{q} c(g_1, \dots, [h, [g, g_i]], \dots, g_q)$$
$$- \sum_{i=1}^{q} \big(gc(g_1, \dots [h, g_i] \dots g_q) + hc(g_1, \dots [g, g_i] \dots g_q)\big) + hgc(g_1, \dots, g_q);$$

$$(\pi([g, h])c)(g_1, \dots, g_q) = \sum_{i=1}^{q} c(g_1, \dots, [[g, h], g_i], \dots, g_q) - [g, h]c(g_1, \dots, g_q).$$

It is seen from this formula that $\pi([g, h])c = \pi(g) \circ \pi(h)c - \pi(h) \circ \pi(g)c$ (checking involves the equality $[g, h] = gh - hg$ and the Jacobi identity).

*Proof of* (2).

$$\delta^{q-1}\iota(g)c(g_1, \dots, g_q) = \sum_{1 \le s < t \le q} (-1)^{s+t-1} c(g, [g_s, g_t], g_1, \dots \widehat{g}_s \dots \widehat{g}_t \dots, g_q)$$
$$+ \sum_{1 \le u \le q} (-1)^u g_u c(g, g_1, \dots \widehat{g}_u \dots, g_q);$$

$$\begin{aligned}\iota(g)\delta^q c(g_1,\ldots,g_q) &= \delta^q c(g,g_1,\ldots,g_q)\\ &= \sum_{1\le t\le q}(-1)^{1+(t+1)-1}c([g,g_t],g_1,\ldots\widehat{g}_t\ldots,g_q)\\ &+\sum_{1\le s<t\le q}(-1)^{(s+1)+(t+1)-1}c([g_s,g_t],g,g_1,\ldots\widehat{g}_s\ldots\widehat{g}_t\ldots,g_q)\\ &- gc(g_1,\ldots,q_q)+\sum_{1\le u\le q}(-1)^{u+1}g_u c(g,g_1,\ldots\widehat{g}_u\ldots,g_q).\end{aligned}$$

In the sum $\delta^{q-1}\iota(g)c(g_1,\ldots,g_q)+\iota(g)\delta^q c(g_1,\ldots,g_q)$, the two terms in the first formula cancel with the second and fourth terms in the second formula. What remains is

$$\sum_{1\le t\le q}(-1)^{t+1}c([g,g_t],g_1,\ldots\widehat{g}_t\ldots,g_q)-gc(g_1,\ldots,g_q)=\pi(g)c(g_1,\ldots,g_q).$$

*Proof of* (3). We want to prove that the diagram

$$\begin{array}{ccc} C^{q-1}(\mathfrak{g};V) & \xrightarrow{\pi(g)} & C^{q-1}(\mathfrak{g};V)\\ \Big\downarrow{\scriptstyle\delta^{q-1}} & & \Big\downarrow{\scriptstyle\delta^{q-1}}\\ C^{q}(\mathfrak{g};V) & \xrightarrow{\pi(g)} & C^{q}(\mathfrak{g};V)\end{array}$$

is commutative. Here is the proof: $\pi(g)\circ\delta^{q-1}=(\iota(g)\circ\delta^q+\delta^{q-1}\circ\iota(g))\circ\delta^{q-1}=\delta^{q-1}\circ\iota(g)\circ\delta^{q-1}$ and $\delta^{q-1}\circ\pi(g)=\delta^{q-1}\circ(\iota(g)\circ\delta^{q-1}+\delta^{q-2}\circ\iota(g))=\delta^{q-1}\circ\iota(g)\circ\delta^{q-1}$.

Since $\pi$ provides a structure of a $\mathfrak{g}$-module in the cochain spaces of $\mathfrak{g}$ and this siructure is compatible with the coboundary operator, there arises a $\mathfrak{g}$-module structure in the cohomology of $\mathfrak{g}$. However, this structure is always trivial.

Proposition 2. *For every $\gamma\in H^q(\mathfrak{g};V)$ and every $g\in\mathfrak{g}$, $\pi(g)\gamma=0$.*

*Proof.* We need to show that if $c\in C^q(\mathfrak{g}:V)$ is a cocycle, then for every $g\in\mathfrak{g}$, $\pi(g)c$ is a coboundary. Here is a proof:

$$\pi(g)c=\iota(g)\delta c+\delta\iota(g)c=\delta\iota(g)c.$$

**3.2.4. Cohomology of semisimple Lie algebras. Whitehead's lemma.** Here is a brief account of results concerning the cohomology of semisimple Lie algebras. First of all, there is a theorem known as the *Second Whitehead Lemma* which states that for every representation $V$ of a semisimple Lie algebra $\mathfrak{g}$ (both are assumed finite dimensional) the inclusion map $V^\rho\to V$ of the invariant space induces a cohomology isomorphism: for any $q\ge 0$,

$$H^q(\mathfrak{g};V)\cong H^q(\mathfrak{g};V^\rho)=\underbrace{H^q(\mathfrak{g};\mathbb{K})\oplus\ldots\oplus H^q(\mathfrak{g};\mathbb{K})}_{\dim V^\rho}.$$

As to $H^q(\mathfrak{g};\mathbb{R})$, if $\mathfrak{g}$ is the Lie algebra of a compact Lie group $G$, then

$$H^q(\mathfrak{g};\mathbb{R})=H^q_{DR}(G)$$

(de Rham cohomology of the manifold $G$). (A reader familiar with the de Rham cohomology will be able to prove this using the group averaging.) We will prove later (in Section 4.2) that for a real semisimple Lie algebra $\mathfrak{g}$, there exists a compact Lie group $\overline{G}$ such that there is an isomorphism between the complexifications:

$$(\operatorname{Lie}\overline{G}) \otimes \mathbb{C} \cong \mathfrak{g} \otimes \mathbb{C},$$

which shows that

$$H^q(\mathfrak{g}; \mathbb{R}) \cong H^q_{DR}(\overline{G});$$

for example,

$$\mathfrak{sl}(n; \mathbb{R}) \otimes \mathbb{C} = \mathfrak{su}(n) = \mathfrak{gl}(n; \mathbb{C}),$$

and, consequently,

$$H^q(\mathfrak{sl}(n, \mathbb{R}); \mathbb{R}) = H^q(\mathfrak{su}(n); \mathbb{R}) = H^q_{DR}(SU(n)) = H^q_{DR}(S^3 \times S^5 \times \ldots \times S^{2n-1}).$$

Also,

$$H^q(\mathfrak{sl}(n, \mathbb{C}); \mathbb{C}) = H^q_{DR}(SU(n); \mathbb{C}) = H^q_{DR}(S^3 \times S^5 \times \ldots \times S^{2n-1}; \mathbb{C}).$$

It is true also that for every real and complex semisimple Lie algebra its cohomology with coefficients in the trivial 1-dimensional module is the same as the de Rham cohomology of a product of spheres of odd dimensions $\geq 3$. In particular, $H^q(\mathfrak{g}; V) = 0$ for any semisimple $\mathfrak{g}$, any (finite-dimensional) $V$, and $q = 1, 2$. This statement is called the *First Whitehead Lemma*, and this is what we are going to prove here.

THEOREM. *For any semisimple* $\mathfrak{g}$ *and any* $V$, $H^1(\mathfrak{g}; V) = H^2(\mathfrak{g}; V) = 0$.

*Proof for* $H^1$. Let $\operatorname{Ker}\delta^1 = Z$, $\operatorname{Im}\delta^0 = B$. By Theorem from Section 2.5.3.2, as a $\mathfrak{g}$-module, $Z = Z^\pi \oplus Z_\pi$. If $c \in Z$ and $g \in \mathfrak{g}$, then $\pi(g)c = \delta^0\iota(g)c \in B$, so $Z_\pi \subset B$ (we already noticed this in the proof of Proposition 2 in Section 3.2.3). Furthermore, if $C \in Z^\pi$ then for $g, h \in \mathfrak{g}$, $\delta^0\iota(g)c(h) = c([g, h]) = 0$, thus $c(\mathfrak{g}') = 0$ and hence $c = 0$, since $\mathfrak{g}' = \mathfrak{g}$ (Corollary 2 in Section 2.5.1). Thus, $Z_\pi = 0$, $Z \subset B$, and $Z/B = H^1(\mathfrak{g}; V) = 0$.

*Proof for* $H^2$. Let $\operatorname{Ker}\delta^2 = Z$, $\operatorname{Im}\delta^1 = B$. As before, $Z = Z^\pi \oplus Z_\pi$ and $Z_\pi \subset B$. Let $c \in Z^\pi$. Then, for a $g \in \mathfrak{g}$, $\pi(g)c = 0$, hence $\delta^1\iota(g) = 0$, and, since $H^1(\mathfrak{g}; V) = 0$, $\iota(g)c = \delta^0 a$ for some $a \in C^0(\mathfrak{g}; V) = V$. This $b$ is defined up to an element of $\operatorname{Ker}\delta^0 = V^\rho$ ($\rho$ is the representation of $\mathfrak{g}$ in $V$), hence, $a$ becomes unique, if we require that $a \in V_\rho$. In other words, there exists a (unique) linear map $b\colon \mathfrak{g} \to \mathfrak{V}_\pi$ such that $\iota(g)c = \delta^0 b(g)$ for every $g \in \mathfrak{g}$. This $b$ may be regarded as an element of $C^1(\mathfrak{g}; V_\rho) \subset C^1(\mathfrak{g}; V)$. We are going to show that $c = -\delta^1 b$. From this it will follow that $Z^\pi \subset B$, and $H^2(\mathfrak{g}; V) = 0$.

The equality $\pi(g)c(h, k) + \pi(h)c(k, g) + \pi(k)c(g, h) = 0$ means that

$$\begin{aligned} &c([g, h], k) + c(h, [g, k]) - gc(h, k) \\ +&c([h, k], g) + c(k, [h, g]) - hc(k, g) \\ +&c([k, g], h) + c(g, [k, h]) - kc(g, h) = 0. \end{aligned}$$

The first and the third column add up to $\delta^2 c(g, h, k) = 0$. Thus,

$$c(h, [g, k]) + c(k, [h, g]) + c(g, [k, h]) = 0$$

or

$$c(h,[g,k])+c([g,h],k)=c(g,[h,k]).$$

Plugging this into the equality $\pi(g)c(h,k)=c([g,h],k)+c(h,[g,k])-gc(h,k)=0$, we get

$$c(g,[h,k])=gc([h,k]).$$

We want to prove that

$$c=\delta^1 b, \text{ that is, } c(g,h)=-b(g,h])+gb(h)-hb(g),$$

and we have

$$\iota(g)c=\delta^0 b(g), \text{ that is, } c(g,h)=-hb(g).$$

Together, this means that

$$-b([g,h])+gb(h)=0$$

and, since the left hand side of his equality belongs to $V_\rho$, it is sufficient to prove that is belongs to $V^\rho$. Here is a proof:

$$\begin{aligned}k(-b([g,h])+gb(h)) &= c([g,h],k)-kc(h,g)\\ &= c([g,h],k)-c(k,[h,g])=0.\end{aligned}$$

### 3.3. Two important applications of the Lie algebra cohomology.

**3.3.1. The Weyl theorem.** *Every (finite-dimensional) representation of a semisimple Lie algebra Is decomposable into a direct sum of irreducible representations.*

*Proof.* We need to prove the following statement: *if $\rho\colon\mathfrak{g}\to\mathfrak{gl}(V)$ is a representation of a semisimple Lie algebra and $W\subset V$ is a subrepresentation, then there exists a $\mathfrak{g}$-module projection of $V$ onto $W$, that is a linear map $p\colon V\to V$ such that $p(V)=W$, $p^2=p$ and $[p,\rho(\xi)]=0$ for every $\xi\in\mathfrak{g}$.* Indeed, for such $p$, $\operatorname{Ker}p$ is also a subrepresentation of $\rho$, and $V=W\oplus\operatorname{Ker}p$.

Choose an arbitrary linear projection $q$ of $V$ onto $W$ (that is, $q\colon V\to V$, $q(V)=W$, $q^2=q$). Then all other projections have the form $q+r$ where $r\colon V\to V$ satisfies the properties $r(V)\subset W$, $r(W)=0$; let $R\subset\mathfrak{gl}(V)$ be the space of maps with these properties. For a $\xi\in\mathfrak{g}$ and an $r\in R$, $[\rho(\xi),r]\in R$; indeed, $[\rho(\xi),r](V)\subset\rho(\xi)r(V)+r(V)\subset W$ and $[\rho(\xi),r](W)\subset r(W)=\rho(\xi)r(W)=0$. Hence $\sigma(\xi)r=[\rho(\xi),r]$ is a representation of $\mathfrak{g}$ in $R$. Thus, if $V$ is not irreducible, it is a sum of two non-zero subrepresentations, and we proceed by induction with respect to $\dim V$.

For a $\xi\in\mathfrak{g}$, $[q,\rho(\xi)]\in R$; indeed, $[q,\rho(\xi)](V)\subset q(V)+\rho(\xi)\subset W$ and $[q,\rho(\xi)](W)\subset q(W)+\rho(\xi)q(W)=0$. Hence, a function $c(\xi)=[q,\rho(\xi)]$ is a cochain from $C^1(\mathfrak{g};R)$. This is a cocycle:

$$\begin{aligned}c([\xi,\eta])=[q,\rho([\xi,\eta])]=[q,[\rho(\xi),\rho(\eta)]]&=[[q,\rho(\xi)],\rho(\eta)]+[\rho(\xi),[q,\rho(\eta)]]\\ =[\rho(\xi),[q,\rho(\eta)]]-[\rho(\eta),[q,\rho(\xi)]&=\sigma(\xi)c(\eta)-\sigma(\eta)c(\xi).\end{aligned}$$

By the Whitehead lemma, there exists some $r\in R=C^0(\mathfrak{g};)$ such that $\delta^0 r(\xi)=\sigma(\xi)r=c(\xi)$, that is, $[\rho(\xi),r]=[q,\rho(\xi)]$, that is, $[q+r,\rho(\xi)]=0$ for all $\xi$, and we can put $p=q+r$.

**3.3.2. The Levi decomposition.** Let $\mathfrak{h}$ be a (real or complex) Lie algebra. The quotient $\mathfrak{g} = \mathfrak{h}/\operatorname{rad}\mathfrak{h}$ is semisimple (if $\mathfrak{s}$ is a non-trivial solvable ideal of $\mathfrak{g}$, then $p^{-1}(\mathfrak{s})$, where $p$ is a projection $\mathfrak{h} \to \mathfrak{g}$, is a solvable ideal of $\mathfrak{h}$ which is not contained in $\operatorname{rad}\mathfrak{h}$).

THEOREM. *There exists a Lie algebra homomorphism $\nu: \mathfrak{g} \to \mathfrak{h}$ such that $p \circ \nu = \mathrm{id}$; in other words, there exists a Lie subalgebra $\widetilde{\mathfrak{g}}$ of $\mathfrak{h}$ such that the restriction of $p$ to $\widetilde{\mathfrak{g}}$ is a Lie algebra isomorphism $\widetilde{\mathfrak{g}} \to \mathfrak{g}$.*

In algebra, this situation is described by by the words: $\mathfrak{h}$ is a semi-direct product of $\operatorname{rad}\mathfrak{h}$ and $\mathfrak{g}$; thus, every Lie algebra is a semi-direct product of a solvable Lie algebra and a semisimple Lie algebra.

*Proof of Theorem. Case* 1: $\operatorname{rad}\mathfrak{h}$ is commutative. For a $\xi \in \mathfrak{g}$, choose a $\widetilde{\xi} \in \mathfrak{h}$ such that $p(\widetilde{\xi}) = \xi$. Then, for an $\omega \in \operatorname{rad}\mathfrak{h}$, $[\widetilde{\xi}, \omega] \in \operatorname{rad}\mathfrak{h}$ does not depend on the choice of $\widetilde{\xi}$; if $\widetilde{\xi}'$ is a different choice, then $\widetilde{\xi}' - \widetilde{\xi} \in \operatorname{rad}\mathfrak{h}$ and $[\widetilde{\xi}', \omega] - [\widetilde{\xi}, \omega] = [\widetilde{\xi}' - \widetilde{\xi}, \omega] = 0$. The formula $\rho(\xi)\omega = [\widetilde{\xi}, \zeta]$ determines a representation $\rho$ of $\mathfrak{g}$ in $\operatorname{rad}\mathfrak{h}$.

Let $\mu: \mathfrak{g} \to \mathfrak{h}$ is a linear map (not necessarily a Lie algebra homomorphism) such that $p \circ \mu = \mathrm{id}$. (Thus, $\rho(\xi)\omega = [\mu(\xi), \omega]$.) For $\xi, \eta \in \mathfrak{g}$ put

$$\theta(\xi, \eta) = [\mu(\xi), \mu(\eta)] - \mu([\xi, \eta]);$$

we can consider $\theta$ as a cochain in $C^2(\mathfrak{g}; \operatorname{rad}\mathfrak{h})$ where $\operatorname{rad}\mathfrak{h}$ is furnished by a structure of a $\mathfrak{g}$-module by the representation $\rho$; thus, $\mu$ is a Lie algebra homomorphism if and only if $\theta = 0$.

Let us prove that $\theta$ is a cocycle:

$$\begin{aligned}\delta^2\theta(\xi, \eta, \zeta) = {} & [\mu([\xi, \eta]), \mu(\zeta)] - \mu([[\xi, \eta], \zeta]) - [\mu(\xi), [\mu(\eta), \mu(\zeta)] + [\mu(\xi), \mu([\eta, \zeta])] \\ & + [\mu([\eta, \zeta]), \mu(\xi)] - \mu([[\eta, \zeta], \xi]) - [\mu(\eta), [\mu(\zeta), \mu(\xi)] + [\mu(\eta), \mu([\zeta, \xi])] \\ & + [\mu([\zeta, \xi]), \mu(\eta)] - \mu([[\zeta, \xi], \eta]) - [\mu(\zeta), [\mu(\xi), \mu(\eta)] + [\mu(\zeta), \mu([\xi, \eta])] = 0\end{aligned}$$

(the first column cancels with the last column, and two middle columns sum up to zeroes by the Jacobi identity).

By the Whitehead lemma, $\theta = \delta^1\tau$ for some $\tau \in C^1(\mathfrak{g}; \operatorname{rad}\mathfrak{h}) = \operatorname{Hom}(\mathfrak{g}, \operatorname{rad}\mathfrak{h})$ that is,

$$\theta(\xi, \eta) = \tau([\xi, \eta]) - [\mu(\xi), \tau(\eta)] + [\mu(\eta), \tau(\xi)].$$

Thus,

$$\begin{aligned}&[\mu(\xi), \mu(\eta)] + [\mu(\xi), \tau(\eta)] + [\tau(\xi), \mu(\eta)] - \mu([\xi, \eta]) - \tau([\xi, \eta]) \\ &\qquad = [(\mu + \tau)(\xi), (\mu + \tau)(\eta)] - (\mu + \tau)([\xi, \eta]) = 0\end{aligned}$$

(we use the fact that $[\tau(\xi), \tau(\eta)] = 0$), and we can take $\mu + \tau$ for $\nu$.

*Case 2:* $\operatorname{rad}\mathfrak{h}$ *is not commutative.* We use the induction with respect to $\dim\operatorname{rad}\mathfrak{h}$ (if this is 0, then $\mathfrak{h} = \mathfrak{g}$ and we have nothing to prove). Consider $\widetilde{\mathfrak{h}} = \mathfrak{h}/(\operatorname{rad}\mathfrak{h})'$ ($(\operatorname{rad}\mathfrak{h})'$ is an ideal in $\mathfrak{h}$; actually, it is true that if $\mathfrak{k}$ is an ideal in a Lie algebra $\mathfrak{h}$, then $(\mathfrak{k})'$ is also an ideal in $\mathfrak{h}$: is $\xi, \eta \in \mathfrak{k}, \zeta \in \mathfrak{h}$, then $[\zeta, [\xi.\eta]] = [[\zeta, \xi], \eta] + [\xi, [\zeta, \eta]] \in \mathfrak{k}'$). Since

$\widetilde{\mathfrak{h}}/((\operatorname{rad}\mathfrak{h})/(\operatorname{rad}\mathfrak{h})') = \mathfrak{h}/\operatorname{rad}\mathfrak{h} = \mathfrak{g}$ is semisimple, and $\operatorname{rad}\mathfrak{h}/(\operatorname{rad}\mathfrak{h})'$ is solvable (commutative), $\operatorname{rad}\widetilde{\mathfrak{h}} = \operatorname{rad}\mathfrak{h}/(\operatorname{rad}\mathfrak{h})'$. Then, by the induction hypothesis (or by Case 1), there exists a Lie algebra homomorphism $\widetilde{\nu}\colon \mathfrak{g} \to \widetilde{\mathfrak{h}}$ whose composition with the projection $\widetilde{\mathfrak{h}} \to \mathfrak{g}$ is the identity. Let $\widetilde{\mathfrak{g}} \subset \mathfrak{h}$ be the inverse image of $\widetilde{\nu}(\mathfrak{g})$ with respect to the projection $\mathfrak{h} \to \widetilde{\mathfrak{h}}$. Then $\widetilde{\mathfrak{g}}/(\operatorname{rad}\mathfrak{h})' = \widetilde{\nu}(\mathfrak{g}) \cong \mathfrak{g}$ is semisimple, hence $\operatorname{rad}\widetilde{\mathfrak{g}} = (\operatorname{rad}\mathfrak{h})'$, and, again by the induction hypothesis, there exists a Lie homomorphism $\lambda\colon \widetilde{\nu}(\mathfrak{g}) \to \widetilde{\mathfrak{g}} \subset \mathfrak{h}$ whose composition with the projection $\widetilde{\mathfrak{g}} \to \widetilde{\nu}(\mathfrak{g})$ is the identity. It remains to put $\nu = \lambda \circ \widetilde{\nu}$.

**3.4. Roots and Cartan algebras for a semisimple Lie algebra.** Let $\mathfrak{g}$ be a complex semisimple Lie algebra, and let $\mathfrak{h} \subset \mathfrak{g}$ be a Cartan algebra.

Theorem 1. *Let $\alpha, \beta$ be two roots of $\mathfrak{g}$. Then*

(1) *if $\alpha + \beta \neq 0$, then $\mathfrak{g}^\alpha \perp \mathfrak{g}^\beta$;*

(2) *if $\alpha + \beta = 0$, then the form $\langle \mathfrak{g}^\alpha, \mathfrak{g}^\beta \rangle$ is non-degenerate.*

*Proof of* (1). For $g_1 \in \mathfrak{g}^\alpha$, $g_2 \in \mathfrak{g}^\beta$, the composition $\operatorname{ad} g_1 \circ \operatorname{ad} g_2$ maps $\mathfrak{g}^\gamma$ into $\mathfrak{g}^{\gamma+\alpha+\beta}$. Hence, if $\alpha + \beta \neq 0$, then the matrix of the operator $\operatorname{ad} g_1 \circ \operatorname{ad} g_2$ with respect to a basis composed of bases of the root spaces has no non-zero diagonal entries, and its trace is 0.

*Proof of* (2). If $\langle g, g' \rangle = 0$ for some $g \in \mathfrak{g}^\alpha$ and all $g' \in \mathfrak{g}^{-\alpha}$, then $g \perp \mathfrak{g}$, and hence $g = 0$ (the Killing form of a semisimple Lie algebra is non-degenerate).

Corollary of (2): $d(\alpha) = d(-\alpha)$; *in particular, if $\alpha$ is a root, then $-\alpha$ is a root.*

Theorem 2. (1) *The restriction of the Killing form to $\mathfrak{h}$ is non-degenerate;*

(2) *roots span $\mathfrak{h}^*$;*

(3) *$\mathfrak{h}$ is commutative.*

*Proof.* (1) follows from part (2) of Theorem 1. If for an $h \in \mathfrak{h}$, $\alpha(h) = 0$ for all roots $\alpha$, then Proposition 3 of Section 3.1.2 shows that $\langle h, \mathfrak{h} \rangle = 0$ which implies, in virtue of (1), that $h = 0$; this yields (2). Finally, every root is zero on $\mathfrak{h}'$ (Proposition 2 of Section 3.1.2), which shows, by (2), that $\mathfrak{h}' = 0$.

The restriction of the Killing form to $\mathfrak{h}$ (non-degenerate by Part (1) of Theorem 2) gives rise to the canonical isomorphism $\mathfrak{h}^* \leftrightarrow \mathfrak{h}$ ($\lambda \in \mathfrak{h}^* \leftrightarrow h_\lambda \in \mathfrak{h}, \lambda(h) = \langle h_\lambda, h \rangle$). This gives a non-degenerate symmetric bilinear form on $\mathfrak{h}^*$: $\langle \lambda, \mu \rangle = \langle h_\lambda, h_\mu \rangle = \lambda(h_\mu) = \mu(h_\lambda)$.

Theorem 3. (1) *If $\alpha$ is a non-zero root and $e \in \mathfrak{g}^\alpha, f \in \mathfrak{g}^{-\alpha}$, then $[e, f] = \langle e, f \rangle h_\alpha$. In particular, $h_\alpha \in [\mathfrak{g}^\alpha, \mathfrak{g}^{-\alpha}]$.*

(2) *If $\alpha$ is a non-zero root, then $\langle \alpha, \alpha \rangle \neq 0$.*

*Proof.* (1) For an $h \in \mathfrak{h}$, $\langle h, [e, f] \rangle = \langle [h, e], f \rangle = \alpha(h) \langle e, f \rangle = \langle h, \langle e, f \rangle h_\alpha \rangle$. Thus $[e, f] = \langle e, f \rangle h_\alpha$. Since $\langle e, f \rangle \neq 0$ for some $e \in \mathfrak{g}^\alpha, f \in \mathfrak{g}^{-\alpha}$ (Part (2) of Theorem 1), $h_\alpha \in [\mathfrak{g}^\alpha, \mathfrak{g}^{-\alpha}]$.

(2) $\langle \alpha, \alpha \rangle = \alpha(h_\alpha)$. Proposition in Section 3.1.4 if $\alpha(h_\alpha) = 0$, then $\beta(h_\alpha) = 0$ for an arbitrary root $\beta$. Thus, in this case $h_\alpha = 0$, a contradiction.

Theorem 4. *Let $\alpha$ be a non-zero root. Then* (1) $\dim \mathfrak{g}^\alpha = 1$ *and* (2) *$k\alpha$ is a root only if $k = -1, 0, 1$.*

*Proof.* Choose arbitrary $e \in \mathfrak{g}^\alpha$ and $f \in \mathfrak{g}^{-\alpha}$ with $\langle e, f \rangle = 1$, that is, $[e, f] = h_\alpha$. Consider the space $V = \mathbb{C}f \oplus \mathbb{C}h_\alpha \oplus \left[ \bigoplus_{k \geq 1} \mathfrak{g}^{k\alpha} \right] \subset \mathfrak{g}$. Obviously, $V$ is invariant with respect to $e, f$, and $h_\alpha$. Hence,

$$0 = \operatorname{Tr} h_\alpha|_V = \alpha(h_\alpha)(-1 + 0 + d(\alpha) + 2d(2\alpha) + 3d(3\alpha) + \dots)$$

(where, as before, $d(\beta) = \dim \mathfrak{g}^\beta$) which shows, since $\alpha(h_\alpha) = \langle\alpha, \alpha\rangle \neq 0$, that $d(\alpha) = 1, d(k\alpha) = 0$ for $k > 1$.

THEOREM 5. *A subalgebra* $\mathfrak{h}$ *of a complex semisimple Lie algebra* $\mathfrak{g}$ *is a Cartan algebra, if and only if it satisfies two conditions:* (A) $\mathfrak{h}$ *is a maximal commutative subalgebra of* $\mathfrak{g}$*;* (B) *for every* $h \in \mathfrak{h}$*, the operator* $\operatorname{ad} h \colon \mathfrak{g} \to \mathfrak{g}$ *is diagonalizable.*

*Proof.* Cartan $\Rightarrow$ (A)&(B). If $\mathfrak{h}$ is a Cartan algebra, then it is commutative by Part (3) of Theorem 2, and is maximal commutative, because it is maximal nilpotent. Since all root spaces (except $\mathfrak{h}$) are one-dimensional, we can form a basis in $\mathfrak{g}$ from an arbitrary basis of $\mathfrak{h}$ adding one non-zero vector from every $\mathfrak{g}^\alpha, \alpha \neq 0$. For every $h \in \mathfrak{h}$, the operator $\operatorname{ad} h$ has a diagonal matrix in this basis.

(A)&(B) $\Rightarrow$ Cartan. If $\mathfrak{h}$ satisfies Conditions (A) and (B), then, with respect to $\mathfrak{h}$, $\mathfrak{g} = \mathfrak{g}^0 \oplus \left[\bigoplus_{\alpha\neq 0}\right]$, from (B) follows that $\mathfrak{g}^\alpha$ consists of eigenvectors of $\operatorname{ad} h$ with eigenvalue $\alpha(h)$ for every $h \in \mathfrak{h}$, hence, $\mathfrak{g}^0$ commutes with $\mathfrak{h}$, and if there is a $g \in \mathfrak{g}^0 - \mathfrak{h}$, then $\mathfrak{h} + \mathbb{C}g$ is commutative, in contradiction with (1).

We will finish this section by making more explicit a construction which was implicitly used in several proofs above. Let $\alpha$ be a nonzero root. Let

$$\mathfrak{s}_\alpha = \mathfrak{g}^\alpha \oplus \mathbb{C}h_\alpha \oplus \mathfrak{g}^{-\alpha} \subset \mathfrak{g}.$$

This 3-dimensional subspace of $\mathfrak{g}$ is a Lie subalgebra. Moreover, this Lie subalgebra is isomorphic to $\mathfrak{sl}(2, \mathbb{C})$. Indeed, choose $e_\alpha \in \mathfrak{g}^\alpha$ and $e_{-\alpha} \in \mathfrak{g}^{-\alpha}$ such that $\langle e_\alpha, e_{-\alpha}\rangle = -1$. Then $[e_\alpha, e_{-\alpha}] = -h_\alpha$, $[h_\alpha, e_\alpha] = \langle\alpha, \alpha\rangle e_\alpha$, $[h_\alpha, e_{-\alpha}] = -\langle\alpha, \alpha\rangle e_{-\alpha}$. The Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ has a basis $e = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $h = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $f = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ with the commutator relations $[e, f] = h, [h, e] = 2e, [h, f] = -2f$, and the formulas $e_\alpha \mapsto \lambda e, h_\alpha \mapsto \mu h, e_{-\alpha} \mapsto \nu f$ with $\mu = \dfrac{\langle\alpha, \alpha\rangle}{2}$, $\lambda\nu = -\mu$ establish an isomorphism $\mathfrak{s}_\alpha \cong \mathfrak{sl}(2, \mathbb{C})$.

**3.5. A deviation: representations of $\mathfrak{sl}(2)$.** The last observation in the previous section shows that every complex semisimple Lie algebra $\mathfrak{g}$ contains subalgebras isomorphic to $\mathfrak{sl}(2, \mathbb{C})$, so there arise representations of $\mathfrak{sl}(2, \mathbb{C})$ is $\mathfrak{g}$. We can expect (and these expectations will come true) that the representation theory of $\mathfrak{sl}(2, \mathbb{C})$ may be important for studying semisimple Lie algebras. We will give here a full description of finite-dimensional representations of $\mathfrak{sl}(2, \mathbb{C})$. Actually, similar results exist for $\mathfrak{sl}(2, \mathbb{F})$ for any field $\mathbb{F}$ of zero characteristic (we will briefly address this subject in exercises in the end of the section).

By the Weyl theorem (Section 3.3.1), every representation of $\mathfrak{sl}(2, \mathbb{C})$ is a direct sum of irreducible representations, so we can restrict ourselves to studying irreducible representations. Remind that $\mathfrak{sl}(2, \mathbb{C})$ has a basis $f, e, h$ with the commutator $[e, f] = h, [h, e] = 2e, [h, f] = -2f$.

Let $n \geq 0$. We define an $(n+1)$-dimensional representation $V_n$ as the space of homogeneous polynomials of two variables $x, y$ of degree $n$ with $e, f, h$ acting as, respectively,

$y\dfrac{\partial}{\partial x}, x\dfrac{\partial}{\partial y}, x\dfrac{\partial}{\partial x} - y\dfrac{\partial}{\partial y}$. If we denote $x^k y^{n-k}$ as $v_k$ $(0 \le k \le n)$, then the action of $e, f, h$ becomes $ev_k = kv_{k-1}$, $fv_k = (n-k)v_{k+1}$, $hv_k = (n-2k)v_k$. This representation is irreducible: if a subrepresentation $W$ of $V$ contains a polynomial $p(x,y) = x^k y^{n-k} + \dots$ where "..." is the sum of monomials of degrees less than $k$ in $x$, then $W$ also contains $y^n = \dfrac{1}{k!}\left(y\dfrac{\partial}{\partial x}\right)^k p$, and also $x^\ell y^{n-\ell}$ (for every $\ell$) as $\dfrac{n!}{(n-\ell)!}\left(x\dfrac{\partial}{\partial y}\right)^\ell y^n$.

THEOREM. *Every irreducible representation of $\mathfrak{sl}(2,\mathbb{C})$ is isomorphic to one of the representations $V_n$.*

*Proof.* Let $V$ be an irreducible representation of $\mathfrak{sl}(2,\mathbb{C})$. Let $w$ be an eigenvector of the operator $h$, $hw = \mu w$. Then $hew = ehw + 2ew = (\mu+2)ew$, thus $ew$ also is an eigenvector of $h$ with an eigenvalue $\mu + 2$. Similarly, $fw$ is an eigenvector of $h$ with the eigenvalue $\mu - 2$. Hence, $e^m w$ and $f^m w$ are eigenvectors of $h$ with eigenvalues $\mu + 2m$ and $\mu - 2m$. Since the operator $h$ has finitely many eigenvalues, the sequence $w, ew, e^2w, \dots$ becomes zero after some $e^m w \neq 0$, that is, $e^{m+1}w = 0$. Put $e^m w = v$ and consider the sequence $v, fv, f^2v, \dots$; it also becomes a sequence of zeroes after some $f^n v \neq 0$, that is, $f^{n+1}v = 0$. The vectors $v, fv, \dots, f^n v$ are linearly independent (they are eigenvalues of $h$ with different eigenvalues, $\lambda, \lambda-2, \dots, \lambda-2n$ where $\lambda = \mu+2m$) and they span an $(n+1)$-dimensional subspace of $V$ invariant with respect to $e, f, h$. Since our representation is irreducible, this subspace must be $V$, and $\{v, fv, f^2v, \dots, f^n v\}$ is a basis in it. Obviously, for $k = 1, \dots, n$, $e(f^k v) = a_k f^{k-1} v$ for some $a_k \in \mathbb{C}$. There are relations between these $a_k$:

$$a_k f^{k-1}v = ef^k v = (ef)f^{k-1}v = (fe+h)f^{k-1}v = (a_{k+1} + \lambda - 2(k-1))f^{k-1}v$$

which yields the equation

$$a_k - a_{k-1} = \lambda - 2(k-1). \tag{$*$}$$

Also,

$$hv = \lambda v = (ef - fe)v = a_1 v,\ \ hf^n v = (\lambda - 2n)f^n v = (ef - fe)f^n v = -a_n f^n v$$

which shows that

$$a_1 = \lambda,\ \ a_n = 2n - \lambda.$$

Thus,

$$a_k = \lambda + (\lambda - 2) + (\lambda - 4) + \dots + (\lambda - 2(k-1)) = k\lambda - k(k-1)$$

and the equality $a_n = 2n - \lambda$ becomes $n\lambda - n(n-1) = 2n - \lambda$, that is, $(n+1)\lambda = n(n+1)$, that is, $\lambda = n$, so $hf^k v = (\lambda - 2k)f^k v = (n-2k)f^k v$ and $a_k = kn - k(k-1) = k(n-k+1)$. Put $v_k = (n-k)!f^k v$; we have:

$$ev_k = (n-k)!ef^k v = (n-k)!k(n-k+1)f^{k-1}v = k(n-k+1)!f^{k-1}v = kv_{k-1};$$
$$fv_k = (n-k)!ff^k v = (n-k)!f^{k+1}v = (n-k)(n-k-1)!f^{k+1}v = (n-k)v_{k+1};$$
$$hv_k = (n-k)!hf^k v = (n-k)!(n-2k)f^k v = (n-2k)(n-k)!f^k v = (n-2k)v_k.$$

Thus, our representation is isomorphic to $V_n$.

EXERCISES 11. Generalize Theorem to the case of an arbitrary field of zero characteristic. (In the case of an algebraically closed field, no changes are needed; in general case, $\lambda$ will belong to the algebraic closure of the given field, which requires an additional attention.)

12. For an integer $n$, consider an (infinite-dimensional) representation of $\mathfrak{sl}(2,\mathbb{C})$ spanned by vectors $f^k v, k \geq 0$, with the obvious action of $f$ and the action of $e$ and $h$ given by the familiar formulas $h(f^k v) = (n-2k)f^k v$, $e(f^k v) = k(k-n-1)f^{k-1}v$ (thus, if $ev = 0$). Prove that if $n < 0$, then this representation is irreducible. If $n \geq 0$, then there is precisely one subrepresentation. It is spanned by $f^k v$ with $k > n$. The quotient over this subrepresentation is our $V_n$. Since the whole representation is not decomposable into any direct sum of representations, this example shows that the Weyl theorem does not hold for infinite-dimensional representations, even of $\mathfrak{sl}(2,\mathbb{C})$.

13. The case of the finite characteristic is more complicated. The previous theorem holds for representations of dimensions less that the characteristic of the field. Consider the case of the field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. Show by examples that neither the previous classification nor the Weyl theorem, are valid for (finite-dimensional) representations of $\mathfrak{sl}(2,\mathbb{F}_p)$.

In conclusion, let us discuss arbitrary (maybe, reducible) finite-dimensional representations of $\mathfrak{sl}(2,\mathbb{C})$. Let $V$ be such a representation. Then, by Theorem,

$$V = V_{n_1} \oplus V_{n_2} \oplus \ldots \oplus V_{n_m} \tag{$**$}$$

where we can assume that $n_1 \geq n_2 \geq \ldots \geq n_m$. The operator $h\colon V \to V$ is diagonalizable, and all its eigenvalues are integers. Moreover, if $d_k(V)$ is the multiplicity of the eigenvalue $k$, then $d_k(V) = d_{-k}(V)$ and, for every $k$, $d_k(V) \geq d_{k+2}$ (this is true for $V$, because it is true for every $V_n$).

PROPOSITION. *The number of occurrences of $V_n$ in the decomposition $(**)$ equals $d_n(V) - d_{n+2}(V)$.*

*Proof.* Indeed, $d_n(V_n) - d_{n+2}(V_n) = 1$ and $d_k(V_n) - d_{k+2}(V_n) = 0$ for all non-negative $k$ not equal to $n$.

EXERCISES 14. Prove that if $n \geq m$, then

$$V_n \otimes V_m \cong V_{n+m} \oplus V_{n+m-2} \oplus V_{n+m-4} \oplus \ldots \oplus V_{n-m+2} \oplus V_{n-m}.$$

(The tensor product of representations of a Lie algebra has a natural structure of a representation of the same Lie algebra: $g(v \otimes w) = gv \otimes w + v \otimes gw$.)

15. Find the decomposition of the representations $S^2 V_n$ and $\Lambda^2 V_n$ into the sums of irreducible representations.

**3.6 Applications to roots of complex semisimple Lie algebras.** Let $\alpha, \beta$ be roots of a complex semisimple Lie algebra $\mathfrak{g}$, and $\alpha \neq 0$. Following section 3.1.4, consider the root series $\{\beta + k\alpha\}$, $p_{\beta,\alpha} = p \leq k \leq q = q_{\beta,\alpha}$ where $p \leq 0$ and $q \geq 0$ are integers. Let $\mathfrak{g}_{\beta,\alpha} = \bigoplus_{k=p}^{q} \mathfrak{g}^{\beta+k\alpha} \subset \mathfrak{g}$. This space is invariant with respect to the operators $\operatorname{ad} e_\alpha, \operatorname{ad} h_\alpha, \operatorname{ad} f_\alpha$ and, hence, is a space of representation of the Lie algebra $\mathfrak{sl}(2,\mathbb{C}) =$

span$(e, h, f)$ where $e$ and $f$ are multiples of $\operatorname{ad} e_\alpha$ and $\operatorname{ad} f_\alpha$, and $h = \dfrac{2}{\langle \alpha, \alpha \rangle} \operatorname{ad} h_\alpha$. The operator $\operatorname{ad} h_\alpha$ acts on $\mathfrak{g}^{\beta + k\alpha}$ as the multiplication by $(\beta + k\alpha)(h_\alpha) = \langle \beta + k\alpha, \alpha \rangle = \langle \beta, \alpha \rangle + k\langle \alpha, \alpha \rangle$, hence, $h \in \mathfrak{sl}(2, \mathbb{C})$ acts on $\mathfrak{g}^{\beta+k\alpha}$ as the multiplication by $\dfrac{2\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle} + 2k$. According to Section 3.5, the last number must be an integer. Moreover, the results of Section 3.5 provide a computation of this integer. Indeed, our representation of $\mathfrak{sl}(2, \mathbb{C})$ is $V_n$ where $n = q_{\beta,\alpha} - p_{\beta,\alpha}$. The space $\mathfrak{g}^{\beta+q\alpha}$ is annihilated by $\operatorname{ad} e_\alpha$ and $\mathfrak{g}^\beta = f^q \mathfrak{g}^{\beta+q\alpha}$. Hence

$$\frac{2\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle} = (q_{\beta,\alpha} - p_{\beta,\alpha}) - 2q_{\beta,\alpha} = -(p_{\beta,\alpha} + q_{\beta,\alpha}),$$

so

$$\langle \beta, \alpha \rangle = -\frac{p_{\beta,\alpha} + q_{\beta,\alpha}}{2} \langle \alpha, \alpha \rangle.$$

This is true for any two roots $\beta, \alpha$ of $\mathfrak{g}$ with $\alpha \neq 0$. Notice also that the two extreme eigenvalues of $h$ in $\mathfrak{g}_{\beta,\alpha}$ are $p_{\beta,\alpha} - q_{\beta,\alpha}$ and $q_{\beta,\alpha} - p_{\beta,\alpha}$.

Let us derive some immediate corollaries from our results.

Proposition 1. *Let $\beta, \alpha$ be roots, and $\alpha \neq 0$.*

(1) *If $\beta + \alpha$ is a root, then $[\mathfrak{g}^\beta, \mathfrak{g}^\alpha] = \mathfrak{g}^{\beta+\alpha}$.*

(2) *If neither of $\beta \pm \alpha$ is a root, then $\langle \beta, \alpha \rangle = 0$.*

*Proof:* (1) follows from the description of action of $e$ in $V_n$ given in Section 3.5; in the case (2), $p_{\beta,\alpha} = q_{\beta,\alpha} = 0$.

Proposition 2. *Let $\alpha$ be a non-zero root of $\mathfrak{g}$, and let $\lambda\alpha$ with some $\lambda \in \mathbb{C}$ be also a root. Then $\lambda = 0$ or $\pm 1$.*

*Proof.* Since $\dfrac{2\langle \lambda\alpha, \alpha \rangle}{\langle \alpha, \alpha \rangle} = 2\lambda \in \mathbb{Z}$, $\lambda$ must be an integer or a half-integer. The case of $\lambda \in \mathbb{Z}$ was settled by Theorem 4 of Section 3.4. Let $\lambda$ be a half-integer, that is, $2\lambda$ is odd. The set of eigenvalues of $h$ in $\mathfrak{g}_{\lambda\alpha,\alpha}$ has the form $\ldots, 2\lambda - 2, 2\lambda, 2\lambda + 2, \ldots$, and it is symmetric in 0. Thus, it contains 1, which means that $\dfrac{1}{2}\alpha$ is a root. But then $\alpha$ is not a root, a contradiction.

Now, let us do some additional computations. For an arbitrary $\xi \in \mathfrak{h}$,

$$\langle \xi, \xi \rangle = \sum_{\beta \in \text{roots}} \left( \dim \mathfrak{g}^\beta \right) \beta(\xi)^2 = \sum_{\beta \in \text{roots}} \beta(\xi)^2.$$

In particular, fro a root $\alpha$,

$$\langle \alpha, \alpha \rangle = \langle h_\alpha, h_\alpha \rangle = \sum_{\beta \in \text{roots}} \beta(h_\alpha)^2 = \sum_{\beta \in \text{roots}} \langle \beta, \alpha \rangle^2 = \sum_{\beta \in \text{roots}} \left( \frac{p_{\beta,\alpha} + q_{\beta,\alpha}}{2} \right)^2 \langle \alpha, \alpha \rangle^2,$$

thus

$$\langle \alpha, \alpha \rangle = \left( \sum_{\beta \in \text{roots}} \left( \frac{p_{\beta,\alpha} + q_{\beta,\alpha}}{2} \right)^2 \right)^{-1}.$$

We see that $\langle\alpha,\alpha\rangle \in \mathbb{Q}_{>0}$ for every non-zero root $\alpha$ and $\langle\beta,\alpha\rangle \in \mathbb{Q}$ for all roots $\beta,\alpha$.

Let $\mathfrak{h}_{\mathbb{R}}$ and $\mathfrak{h}_{\mathbb{Q}}$ be the sets of all real and all rational linear combinations of $h_\alpha$ for all roots $\alpha$.

PROPOSITION 3. (1) $\dim_{\mathbb{R}} \mathfrak{h}_{\mathbb{R}} = \dim_{\mathbb{Q}} \mathfrak{h}_{\mathbb{Q}} = \dim_{\mathbb{C}} \mathfrak{h}$.

(2) *The restriction of the Killing form to $\mathfrak{h}_{\mathbb{Q}}$ is rational and positive definite; hence the restriction of the Killing form to $\mathfrak{h}_{\mathbb{R}}$ is real and positive definite.*

*Proof.* Let $r = \dim \mathfrak{h}$, and let $\alpha_1,\dots,\alpha_r$ be such roots that $h_{\alpha_1},\dots,h_{\alpha_r}$ is a basis in $\mathfrak{h}$ (exists by Part (2) of Theorem 2 in Section 3.4). Then every root is a rational linear combination of $\alpha_1,\dots,\alpha_r$: indeed, if $\beta = \sum u_i\alpha_i$, then $(u_1,\dots,u_r)$ is a solution of rational system $\sum_j \langle\alpha_i,\alpha_j\rangle u_j = \langle\alpha_i,\beta\rangle,\ i = 1,\dots,r$, and hence all $u_i$ are rational. This implies (1). To prove (2), notice that if $\xi \in \mathfrak{h}_{\mathbb{Q}}$, then $\beta(\xi)$ is rational for every root $\beta$, and hence

$$\langle\xi,\xi\rangle = \sum_{\beta\in\text{roots}} \beta(\xi)^2$$

is rational and is positive, if $\xi \neq 0$. The statement for $\mathfrak{h}_{\mathbb{R}}$ follows (or can be proven in the same way).

The number $r = \dim \mathfrak{h}$ is called the *rank* of the semisimple Lie algebra $\mathfrak{g}$ and is denoted as $\operatorname{rank}\mathfrak{g}$. For a real semisimple Lie algebra $\mathfrak{g}$, we put $\operatorname{rank}\mathfrak{g} = \operatorname{rank}(\mathfrak{g}\otimes\mathbb{C})$. For example, $\operatorname{rank}\mathfrak{sl}(n,\mathbb{R}\text{ or }\mathbb{C}) = n-1$ and $\operatorname{rank}\mathfrak{su}(n) = n-1$.

EXERCISE 16. Prove that $\operatorname{rank}\mathfrak{so}(n) = \left[\dfrac{n}{2}\right]$.

**3.7. Geometry of roots.** Part (2) of Proposition 3 in Section 3.6 gives us a possibility to apply to roots of a semisimple Lie algebra the notions of the usual Euclidean geometry. Some things can be deduced from the formulas of the previous section immediately. Below we assume that some semisimple real or complex Lie algebra $\mathfrak{g}$ is fixed, and, speaking of *roots* we mean roots of this Lie algebra or (if it is real) of its complexification.

PROPOSITION. (1) *The angle between two non-collinear roots must be one of the following:* $30°, 45°, 60°, 90°, 120°, 135°, 150°$.

(2) *If the angle between two roots is* $60°$ *or* $120°$*, their lengths are the same; If the angle between two roots is* $45°$ *or* $135°$*, then the ratio of their length is* $\sqrt{2}$*; If the angle between two roots is* $30°$ *or* $150°$*, then the ratio of their length is* $\sqrt{3}$.

(3) *The length of a series of roots never exceeds* 4.

*Proof.* If $\beta,\alpha$ are two non-collinear roots, then

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

is the product of two integers or half-integers, and hence is some integer divided by 4. Since it is less than 1, it can be $0, \dfrac{1}{4}, \dfrac{1}{2}$, or $\dfrac{3}{4}$. This proves (1).

The square of the ratio of the lengths of $\alpha$ and $\beta$ is (if $\langle\beta,\alpha\rangle \neq 0$) the ratio of these two numbers. If the angle is $60°$ or $120°$, then these two numbers are both $\dfrac{1}{2}$ or $-\dfrac{1}{2}$, and the ratio is 1. If the angle is 45° or 135°, then the two numbers are $1, \frac{1}{2}$ or $-1, -\frac{1}{2}$, and the ratio is 2. Finally, if the angle is 30° or 150°, then the two numbers are $\frac{1}{2}, \frac{3}{2}$ or $-\frac{1}{2}, -\frac{3}{2}$, and the ratio is 3. This proves (2).

To prove (3), it is sufficient to notice that, according to the preceding computations, the absolute value of $\frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle} = -(p_{\beta,\alpha} + q_{\beta,\alpha})$ never exceeds 3.

If $r = 1$, then we have, according to Proposition 2 of Section 3.6, only one possibility for the roots: $\alpha, 0, -\alpha$. The pairs $(p_{\beta,\alpha}, q_{\beta,\alpha})$ for $\beta = \alpha, 0, -\alpha$ are, respectively, $(-2,0), (-1,1), (0,2)$ which makes

$$\langle\alpha,\alpha\rangle = \left((-1)^2 + 0 + 1^2)\right)^{-1} = \frac{1}{3}.$$

This the root system of the Lie algebra $\mathfrak{sl}(2)$.

If our Lie algebra is a product of two or more simple Lie algebras, then the root system is *reducible*: the Cartan algebra is the sum of Cartan algebras of these simple Lie algebras, and the root system is the union of roots system of several mutually orthogonal spaces. We will restrict our attention to *irreducible* root systems.

Suppose that $\alpha, \beta$ are two roots which form the angle of $120°, 135°, 150°$, and the length of $\beta$ equals the length of $\alpha$ times, respectively, $1, \sqrt{2}, \sqrt{3}$. Then the root series determined by $\beta$ and $\alpha$ is $\beta, \beta+\alpha, \ldots, \beta+q\alpha$ with $q$, respectively, being $1, 2, 3$. The configurations of roots arising are shown below.

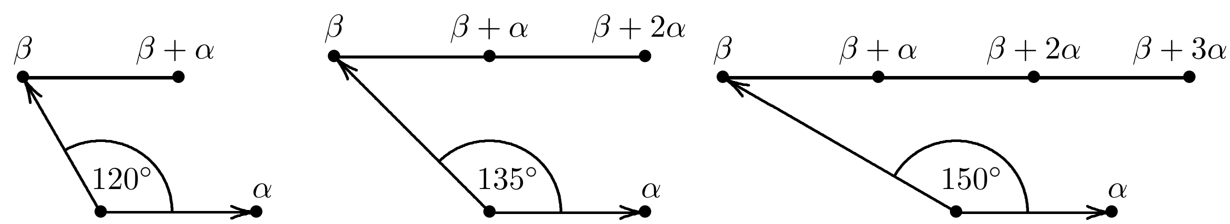



Fig. 1

Consider the case $r = 2$. Let us show that there are three possibilities for an irreducible root system, and they are the ones shown in Figure 2 below.

If the maximal obtuse angle between the roots is 120°, then the root system must contain the configuration in Figure 1, left, and also the roots $-\beta, -\beta-\alpha, -\alpha$. Any additional non-zero root would create an obtuse angle between roots exceeding 120°, so in this case we must have a root system shown in Figure 2, upper left.

If the maximal obtuse angle is 135°, then the root system contains the central configuration of Figure 1. Also it contains the four roots opposite to the roots of this configuration and again no more non-zero roots. This gives the upper right system of Figure 2.

Finally, in the case when this angle is 150°, the root system must contain the configuration in Figure 1 left, and also five roots opposite to the roots of this configuration. But this is not all. The roots $-\beta-2\alpha$ and $\beta+3\alpha$ also form an angle of 150°. This leads to the roots series $\beta+3\alpha, \alpha, -\beta-\alpha, -2\beta-3\alpha$ and, hence, a new root $-2\beta-3\alpha$. The opposite root $2\beta + 3\alpha$ also must appear, and we obtain a system of 12 non-zero roots shown in Figure 2, bottom. Again, there are no more non-zero roots.

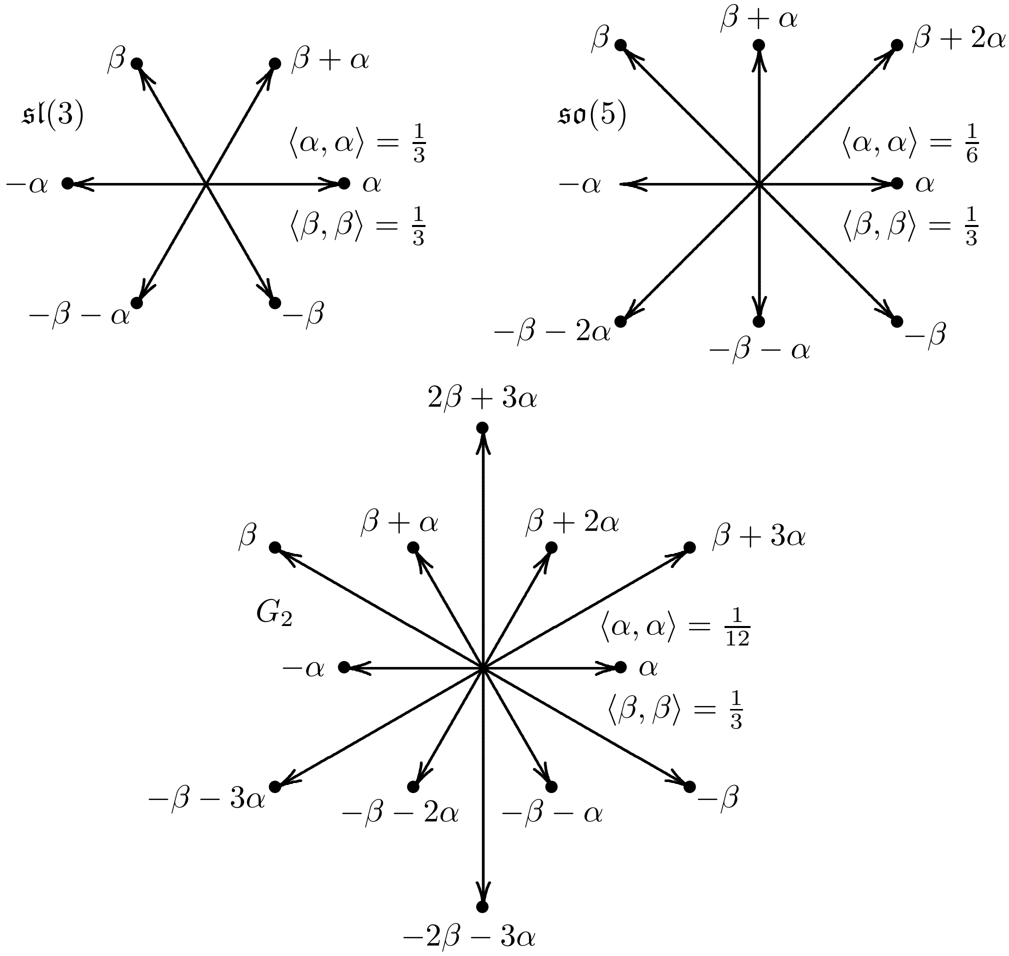



Fig. 2

In the upper left diagram, $(p_{\gamma,\alpha}, q_{\gamma,\alpha})$, for $\gamma = \alpha, \beta+\alpha$, and so on counterclockwise, is $(-2,0), (-1,0), (0,1), (0,2), (0,1)$, and $(-1,0)$, so

$$\langle\alpha,\alpha\rangle = \left(4\cdot\frac{1}{4} + 2\cdot 1\right)^{-1} = \frac{1}{3},$$

and $\langle\beta,\beta\rangle$ is the same. In the upper right diagram, these pairs are (for $\gamma$ ordered counterclockwise starting with $\alpha$) are $(-2,0), (-2,0), (-1,1), (0,2), (0,2), (0,2), (-1,1)$, and $(-2,0)$, so

$$\langle\alpha,\alpha\rangle = (6\cdot 1)^{-1} = \frac{1}{6},$$

and $\langle\beta,\beta\rangle$ is twice as large. Finally, in the bottom diagram these pairs are $(-2,0), (-3,0)$, $(-2,1), (0,0), (-1,2), (0,3), (0,2), (0,3), (-1,2), (0,0), (-2,1), (-3,0)$, so

$$\langle\alpha,\alpha\rangle = \left(4\cdot\frac{9}{4} + 2\cdot 1 + 4\cdot\frac{1}{4}\right)^{-1} = \frac{1}{12},$$

and $\langle\beta,\beta\rangle$ is three times as large.

The dimensions of the corresponding Lie algebras can be found as $\dim\mathfrak{h}=2$ plus the number of non-zero roots. For the three diagrams of Figure 2, these are $8, 10$, and 14. It is easyn to see that the first diagram corresponds to the Lie algebra $\mathfrak{sl}(3)$. The second diagram correspond to the 10-dimensional Lie algebra $\mathfrak{so}(5)$; the verification of this can be regarded as an extension of Exercise 16. The 14-dimensional Lie algebra of the last diagram is the simplest of the five so-called *exceptional* Lie algebras; its standard notation is $G_2$. We will consider exceptional Lie algebras at the appropriate moment later (see Section XX).

If the reader has an impression that it is possible, in a pure geometric way, to classify all the root systems of simple complex Lie algebras, we can say that not only this is right, but we are going to do this right away. Our classification will be actually, not a classification of root systems, but rather a classification of simple complex Lie algebras.

We finish this Section with a brief description of irreducible root systems in the 3-dimensional space. There are three of them.

EXERCISE 17. Give a description of the three irreducible root system in $\mathbb{R}^3$. To make it easier, let me give a geometric description of the final result.

The first of the three systems contains 12 non-zero roots, all of the same lengths. The convex hull of the set of roots is shown in Figure 3 below. It is the semiregular polyhedron with 14 faces, of which 6 are squares and 8 are equilateral triangles. The center of this polyhedron is the origin of the Euclidean space $\mathfrak{h}^*$, and the vertices correspond to the roots. The corresponding Lie algebra is $\mathfrak{sl}(4)\cong\mathfrak{so}(6)$.

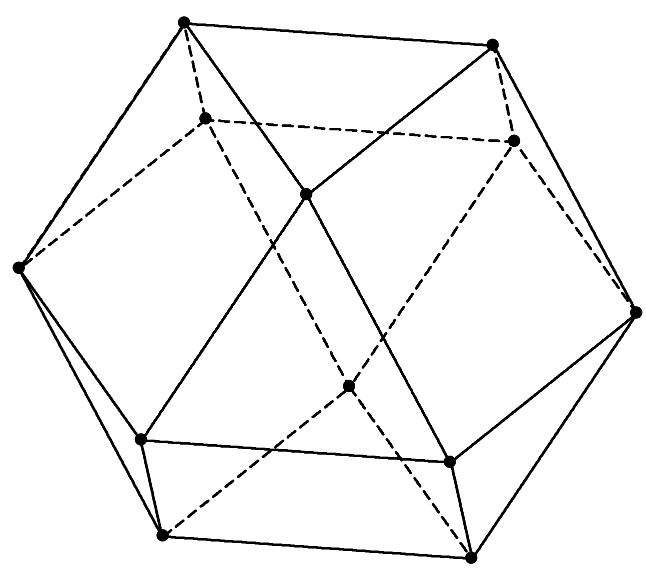


Fig. 3

The other two systems of roots have 18 non-zero roots each. The convex hull of one of them is the same as in the previous example, but the set of roots is different: besides the vertices of the polyhedron is also includes the centers of the square faces (Fig 4, keft). The other one is represented by a regular octahedron; the roots are the vertices and the midpoints of the edges (Fig. 4, right). The corresponding Lie algebras are, respectively,

$\mathfrak{so}(7)$ and $\mathfrak{sp}(3)$ (we will consider these Lie algebras later.

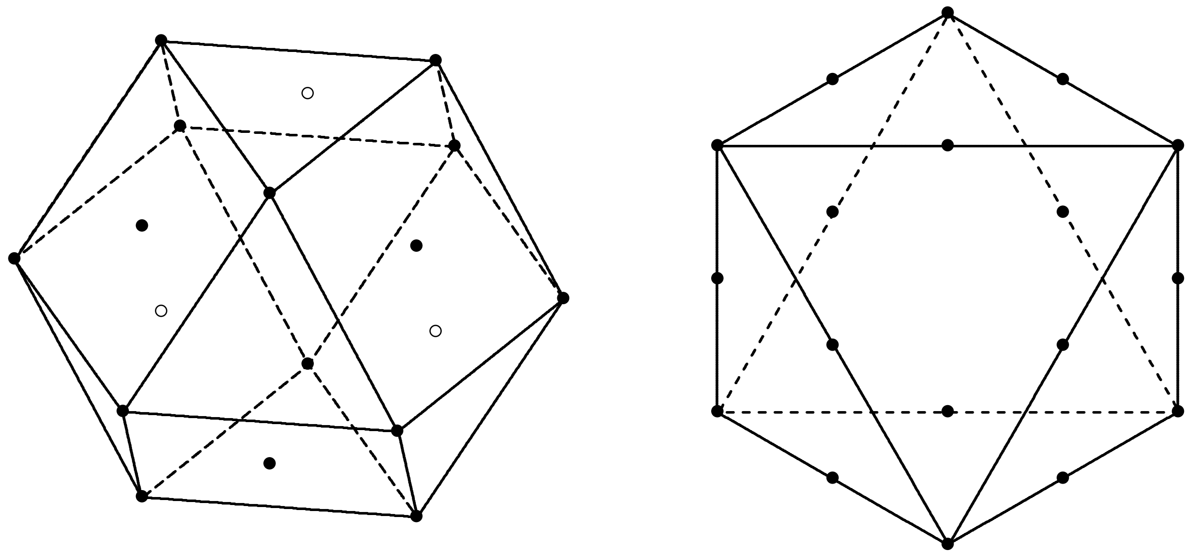


Fig. 4

**3.8. Dynkin diagrams.** In this section, we will give a solution of a problem which, actually, is equivalent to a description of root systems for arbitrary complex simple Lie algebras. We will consider not full root systems, but rather their subsystems consisting of *simple positive roots*.

**3.8.1. Simple positive roots.** Choose an arbitrary hyperplane in $\{F = 0\} \subset \mathfrak{h}^*_{\mathbb{R}}$ (where $F$ is a linear form) not passing through any non-zero root. Call a non-zero root $\alpha$ *positive*, if $\alpha(F) > 0$ and *negative* otherwise. Thus, every non-zero root is either positive or negative, and is $\alpha$ is positive, then $-\alpha$ is negative, and vice versa. A positive root is called a *simple positive root*, if it is not a sum of two other positive roots.

Proposition 1. *If $\alpha, \beta$ are two different simple positive roots, then $\langle \alpha, \beta \rangle \leq 0$.*

*Proof.* If $\langle \alpha, \beta \rangle > 0$, then $p_{\beta,\alpha} + q_{\beta,\alpha} < 0$ (see Section 3.6), and hence, $\beta - \alpha$ is a root, as well as $-(\beta - \alpha) = \alpha - \beta$. One of these roots must be positive, let it be $\beta - \alpha$. Then $\beta = (\beta - \alpha) + \alpha$ is a sum of two positive roots, so it is not simple.

Proposition 2. *Simple positive roots are linearly independent.*

*Proof.* Let $\alpha_1, \dots, \alpha_r$ be all simple positive roots, and let $\sum_i u_i \alpha_i = 0$ for some real $u_i$. Let $I_+ = \{i \mid u_i > 0\}$ and $I_- = \{j \mid u_j < 0\}$. Let $\gamma = \sum_{i \in I_+} u_i \alpha_i = \sum_{j \in I_-} (-u_j) \alpha_j$. Then $\langle \gamma, \gamma \rangle = \sum_{i \in I_+} \sum_{j \in I_-} (-u_i u_j) \langle \alpha_i, \alpha_j \rangle \leq 0$, thus actually $\langle \gamma, \gamma \rangle = 0$, so $\gamma = 0$. But a non-empty linear combination of positive roots with positive coefficients, as well as with negative coeficients, cannot be zero (because the value of $F$ is not zero). Thus the initial linear combination has all the coefficients zero.

Proposition 3. *Simple positive roots form a basis in $\mathfrak{h}^*_{\mathbb{R}}$. Moreover, every positive (negative) root is a linear combination of simple positive roots with all coefficients non-negative (non-positive). In particular, a linear cimbination of simple positive roots with coefficients of both signs is never a root.*

Proof. Let $\alpha$ be a positive root with the smallest value of $F(\alpha)$ which is not a positive linear combination of simple positive roots. Then it is not a sum of two positive roots, hence it is a simple positive root, which is a contradiction.

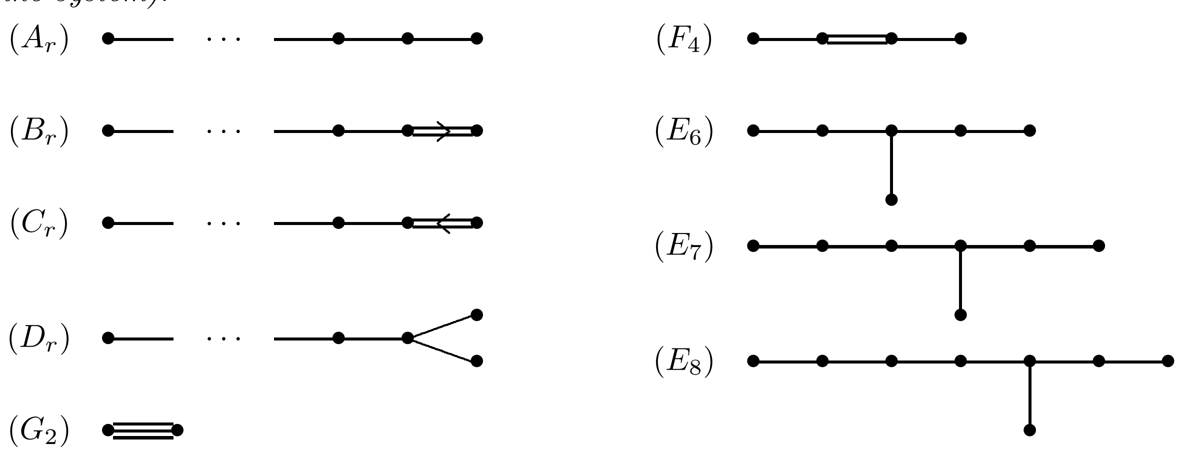
**3.8.2. Admissible system of vectors and their Dynkin diagrams.** Let $V$ be a Euclidean space, and let $\Pi = \{v_1, \dots, v_r\} \subset V$. We call $\Pi$ an *admissible system*, if the vectors $v_1, \dots, v_r$ are linearly independent and for every different $i, j$,

$$\frac{2\langle v_i, v_j\rangle}{\langle v_i, v_1\rangle} = 0, -1, -2, -3;$$

in other words, the angle $\angle(v_1, v_j)$ is equal to $90°, 120°, 135°, 150°$, and in the last three cases the ratio of lengths of $v_i$ and $v_j$ is equal to, respectively, $1, \sqrt{2}, \sqrt{3}$. Obviously, a subset of an admissible system is an admissible system. The system $\Pi$ is called *irreducible*, if there is no decomposition $\Pi = \Pi' \cup \Pi''$ with $v' \perp v''$ for all $v \in \Pi'$, $v'' \in \Pi''$.

The *Dynkin diagram* of an admissible system $v_1, \dots, v_r$ is a graph with $r$ vertices with a fixed bijection with $v_1, \dots, v_r$ in which $v_i$ and $v_j$ are joined by 1, 2, or 3 edges, if the angle $\angle(v_i, v_j)$ is, respectively, $120°$, $135°$, or $150°$ (in particular, if $v_i \perp v_j$, then the corresponding vertices are not joined by any edges. If two vertices are joined by two or three edges, $\bullet\!=\!\bullet$ or $\bullet\!\equiv\!\bullet$, then one of the two vectors involved is $\sqrt{2}$ or $\sqrt{3}$ times longer, than the other one; to indicate this on the diagram, we can use notations like $\bullet\!\Rightarrow\!\bullet$. Obviously, an admissible system is irreducible, if and only if its Dynkin diagram is connected,

THEOREM. *The Dynkin diagram of an irreducible admissible system is one of the following (the notations $A_r, \dots, E_8$ are commonly used notations for these Dynkin diagrams; in particular, the subscript denotes the number of vertices, that is, the number of vectors in the system)*:



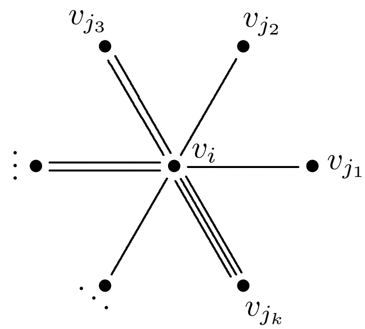

*Proof* will consist of five steps.

*Step One. Dynkin diagram is a tree* (possibly, with multiple edges). Put $w = \sum_i \frac{v_i}{\|v_i\|}$; this is a non-zero vector (since the vectors $v_i$ are linearly independent). Let $S$ be the set of pairs $(i, j)$ with $i < j$ and $\langle v_i, v_j\rangle \neq 0$ (that is the set of all connections in the Dynkin diagram). Notice that if $(i,j) \in S$, then $\cos \angle(v_i, v_j) \le -\frac{1}{2}$. We have:

$$0 < \langle w, w\rangle = \sum_i \frac{\langle v_i, v_i\rangle}{\|v_i\|^2} + \sum_{(i,j)\in S} \frac{2\langle v_i, v_j\rangle}{\|v_i\| \cdot \|v_j\|} = r + \sum_{(i,j)\in S} 2\cos\angle(v_i, v_j) \le r - \operatorname{card}(S).$$

Hence $\operatorname{card}(S) < r$, that is, $\operatorname{card}(S) \le r-1$, which means that actually $\operatorname{card}(S) = r-1$ and our graph is a tree.

*Step Two. No more, than three edges at every vertex* (edges are counted with their multiplicities). Consider a vertex in the Dynkin diagram with all its connections; the picture looks like this:



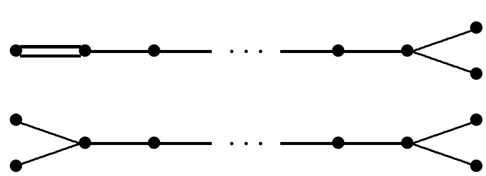

(we label every vertex as a vector from our system to which it is assigned). The vectors $v_{j_1}, \dots, v_{j_k}$ are all orthogonal to each other (no cycles in the Dynkin diagram!); on the contrary, the vector $v_i$ is not orthogonal to any of $v_{j_s}$, but is not their linear combination (the vectors of our system are linearly independent!). In the $(k+1)$-dimensional space $\operatorname{span}(v_i, v_{j_1}, \dots, v_{j_k})$ take a non-zero vector $w$ orthogonal to all $v_{j_s}$ (it is not vector from our system). Then

$$\cos^2\angle(v_i, w) + \sum_{s=1}^{k} \cos^2\angle(v_i, v_{j_s}) = 1.$$

But $\cos^2\angle(v_i, w) > 0$ and $\cos^2\angle(v_i, v_{j_s})$ is one quarter of the number of edges connecting $v_i$ with $v_{j_s}$. Thus the total number of edges emanating from $v_i$ cannot exceed 3.

In particular, we see that if our diagram contain a triple edge, then it cannot contain anything else; so it must be $G_2$.

*Step Three: excluding three more configurations.* Let us prove now that our Dynkin diagram cannot have a subdiagram of any of the following shapes:



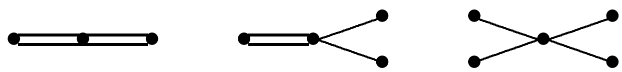
If the Dynkin diagram of some admissible system of vectors contains a subdiagram 

, then we will obtain another admissible system of vectors by replacing all the vectors $v_{i_s}$ by their sum, $v = v_{i_1} + \ldots + v_{i_k}$. Indeed, all the vectors $v_{i_s}$ have the same length; let $\langle v_{i_s}, v_{i_s}\rangle = a$. Then $\langle v_{i_s}, v_{i_{s+1}}\rangle = -\dfrac{a}{2}$ and

$$\langle v, v\rangle = \sum_{s=1}^{k}\langle v_{i_s}, v_{i_s}\rangle + 2\sum_{s=1}^{k-1}\langle v_{i_s}, v_{i_{s+1}}\rangle = ka - (k-1)a = a,$$

and if some $v_j$ from our system (not one of $v_{i_s}$) has a non-zero inner product with some $v_{i_s}$, then there can be only one such $v_{i_s}$ and therefore $\langle v_j, v\rangle = \langle v_j, v_{i_s}\rangle$. Thus we can collapse the subdiagram 

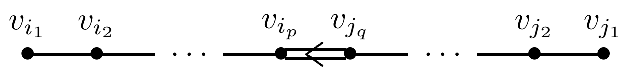
into one point to obtain a Dynkin diagram of another admissible system. But this will turn the three configurations shown above into the configurations



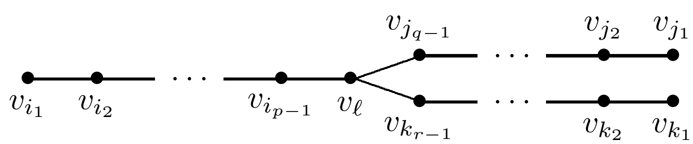
already prohibited in Step Two.

After Step Three, we have, for a Dynkin diagram of an irreducible admissibly system only the following possibilities: $G_2$, $A_r$ and two more:

$(X_{pq})$




$(Y_{pqr})$




Consider these last possibilities.

*Step Four: diagram $X_{pq}$.* Let $\langle v_{i_s}, v_{i_s}\rangle = a$. Then $\langle v_{i_s}, v_{i_{s+1}}\rangle = -\dfrac{a}{2}$, $\langle v_{j_s}, v_{j_s}\rangle = 2a$, $\langle v_{j_s}, v_{i_{s+1}}\rangle = -a$, $\langle v_{i_p}, v_{j_q} = -a$. Put $w = \sum_s s v_{i_s}$, $z = \sum_s s v_{j_s}$. Then

$$\langle w, w\rangle = a\left[(1^2 + 2^2 + \ldots + p^2) - (1\cdot 2 + 2\cdot 3 + \ldots + (p-1)p)\right] = \frac{p(p+1)}{2}a,$$

similarly, $\langle z, z\rangle = q(q+1)$, and, obviously, $\langle w, z\rangle = -pqa$. Apply the Cauchy-Schwarz inequality: $\dfrac{p(p+1)}{2}\cdot q(q+1) > p^2q^2$, hence $(p+1)(q+1) > 2pq$, $pq-p-q < 1$, $(p-1)(q-1) < 2$. Thus, either $p = 0$, or $q = 0$, or $p = q = 1$. This corresponds to the diagrams $B_r, C_r$, and $F_4$.

*Step Five: diagram $Y_{pqr}$.* All the vectors have the same length, let the inner square of each be $a$. Put $x = \sum_s s v_{i_s}$, $y = \sum_s s v_{j_s}$, $z = \sum_s s v_{k_s}$. Then, as in Step Four, $\langle x, x\rangle = \dfrac{p(p-1)}{2}\cdot a$, $\langle y, y\rangle = \dfrac{q(q-1)}{2}\cdot a$, $\langle z, z\rangle = \dfrac{r(r-1)}{2}\cdot a$. Also, $\langle x, w\rangle = -(p-1)\dfrac{a}{2}$, $\langle y, w\rangle = -(q-1)\dfrac{a}{2}$, $\langle z, w\rangle = -(r-1)\dfrac{a}{2}$. Hence, $\cos^2\angle(x, w) = \dfrac{(p-1)^2a^2}{4\cdot\frac{p(p-1)}{2}a^2} = \dfrac{p-1}{2p}$ and similarly $\cos^2\angle(y, w) = \dfrac{q-1}{2q}$, $\cos^2\angle(z, w) = \dfrac{r-1}{2r}$. Since the vectors $x, y, z$ are orthogonal to each other, the sum of this squares of cosines is less than 1 (it would have been 1, if $w$ were a linear combination of $x, y, z$). Thus,

$$\frac{p-1}{2p} + \frac{q-1}{2q} + \frac{r-1}{2r} < 1 \Longrightarrow \frac{1}{p} + \frac{1}{q} + \frac{1}{r} > 1.$$

If we assume that $1 < p \leq q \leq r$, this inequality holds, if $p = q = 2$ and $r$ is arbitrary, and if $p = 2, q = 3, r = 3, 4$, or 5. This corresponds to diagrams $D_{r+2}, E_6, E_7$, and $E_8$.

This completes the proof of Theorem.

**3.9. Classification of complex semisimple Lie algebras.** Our goal in this section is to show that the classification of admissible system of vectors given in Section 3.8 provides, actually a full isomorphism classification of complex semisimple Lie algebras. We already know that, for a complex semisimple Lie algebra, simple positive roots form an admissible system of vectors. According to Section 3.8, this assigns to a complex semisimple Lie algebra a Dynkin diagram, which is a disjoint union of diagrams of the types $A_r, \ldots, E_8$ as listed above. We want two show that if two complex semisimple Lie algebras have identical Dynkin diagrams, then they are isomorphic.

**3.9.1. A transition from Dynkin diagrams to root systems.** First, we show that is two complex semisimple Lie algebras, $\mathfrak{g}$ and $\mathfrak{g}'$ (we hope that the reader will not confuse this notation with the notation for the commutator subalgebra) with Cartan subalgebras $\mathfrak{h}$ and $\mathfrak{h}'$, have identical Dynkin diagrams that there exists an isometry $\mathfrak{h} \to \mathfrak{h}'$ which establishes a bijection between the sets of roots.

First, we reduce the general case to the irreducible case. We leave the following statement to the reader as an exercise.

Exercise 18. Prove that the system $\Pi$ of simple positive roots of a complex semisimple Lie algebra $\mathfrak{g}$ is reducible, $\Pi = \Pi' \cup \Pi''$, $\Pi' \perp \Pi''$, then $\mathfrak{g} = \mathfrak{g}' \oplus \mathfrak{g}''$ with the systems of simple positive roots of $\mathfrak{g}', \mathfrak{g}''$ being $\Pi', \Pi''$.

Next, the full system of roots is determined by the system of simple positive roots. Indeed, if two simple Lie algebras have the same system of simple positive roots (with respect to some isometry between the Cartan subalgebras preserving the positivety/negativity), then they have the same system of roots. Indeed, consider the first difference between the two system of roots, that is let $\alpha$ be a positive root of one of the two Lie algebras which is not a root of the other one, and $(\alpha)$ where $F$ is the linear form determining the positivity/negativity (see Section 3.8.1) be the smallest possible. But then $\alpha$ is not simple, $\alpha = \beta + \gamma$ where $\beta, \gamma$ are positive roots for the both algebras. But the numbers $p_{\beta,\gamma}, q_{\beta,\gamma}$ are determined by the inner products, so they must be the same for the two Lie algebras, and if $\beta + \gamma$ is a root of one of them, it must be also a root of the other one.

Finally, If simple complex Lie algebras $\mathfrak{g}, \mathfrak{g}'$ with Cartan subalgebras $\mathfrak{h}, \mathfrak{h}'$ have identical Dynkin diagrams, then there exists an isometry $\mathfrak{h} \to \mathfrak{h}'$ which establishes a bijection between roots. Indeed, a bijection between Dynkin diagrams is a bijection between simple positive roots, and since simple positive roots form a basis in the Cartan subalgebra, this bijection provides an isomorphism between the Cartan subalgebras. The only additional observation we need is that a Dynkin diagram determines inner products of simple positive roots only up to a (rational) factor, so our isomorphism could be rather a similarity than an isometry. But the formula $\langle\alpha,\beta\rangle = \sum\limits_{\gamma\in\{\text{roots}\}} \langle\gamma,\alpha\rangle\langle\gamma,\beta\rangle$ shows that root systems cannot be proportional without being identical.

**3.9.2. Isomorphism between the root systems implies isomorphism between Lie algebras.** The following is the main result of this Section.

THEOREM. *Let $\mathfrak{g}, \mathfrak{g}'$ be two semisimple complex Lie algebras, and let $\mathfrak{h}, \mathfrak{h}'$ be their Cartan subalgebras. If a linear map $\mathfrak{h} \to \mathfrak{h}'$ establishes a bijection between the roots of $\mathfrak{g}$ and $\mathfrak{g}'$, then it can be extended to a Lie algebra isomorphism $\mathfrak{g} \to \mathfrak{g}'$.*

Before proving Theorem, let us consider one semisimple complex Lie algebra $\mathfrak{g}$ with a Cartan subalgebra $\mathfrak{h}$. For every pair of opposite non-zero roots $\alpha, -\alpha$, fix $e_\alpha \in \mathfrak{g}^\alpha, e_{-\alpha} \in \mathfrak{g}^{(-\alpha)}$ such that $[e_\alpha, e_{-\alpha}] = -h_\alpha$. If $\alpha, \beta, \alpha+\beta$ are non-zero roots, then there exists a non-zero $N_{\alpha,\beta}$ such that $[e_\alpha, e_\beta] = N_{\alpha,\beta} e_{\alpha+\beta}$. If $\alpha$ and $\beta$ are non-zero roots, but $\alpha+\beta$ is not a non-zero root, then we put $N_{\alpha,\beta} = 0$.

LEMMA. *The numbers $N_{\alpha,\beta}$ satisfy the following relations.*

(1) $N_{\beta,\alpha} = -N_{\alpha,\beta}$.

(2) *If $\alpha, \beta, \gamma$ are non-zero roots and $\alpha+\beta+\gamma = 0$, then $N_{\alpha,\beta} = N_{\beta,\gamma} = N_{\gamma,\alpha}$.*

(3) *If $\alpha, \beta, \gamma, \delta$ are non-zero roots, the sum of no two of them is zero, and $\alpha+\beta+\gamma+\delta = 0$, then $N_{\alpha,\beta}N_{\gamma,\delta} + N_{\beta,\gamma}N_{\alpha,\delta} + N_{\gamma,\alpha}N_{\beta,\delta} = 0$.*

(4) $N_{\alpha,\beta}N_{-\alpha,-\beta} = \dfrac{1}{2}\langle\alpha,\alpha\rangle q(1-p)$ *where $p = p_{\alpha,\beta}, q = q_{\alpha,\beta}$* (see Section 3.6).

*Proof of Lemma.* (1) is obvious.

To prove (2), let us use the Jacobi identity:

$$\begin{aligned} 0 &= [e_\alpha, [e_\beta, e_\gamma]] + [e_\beta, [e_\gamma, e_\alpha]] + [e_\gamma, [e_\alpha, e_\beta]] \\ &= [e_\alpha, N_{\beta,\gamma} e_{-\alpha}] + [e_\beta, N_{\gamma,\alpha} e_{-\beta}] + [e_\gamma, N_{\alpha,\beta} e_{-\gamma}] \\ &= -(N_{\beta,\gamma} h_\alpha + N_{\gamma,\alpha} h_\beta + N_{\alpha,\beta} h_\gamma). \end{aligned}$$

But $h_\alpha + h_\beta + h_\gamma = 0$ and the vectors $h_\alpha, h_\beta, h_\gamma$ are not collinear. Hence, $N_{\alpha,\beta} = N_{\beta,\gamma} = N_{\gamma,\alpha}$.

To prove (3), we notice that

$$[e_\alpha, [e_\beta, e_\gamma]] = [e_\alpha, N_{\beta,\gamma} e_{\beta+\gamma}] = N_{\alpha,\beta+\gamma} N_{\beta,\gamma} e_{-\delta} = N_{\delta,\alpha} N_{\beta,\gamma} e_{-\delta}$$

(the last equality follows from (2), since $\alpha + (\beta+\gamma) + \delta = 0$). Thus $[e_\alpha, [e_\beta, e_\gamma]] = -N_{\beta,\gamma} N_{\alpha,\delta} e_{-\delta}$, and (3) follows from the Jacobi identity.

To prove (4), we use formulas for the action of $\mathfrak{sl}(2,\mathbb{C})$ in the space $\bigoplus_{k=p}^{q} \mathfrak{g}^{\beta+k\alpha}$. In the notation of Section 3.5, this space is $V_{q-p}$, $e_\beta \in \mathbb{C}v_q$, and $\operatorname{ad} e_{-\alpha} \circ \operatorname{ad} e_\alpha$ is $-\dfrac{1}{2}\langle\alpha,\alpha\rangle f \circ e$, and $f \circ e(v_q) = q(q-p-(q-1))v_q = q(1-p)v_q$. Thus

$$[e_{-\alpha}, [e_\alpha, e_\beta]] = -\frac{1}{2}\langle\alpha,\alpha\rangle q(1-p) e_\beta.$$

On the other hand,

$$[e_{-\alpha}, [e_\alpha, e_\beta]] = [e_{-\alpha}, N_{\alpha,\beta} e_{\beta+\alpha}] = N_{-\alpha,\alpha+\beta} N_{\alpha,\beta} e_\beta = N_{-\beta,-\alpha} N_{\alpha,\beta} e_\beta$$

(the last equality follow from (2), since $-\alpha+(\alpha+\beta)+(-\beta) = 0$). The equality (4) follows.

*Proof of Theorem.* As before, we assume that there are $e_\alpha \in \mathfrak{g}^\alpha$ such that $[e_\alpha, e_{-\alpha}] = -h_\alpha$ and $[e_\alpha, e_\beta] = N_{\alpha,\beta} e_{\alpha+\beta}$, if $\alpha, \beta, \alpha+\beta$ are non-zero roots of $\mathfrak{g}$. Also, there is a linear map $\varphi: \mathfrak{h} \to \mathfrak{h}'$ which establishes a bijection between the root systems $\Delta, \Delta'$ of $\mathfrak{g}$ and $\mathfrak{g}'$. For $(\varphi^*)^{-1}(\alpha)$ we use the notation $\alpha'$. We want to find, for every $\alpha$, an $e'_{\alpha'} \in \mathfrak{g}'^{(\alpha')} \subset \mathfrak{g}'$ such that

$$[e'_{\alpha'}, e'_{\beta'}] = N_{\alpha,\beta} e'_{\alpha'+\beta'}. \tag{$*$}$$

Then the map $\varphi$ will be extended by $e_\alpha \mapsto e'_{\alpha'}$ to a Lie algebra isomorphism $\mathfrak{g} \to \mathfrak{g}'$, so $\mathfrak{g}' \cong \mathfrak{g}$.

First remark that the map $\varphi$ must be an isometry. Indeed, since $\varphi$ is linear, the lengths of root series are the same for $\Delta$ and $\Delta'$, so $p_{\alpha,\beta} = p_{\alpha',\beta'}$ and $q_{\alpha,\beta} = q_{\alpha',\beta'}$. Hence

$$-\frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle} = p_{\alpha,\beta} + q_{\alpha,\beta} = p_{\alpha',\beta'} + q_{\alpha',\beta'} = -\frac{2\langle\beta',\alpha'\rangle}{\langle\alpha',\alpha'\rangle}$$

which shows that the inner products are proportional: $\langle\alpha,\beta\rangle = \lambda\langle\alpha',\beta'\rangle$ for all $\alpha, \beta$. But then

$$\langle\alpha,\beta\rangle = \langle h_\alpha, h_\beta\rangle = \sum_\gamma \gamma(h_\alpha)\gamma(h_\beta) = \sum_\gamma \langle\gamma,\alpha\rangle\langle\gamma,\beta\rangle = \sum_{\gamma'} \lambda\langle\gamma',\alpha'\rangle\lambda\langle\gamma',\beta'\rangle = \lambda^2\langle\alpha',\beta'\rangle,$$

so $\lambda^2 = \lambda$ and $\lambda = 1$.

Now let us make a choice of $e'_{\alpha'}$. We assume that the roots are ordered by the values of the function $F$ from Section 3.8.1 (we suppose that these values are different for different roots). For a positive root $\rho$, put $\Sigma_\rho = \{\alpha \in \Delta \mid -\rho < \alpha < \rho\}$. If $\sigma$ is the root next after $\rho$, then $\Sigma_\sigma = \Sigma_\rho \cup \{-\rho, \rho\}$.

*Induction hypothesis*: $(*)$ holds, if $\alpha, \beta, \alpha+\beta \in \Sigma_\rho$. We want to extend this to $\Sigma_\sigma$. If $\rho$ is a simple positive root, then we choose $e'_{\rho'} e'_{-\rho'}$ under the only condition $[e'_{\rho'}, e'_{-\rho'}] = -h_{\rho'}$. Otherwise, take any $\alpha, \beta \in \Sigma_\rho$ with $\alpha+\beta = \rho$ and define $e'_{\rho'}$ from the formula $(*)$ and $e'_{-\rho'}$ from the condition $[e'_{\rho'}, e'_{-\rho'}] = -h_{\rho'}$. Now, $e'_{\alpha'}$ is defined for all $\alpha \in \Sigma_\sigma$, and it remains to verify that if $\gamma, \delta, \gamma+\delta \in \Sigma_\sigma$, and $[e'_{\gamma'}, e'_{\delta'}] = N'_{\gamma',\delta'} e'_{\gamma'+\delta'}$, then $N'_{\gamma',\delta'} = N_{\gamma,\delta}$.

If $\gamma, \delta, \gamma+\delta \in \Sigma_\rho$, then $N'_{\gamma',\delta'} = N_{\gamma,\delta}$ by the induction hypothesis.

If $\gamma+\delta = \rho$ and $\{\gamma,\delta\} = \{\alpha,\beta\}$, then $N'_{\gamma',\delta'} = N_{\gamma,\delta}$ by the definition of $e'_{\rho'}$.

If $\gamma+\delta = \rho$, $\gamma, \delta \in \Sigma_\rho$ and $\{\gamma,\delta\} \neq \{\alpha,\beta\}$, then $\alpha+\beta+(-\gamma)+(-\delta) = 0$, pairwise sums are not zero, and, by Part (3) of Lemma,

$$N_{\alpha,\beta} N_{-\gamma,-\delta} + N_{\beta,-\gamma} N_{\alpha,-\delta} + N_{-\gamma,\alpha} N_{\beta,-\delta} = 0,$$
$$N_{\alpha,\beta} N'_{-\gamma',-\delta'} + N_{\beta,-\gamma} N_{\alpha,-\delta} + N_{-\gamma,\alpha} N_{\beta,-\delta} = 0$$

(no primes in the second equality, because each of $\beta-\gamma, \alpha-\delta, \alpha-\gamma, \beta-\delta$ is either not a root, or belongs to $\Sigma_\rho$, and we can use the statements in the previously considered cases); thus, $N'_{-\gamma',-\delta'} = N_{-\gamma,-\delta}$, and $N'_{\gamma',\delta'} = N_{\gamma,\delta}$ in virtue of Part (1) of Lemma.

The case of $\gamma+\delta=-\rho$ is similar.

Finally, of $\gamma, \delta, -(\gamma+\delta)$ no more than one is $\pm\rho$, their sum is zero, and we can finish the proof, using Part (2) of Lemma.

In particular, Theorem can be applied to the case $\mathfrak{g}'=\mathfrak{g}$ and $\varphi=-\operatorname{id}$. There arises a Lie algebra automorphism $\omega\colon \mathfrak{g}\to\mathfrak{g}$ with $\omega(e_\alpha)=\rho_\alpha e_{-\alpha}$. Put $\widetilde{e}_\alpha=\sqrt{\rho_\alpha}e_\alpha$. Then the numbers $N_{\alpha,\beta}$ arising from the condition $[\widetilde{e}_\alpha,\widetilde{e}_\beta]=N_{\alpha,\beta}\widetilde{e}_{\alpha+\beta}$ satisfy, in addition to the equalities from Lemma, the equality $N_{-\alpha,-\beta}=N_{\alpha,\beta}$. These vectors $\widetilde{e}_\alpha$ form, together with an arbitrary basis of $\mathfrak{h}$, a basis in $\mathfrak{g}$ which is called the *Weyl basis*. Remark that for a Weyl basis $\widetilde{e}_\alpha$ the numbers $N_{\alpha,\beta}$ are all real: by Part (4) of Lemma, $N^2_{\alpha,\beta}=N_{\alpha,\beta}N_{-\alpha,-\beta}=\dfrac{1}{2}\langle\alpha,\alpha\rangle q(1-p)\in\mathbb{Q}_{\geq 0}$.

**3.9.3. Classical and exceptional Lie algebras.** To finish the classification, we need to show that for each of the diagrams displayed in Theorem of Section 3.8.2, there exists a simple complex Lie algebra with this Dynkin diagrams. In this section, we will do a part of this work; the reader will find the rest elsewhere.

**3.9.3.1. The series $\mathbf{A_r}$.**



The corresponding Lie algebra is $\mathfrak{sl}(r+1)$, the Lie algebra of matrices of order $r+1$ with zero trace. For the Cartan algebra $\mathfrak{h}$, we choose the space of diagonal matrices which can be described as $\{(\lambda_1,\ldots,\lambda_{r+1})\in\mathbb{C}^{r+1}\mid \lambda_1+\ldots+\lambda_{r+1}=0\}$. The Killing form in $\mathfrak{h}$ is (proportional to) the restrictin to $\mathfrak{h}$ of the standard (bilinear) dot-product in $\mathbb{C}^{r+1}$, $\langle(\lambda_1,\ldots,\lambda_{r+1}),(\mu_1,\ldots,\mu_{r+1})\rangle=\lambda_1\mu_1+\ldots+\lambda_{r+1}\mu_{r+1}$. To see this it is better to take the Killing form not for the adjoint representation, but for the natural representation of $\mathfrak{h}$ in $\mathbb{C}^{r+1}$: indeed,

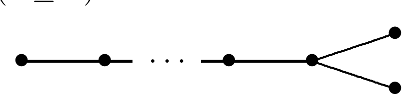
$$\operatorname{Tr}\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_{r+1}\end{bmatrix}\begin{bmatrix}\mu_1 & & \\ & \ddots & \\ & & \mu_{r+1}\end{bmatrix}=\lambda_1\mu_1+\ldots+\lambda_{r+1}\mu_{r+1}.$$

The root spaces are $\mathbb{C}E_{ij}$, $i\neq j$ (where $E_{ij}$ is a one-entry matrix), and the corresponding root is $(\lambda_1,\ldots,\lambda_{r+1})\mapsto\lambda_i-\lambda_j$. For positive roots we can take the roots $\lambda_i-\lambda_j$ with $i<j$ (to check this, we need to verify that for every non-zero root $\alpha$, precisely one of $\alpha,-\alpha$ is positive, and if the sum of two roots is a root, then it is a positive root). Then the simple positive roots are $\alpha_i=\lambda_i-\lambda_{i+1}$, $i=1,\ldots,r$. Obviously,

$$\langle\alpha_i.\alpha_j\rangle=\begin{cases}-1, & \text{if } j=i\pm 1,\\ 2, & \text{if } i=j,\\ 0 & \text{otherwise.}\end{cases}$$

Thus, all $\alpha_i$ have the same lengths, and the angle between $\alpha_i$ and $\alpha_j$ (with $i\neq j$) is $120°$, if $j=i\pm 1$ and is $90°$ otherwise. This corresponds to the Dynkin diagram above.

**3.9.3.2. The series $\mathbf{D_r}$ ($\mathbf{r\geq 4}$).**



The corresponding Lie algebra is $\mathfrak{o}(2r)$, the Lie algebra of skew-symmetric matrices of order $2r$. For the Cartan subalgebra, we can take the space of block diagonal skew-symmetric matrices with $r$ diagonal $2 \times 2$ blocks. The most convenient basis in $\mathfrak{h}$ consists of matrices $h_k = i(E_{2k-1,2k} - E_{2k,2k-1})$.

EXERCISE 19. (a) Prove that the Killing form with respect to this form is (proportional to) the standard dot product.

(b) Prove that the roots are $\pm\lambda_i \pm \lambda_j$, $1 \le i < j \le r$.

(c) Prove that for positive roots we can take the roots $\lambda_i \pm \lambda_j$, $1 \le i < j \le r$.

(d) Prove that the simple positive roots are $\lambda_1 - \lambda_2, \dots, \lambda_{r-1} - \lambda_r, \lambda_{r-1} + \lambda_r$. Check that the corresponding Dynkin diagram is the one shown above.

**3.9.3.3. The series $\mathbf{B_r}$.**

The corresponding Lie algebra is $\mathfrak{o}(2r+1)$, the Lie algebra of skew-symmetric matrices of order $2r+1$. For the Cartan subalgebra, we can take the same space as before in $\mathfrak{o}(2r) \subset \mathfrak{o}(2r+1)$, with the same basis (and inner product) as before.

EXERCISE 20. (a) Prove that the Killing form with respect to this form is proportional to the standard dot product.

(b) Prove that the roots are $\pm\lambda_i \pm \lambda_j (1 \le i < j \le r)$ and $\pm\lambda_i \,(1 \le i \le r)$.

(c) Prove that for positive roots we can take the roots $\lambda_i \pm \lambda_j, 1 \le i < j \le r$ and $\lambda_i \,(1 \le i \le r)$.

(d) Prove that the simple positive roots are $\lambda_1 - \lambda_2, \dots, \lambda_{r-1} - \lambda_r, \lambda_r$. Check that the corresponding Dynkin diagram is the one shown above.

**3.9.3.4. The series $\mathbf{C_r}$.**

The corresponding Lie algebra $\mathfrak{sp}(2r)$ is defined in the following way. Let $\langle\ ,\ \rangle$ be a non-degenerate skew-symmetric bilinear form in $\mathbb{C}^{2r}$. Such a form is, actually, unique up to an automorphism of $\mathbb{C}^{2r}$, but it will be convenient to us to have an explicit definition with respect to a basis. For a basis in $\mathbb{C}^{2r}$, we will use the notation $p_1, \dots, p_r, q_1, \dots, q_r$, and the form is defined by the equalities

$$\begin{aligned} \langle p_i, p_j \rangle &= 0, & \langle p_i, q_j \rangle &= \delta_{ij}, \\ \langle q_i, p_j \rangle &= -\delta_{ij}, & \langle q_i, q_j \rangle &= 0. \end{aligned}$$

A linear transformation $f\colon \mathbb{C}^{2r} \to \mathbb{C}^{2r}$ is called *canonical*, if $\langle f(x), f(y) \rangle = \langle x, y \rangle$ for all $x, y \in \mathbb{C}^{2r}$. Canonical transformations form a (complex) Lie group denoted as $Sp(2r, \mathbb{C})$ and the corresponding Lie algebra is $\mathfrak{sp}(2r)$. It consists of linear transformations $\varphi\colon \mathbb{C}^{2r} \to \mathbb{C}^{2r}$ such that $\langle \varphi(x), y \rangle + \langle x, \varphi(y) \rangle = 0$ for all $x, y \in \mathbb{C}^{2r}$ (it is seen from this description, if from nothing else, that $\mathfrak{sp}(2r)$ is a *complex* Lie algebra. More directly, $\mathfrak{sp}(2r)$ can be described as the Lie algebra of block matrices (with $r \times r$ blocks)

$$\begin{bmatrix} A & B \\ C & -A^t \end{bmatrix} \text{ with } B^t = B,\ C^t = C.$$

A Cartan subalgebra is an $r$-dimensional space with coordinates $\lambda_1, \dots, \lambda_r$; it consists of diagonal matrices with the diagonal entries $\lambda_1, \dots, \lambda_r, -\lambda_1, \dots, -\lambda_r$.

EXERCISE 21. (a) Prove that the Killing form with respect to this form is proportional to the standard dot product.

(b) Prove that the roots are $\lambda_i - \lambda_j$ $(i \neq j)$ and $\pm(\lambda_i + \lambda_j)$ $(i \leq j)$.

(c) Prove that for positive roots we can take the roots $\lambda_i - \lambda_j$, $1 \leq i < j \leq r$ and $\lambda_i + \lambda_j$ $(1 \leq i \leq j \leq r)$.

(d) Prove that the simple positive roots are $\lambda_1 - \lambda_2, \dots, \lambda_{r-1} - \lambda_r, 2\lambda_r$. Check that the corresponding Dynkin diagram is the one shown above.

**3.9.3.5. Exceptional Lie algebras.** There exist constructions of the simple complex Lie algebras with Dynkin diagrams $G_2, F_4, E_6, E_7$, and $E_8$ (and the corresponding Lie groups). These Lie algebras are called *exceptional*. Their dimensions (respectively) are 14, 52, 78, 133, and 248.

The Lie algebra of the type $G_2$ may be constructed as the complexification of the Lie algebra of derivations of the *octonion algebra*: a non-asociative 8-dimensional real division algebra. The constructions of the other exceptional Lie algebras (and corresponding Lie groups) are more complicated; they can be easily found in the literature (see, for example, "Exceptional Lie algebras" by N. Jacobson).

**4. Real Lie algebras.**

**4.1 Complexification and decomplexification.** If $\mathfrak{g}$ is a real Lie algebra, then the complexification $\mathbb{C}\mathfrak{g} = \mathfrak{g} \otimes \mathbb{C}$ is a complex Lie algebra. Certainly, $\mathfrak{g} \subset \mathbb{C}\mathfrak{g}$. If $\mathfrak{g}$ is (semi)simple, then $\mathbb{C}\mathfrak{g}$ is also (semi)simple, and the complexification of any Cartan subalgebra of $\mathfrak{g}$ is a Cartan subalgebra of $\mathbb{C}\mathfrak{g}$.

Moving backward, for a complex Lie algebra $\mathfrak{g}$, we may want to find a real subalgebra $\mathfrak{g}_{\mathbb{R}}$, of which $\mathfrak{g}$ is a complexification. This $\mathfrak{g}_{\mathbb{R}}$ is called a *real form* of $\mathfrak{g}$. It is known from linear algebra that if $W$ is a real vector subspace of a complex vector space $V$ such that $\dim_{\mathbb{R}} W = \dim_{\mathbb{C}} V$ and $V = \operatorname{span}_{\mathbb{C}} W$, then $V \cong_{\mathbb{R}} W \oplus iW$, and the *complex conjugation operator* $\sigma = \mathrm{id}_W \oplus (-\mathrm{id}_{iW}) \colon V \to V$ is an antilinear automophism such that $\sigma^2 = \mathrm{id}$ and $W = \operatorname{Fix}(\sigma)$. For Lie algebras, we just need to add the requirement that $\sigma$ is a Lie algebra automorphism: conjugation classes of real forms of a complex Lie algebra $\mathfrak{g}$ correspond bijectively to antilinear involutive (square = id) automorphisms of $\mathfrak{g}$.

There are well known examples of pairs $\mathfrak{g}, \mathbb{C}\mathfrak{g}$. The Lie algebra $\mathfrak{sl}(n, \mathbb{C})$ is the complexification of each of real subalgebras $\mathfrak{sl}(n, \mathbb{R})$ and $\mathfrak{su}(n)$: the corresponding antilinear involutions are $A \mapsto \overline{A}$, $A \mapsto -\overline{A}^t$. The Lie algebra $\mathfrak{o}(n, \mathbb{C})$ of complex skew-symmetric $n \times n$ matrices is the complexification of the Lie algebra $\mathfrak{o}(n, \mathbb{R})$ of real skew-symmetric matrices. Also, $\mathfrak{o}(n, \mathbb{C})$ is the complexification of the Lie algebra $\mathfrak{o}(p, q)$ where $p + q = n$; this is the Lie algebra of matrices $A$ satisfying the condition $\langle Av, w\rangle_{p,q} + \langle v, Aw\rangle_{p,q}$ for all $v, w \in \mathbb{R}^n$ where $\langle\ ,\ \rangle_{p,q}$ is a symmetric bilinear form of signature $(p, q)$. This Lie algebra has a convenient matrix description: it consists of matrices of the form

$$\begin{array}{c} p \\ q \end{array}\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right], \quad A^t = -A, B^t = C, D^t = -D.$$

Exercise 1. Prove that $\mathbb{C}\mathfrak{o}(p,q) = \mathfrak{o}(n,\mathbb{C})$.

**4.2. Compact Lie algebras.**

**4.2.1. Definition and the most important properties.** A real Lie algebra is called *compact*, if its Killing form is negative definite. (Question: why not positive definite? Answer: the Killing form cannot be positive definite: if it is positive definite, then it is non-degenerate, hence, the Lie algebra is semisimple; but according to Theorem 1 in Section 3.4, every root space corresponding to a non-zero root is orthogonal to itself, which is impossible for a positive definite form.) In particular, the Killing form of a compact Lie algebra is non-degenerate, so every compact Lie algebra is semisimple.

Theorem. (a) *If $G$ is a compact Lie group, then the Killing form of* $\operatorname{Lie} G$ *is non-positive. If $G$ is compact and semisimple (which means that* $\operatorname{Lie} G$ *is semisimple), then* $\operatorname{Lie} G$ *is compact.*

(b) *Every compact Lie algebra is a Lie algebra of a compact group.*

*Proof.* (a) If $G$ is compact, then $T_eG$ possesses an Ad-invariant Euclidean structure. Indeed, take an arbitrary Euclidean structure, $\langle\ ,\ \rangle'$, on $T_eG$, and take its average:

$$\langle v, w\rangle = \frac{\int_G \langle \mathrm{Ad}(g)v, \mathrm{Ad}(g)w\rangle' dg}{\operatorname{vol} G}$$

where $dg$ is an invariant measure on $G$ (its existence is a classical analytical fact; we will not prove it here). Then $\mathrm{Ad}(\mathfrak{g}) \subset O(\dim \mathfrak{g})$ and hence, with respect to an orthonormal basis in $\mathfrak{g} = T_eG$, the image of $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ consists of skew-symmetric matrices. But the trace of the square of a real skew-symmetric matrix, $\sum_{i,j} a_{ij}a_{ji} = -\sum_{ij} a_{ij}^2$ is never positive. If, in addition, he Lie algebra $\mathfrak{g}$ is semisimple, then the Killing form is non-degenerate, and, hence, negative definite.

(b) Let $G$ be a Lie group with $\operatorname{Lie} G = \mathfrak{g}$. As we have seen before, there arises a homomorphism $\mathrm{Ad}\colon G \to O(\dim \mathfrak{g})$, and the differential $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ has no kernel (because $\mathfrak{g}$ is semisimple). The kernel $\Gamma = \operatorname{Ker} \mathrm{Ad} \subset G$ is a central subgroup, and it is discrete (since $\operatorname{Ker} \mathrm{ad} = 0$). The image $\mathrm{Ad}(G/\Gamma)$ is a virtual subgroup of $O(\dim \mathfrak{g})$ with the Lie algebra $\mathfrak{g}$, and, since $\mathfrak{g}' = \mathfrak{g}$, it is a Lie subgroup (this can be easily deduced from the description of virtual subgroups given in Section 1.3.4.D). Thus, $G/\Gamma$ is compact, and $\operatorname{Lie}(G/\Gamma) = \mathfrak{g}$.

**4.2.2. The existence of a compact real form.** Theorem. *Every complex semisimple Lie algebra possesses a compact real form.*

*Proof.* Let $\mathfrak{g} = \mathfrak{h} \oplus [\bigoplus_\alpha \mathfrak{g}^\alpha]$ be a root decomposition of our Lie algebra. Let $\{h_i\,(1 \le i \le r), e_\alpha\,(\alpha \in \{\text{roots}\}\}$ be a basis in $\mathfrak{g}$ consisting of an arbitrary basis in $\mathfrak{h}_\mathbb{R} \subset \mathfrak{h}$ and the Weyl basis $\{e_\alpha \in \mathfrak{g}^\alpha\}$ (Sections 3.6 and 3.9.2). Define a map $\sigma\colon \mathfrak{g} \to \mathfrak{g}$ by the formula

$$\sigma\left(\sum\nolimits_i \mu_i h_i + \sum\nolimits_\alpha \lambda_\alpha e_\alpha\right) = -\sum\nolimits_i \overline{\mu}_i h_i + \sum\nolimits_\alpha \overline{\lambda}_\alpha e_{-\alpha}.$$

It is obvious that $\sigma$ is antilinear and that $\sigma^2 = \mathrm{id}$. Let us prove that $\sigma$ is a Lie algebra automorphism and that the restriction of the Killing form to $\operatorname{Fix} \sigma$ is negative definite.

To prove that $\sigma$ is a Lie algebra automorphism, it is sufficient to prove that

$$[\sigma(h_i), \sigma(h_j)] = \sigma[h_i, h_j],\ [\sigma(e_\alpha), \sigma(h_i)] = \sigma[e_\alpha, h_i], [\sigma(e_\alpha), \sigma(e_\beta)] = \sigma[e_\alpha, e_\beta].$$

The first is obvious: the both sides of the equality are equal to zero. Prove the second:

$$\begin{aligned}[\sigma(e_\alpha), \sigma(h_i)] &= [e_{-\alpha}, -h_i] = -\alpha(-h_i)e_{-\alpha} = \alpha(h_i)e_{-\alpha},\\ \sigma[e_\alpha, h_i] &= \sigma(\alpha(h_i)e_\alpha) = \overline{\alpha(h_i)}e_{-\alpha},\end{aligned}$$

and the two results match, since $\alpha(h_i)$ is real (see Section 3.6). Prove the third:

$$\begin{aligned}[\sigma(e_\alpha), \sigma(e_\beta)] &= [e_{-\alpha}, e_{-\beta}] = N_{-\alpha,-\beta}e_{-\alpha-\beta},\\ \sigma[e_\alpha, e_\beta] &= \sigma(N_{\alpha,\beta}e_{\alpha+\beta}) = \overline{N_{\alpha,\beta}}e_{-\alpha-\beta},\end{aligned}$$

and the two results match, since $N_{\alpha,\beta}$ is real and coincides with $N_{-\alpha,-\beta}$.

To prove that the restriction of the Killing form to $\operatorname{Fix}\sigma$ is negative definite, take $\xi = \sum_i \mu_i h_i + \sum_\alpha \lambda_\alpha e_\alpha$ and compute $\langle \xi, \sigma(\xi)\rangle$, that is

$$\left\langle \sum_i \mu_i h_i + \sum_\alpha \lambda_\alpha e_\alpha, -\sum_i \overline{\mu}_i h_i + \sum_\alpha \overline{\lambda}_\alpha e_{-\alpha}\right\rangle.$$

This expression is the sum of four inner products. The first is

$$\langle \sum_i \mu_i h_i, -\sum_i \overline{\mu}_i h_i\rangle = -\sum_\alpha \alpha\left(\sum_i \mu_i h_i\right)\alpha\left(\sum_i \overline{\mu}_i h_i\right) = -\sum_\alpha \left|\alpha\left(\sum_i \mu_i h_i\right)\right|^2 < 0$$

(we used the fact that $\alpha(h_i)$ is real, see Section 3.6). Next,

$$\left\langle \sum_\alpha \lambda_\alpha e_\alpha, \sum_\alpha \overline{\lambda}_\alpha e_{-\alpha}\right\rangle = \sum_\alpha \lambda_\alpha \overline{\lambda}_\alpha \langle e_\alpha, e_{-\alpha}\rangle = -\sum_\alpha |\lambda_\alpha|^2 < 0$$

(we used the facts that $\langle e_\alpha, e_{-\alpha}\rangle = -1$ and $\langle e_\alpha, e_\beta\rangle = 0$, if $\alpha + \beta \neq 0$, see Section 3.4). And the remaining two inner products are zeroes, since for $\alpha \neq 0$, $e_\alpha$ is orthogonal to $\mathfrak{h}$ (see again Theorem 1 of Section 3.4).

Remarks. (1) The last part of Proof shows, actually, that $x, y \mapsto \langle x, \sigma(y)\rangle$ is a negative definite Hermitian form on $\mathfrak{g}$.

(2) Obviously, $\sigma$ does not depend on the choice of a basis in $\mathfrak{h}_\mathbb{R}$.

(3) The most important result in Theorem is that the real form constructed is compact. Without the requirement of compactness, it is very easy to find a real (even a rational) form: take, for example, the space of all real (or rational) linear combinations of $h_\alpha$'s and $e_\alpha$'s.

Exercise 2. Prove that for $\mathfrak{sl}(n, \mathbb{C})$, the compact real form constructed in the last proof is $\mathfrak{su}(n)$, while the real and rational forms from Remark (3) above are $\mathfrak{sl}(n, \mathbb{R})$ and $\mathfrak{sl}(n, \mathbb{Q})$.

**4.3. Some further results.** In this section we will discuss some further results concerning real forms of complex semisimple Lie algebras. The reader can regard these facts as exercises. The proofs and a lot of additional information is contained in a book by A. Onishchik and E. Vinberg "*Lie groups and algebraic groups,*" Springer Verlag, 1990 (available on the Web).

PROPOSITION 1. *For a complex semisimple Lie algebra, its compact form is unique up to a conjugation.*

This means the following. If $\sigma, \sigma'\colon \mathfrak{g} \to \mathfrak{g}$ are two antilinear involutions of a semisimple complex Lie algebra $\mathfrak{g} = \operatorname{Lie} G$ with compact $\operatorname{Fix}\sigma, \operatorname{Fix}\sigma'$, then there exists a $g \in G$ such that the automorphism $\operatorname{Ad} g\colon \mathfrak{g} \to \mathfrak{g}$ takes $\sigma$ into $\sigma'$, that is, $\sigma' = \operatorname{Ad} g \circ \sigma \circ \operatorname{Ad} g^{-1}$. Consequently, $\operatorname{Ad} g$ takes $\operatorname{Fix}\sigma$ into $\operatorname{Fix}\sigma'$, in particular, the compact real Lie algebras $\operatorname{Fix}\sigma, \operatorname{Fix}\sigma'$ are isomorphic.

A full classification of real forms of (semi)simple complex Lie algebras is known. Before stating the results, we will describe the most important technical means for this classification.

Two real forms corresponding to antilinear involutions $\sigma, \tau\colon \mathfrak{g} \to \mathfrak{g}$ are *compatible*, if $\sigma \circ \tau = \tau \circ \sigma$ (equivalently, $\sigma(\operatorname{Fix}\tau) \subset \operatorname{Fix}\tau$, $\tau(\operatorname{Fix}\sigma) \subset \operatorname{Fix}\sigma$).

PROPOSITION 2. *Every linear form of a semisimple complex Lie algebra is compatible with a unique compact real form. More precisely, for every antilinear involution $\sigma\colon \mathfrak{g} \to \mathfrak{g}$ there exists a unique antilinear involution $\tau_\sigma\colon \mathfrak{g} \to \mathfrak{g}$ such that $\sigma \circ \tau_\sigma = \tau_\sigma \circ \sigma$ and $\operatorname{Fix}\tau_\sigma$ is compact.*

For a given $\sigma$, the composition $\sigma \circ \tau_\sigma$ is a complex Lie algebra involution.

PROPOSITION 3. *The correspondence $\sigma \mapsto \sigma \circ \tau_\sigma$ provides a bijection between the conjugacy classes of real forms of a semisimple complex Lie algebra and involutive automorphisms of this Lie algebra.*

Involutive automorhisms can be classified in terms of root systems. We will formulate, at least for classical Lie algebras, a final result. We will denote by $\mathbb{H}$ the quaternion algebra, by $E_n$ the identity matrix of order $n$ and also use the following notations:

$$S_m = \left[\begin{array}{c|c} 0 & -E_m \\ \hline E_m & 0 \end{array}\right]; \quad I_{p,q} = \left[\begin{array}{c|c} E_p & 0 \\ \hline 0 & -E_q \end{array}\right]; \quad K_{p,q} = \left[\begin{array}{c|c} I_{p,q} & 0 \\ \hline 0 & I_{p,q} \end{array}\right].$$

PROPOSITION 4. *The following involutive automorphisms $\theta\colon \mathfrak{g} \to \mathfrak{g}$ form full systems of representatives of conjugacy classes of involutive automorphisms of classical simple complex Lie algebras.*

(1) $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C}), n \geq 2$:
- (a) $\theta(X) = -X^t$;
- (b) $\theta(x) = -\operatorname{Ad} S_m(X^t)$ $(n = 2m)$;
- (c) $\theta = \operatorname{Ad} I_{p,n-p}$ $(p = 0, 1, \dots, [n/2])$.

(2) $\mathfrak{g} = \mathfrak{so}(n, \mathbb{C}), n = 3$ or $n \geq 5$:
- (a) $\theta = \operatorname{Ad} I_{p,n-p}$ $(p = 0, 1, \dots, [n/2])$;
- (b) $\theta(x) = \operatorname{Ad} S_m$ $(n = 2m)$.

(3) $\mathfrak{g} = \mathfrak{sp}(n, \mathbb{C}), n = 2m \geq 2$:
- (a) $\theta(x) = \operatorname{Ad} S_m$ $(n = 2m)$.
- (b) $\theta = \operatorname{Ad} K_{p,m-p}$ $(p = 0, 1, \dots, [m/2])$.

The corresponding real forms (or, at least, the notations for them) are listed in the following statement.

PROPOSITION 5. *Every real form of the simple complex Lie algebra $\mathfrak{g}$ is isomorphic to precisely one of the following.*

(1) $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C}), n \geq 2$:
- (a) $\mathfrak{sl}(n, \mathbb{R})$;
- (b) $\mathfrak{sl}(m, \mathbb{H})$ $(n = 2m)$;
- (c) $\mathfrak{su}(p, n-p)$ $(p = 0, 1, \dots, [n/2])$.

(2) $\mathfrak{g} = \mathfrak{so}(n, \mathbb{C}), n = 3$ or $n \geq 5$:
- (a) $\mathfrak{so}(p, n-p)$ $(p = 0, 1, \dots, [n/2])$;
- (b) $\mathfrak{u}(m, \mathbb{H})$ $(n = 2m)$.

(3) $\mathfrak{g} = \mathfrak{sp}(n, \mathbb{C}), n = 2m \geq 2$:
- (a) $\mathfrak{sp}(n, \mathbb{R})$ $(n = 2m)$.
- (b) $\mathfrak{sp}(p, m-p)$ $(p = 0, 1, \dots, [m/2])$.

Real forms of exceptional complex Lie algebras are also all known. Here, we restrict ourselves to mentioning that the numbers of non-isomorphic real forms of the Lie algebras of the types $G_2, F_4, E_6, E_7, E_8$ are, respectively, 1, 2, 4, 4, 2.

**5. Survey of the representation theory.** In this section, we will briefly discuss some classical results regarding finite-dimensional representations of finite-dimensional semisimple complex Lie algebra. We may occasionally mention the infinite-dimensional case, but, if the opposite is not explicitly stated, $\mathfrak{g}$ denotes a finite-dimensional semisimple complex Lie algebra, $\mathfrak{h}$ denotes a Cartan subalgebra with a specified sets of positive and negative roots: this sets will be denoted as $\Delta^+$ and $\Delta^-$. Then $\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+$ where $\mathfrak{n}_\pm = \oplus_{\alpha \in \Delta^\pm} \mathfrak{g}^\alpha$. These $\mathfrak{n}_-, \mathfrak{h}, \mathfrak{n}_+$ are subalgebras of $\mathfrak{g}$, $\mathfrak{n}_-$ and $\mathfrak{n}_+$ are nilpotent, $\mathfrak{h}$ is commutative, and $[\mathfrak{h}, \mathfrak{n}_\pm] \subset \mathfrak{n}_\pm$. If the opposite is not stated, we assume the representations considered finite-dimensional.

**5.1. Weights and highest weights.**

**5.1.1. Definitions and main results.** Let $V$ be a representation of $\mathfrak{g}$, and let $\lambda\colon \mathfrak{h} \to \mathbb{C}$ be a linear map. A non-zero vector $v \in V$ is called a weight vector of type $\lambda$, or a vector of weight $\lambda$, if $hv = \lambda(h)v$ for every $h \in \mathfrak{h}$ (this definition should replace a definition of weights considered in Sections 2 and 3). The set $V^\lambda$ of all vectors of a given weight $\lambda$ (plus the zero vector) is called the *weight space*. We will refer to $\lambda$'s with $V^\lambda \neq 0$ as to *weights* of the representation. A weight vector $v$ is called a *highest* weight vector, if, in addition of the assumption above, $gv = 0$ for every $g \in \mathfrak{n}_+$.

PROPOSITION 1. *If* $\dim V < \infty$*, then* $V$ *contains a highest weight vector.*

*Proof.* First, we construct a weight vector. Let $h_1, h_2, \dots h_r$ be a basis in $\mathfrak{h}$. The operator $h_1\colon V \to V$ has at least one eigenvector; let the corresponding eigenvalue be $\lambda_1$. Let $V_1 = \{v \in V \mid h_1 v = \lambda_1 v\}$; this is a non-zero subspace of $V$. Since $[h_1, h_2] = 0$, the space $V_1$ is $h_2$-invariant: for $v \in V_1$, $h_1(h_2 v) = h_2(h_1 v) = h_2(\lambda_1 v) = \lambda_1 h_2 v$, so $h_2 v \in V_1$. The operator $h_2\colon V_1 \to V_1$ has at least one eigenvector; let the corresponding eigenvalue be $\lambda_2$ and let $V_2 = \{v \in V_1 \mid h_2 v = \lambda_2 v\}$; this is a non-zero subspace of $V_2$. Proceeding in the same way with $h_3, \dots, h_r$, we obtain a chain of subspaces $V \supset V_1 \supset V_2 \supset \dots \supset V_r \neq 0$ and numbers $\lambda_1, \lambda_2, \dots, \lambda_r$ such that $h_i v = \lambda_i v$ for every $v \in V_r$ and every $i$. Then every non-zero element of $V_r$ is a vector of weight $\lambda$ where $\lambda(h_i) = \lambda_i$.

Now let us pass to highest weight vectors. Choose an $h_0 \in \mathfrak{h}_{\mathbb{R}} \subset \mathfrak{h}$ such that $\alpha(h_0)$ is a positive real number for every $\alpha \in \Delta^+$. Choose a non-zero $e_\alpha$ in every $\mathfrak{g}^\alpha$ (we can use a Weyl basis, but it is not necessary at the moment). Let $v \in V$ be a vector of weight $\lambda$. If $e_\alpha v = 0$ for every $\alpha \in \Delta^+$, then $v$ is a highest weight vector, and we have nothing to do. Let $e_\alpha v \neq 0$ for some $\alpha \in \Delta^+$. Then the vector $e_\alpha v$ is a weight vector of type $\lambda + \alpha$. Indeed, $h e_\alpha v = e_\alpha h v + [h, e_\alpha] v = e_\alpha(\lambda(h) v) + \alpha(h) e_\alpha v = (\lambda(h) + \alpha(h)) e_\alpha v$. Again, if $e_\beta(e_\alpha v) = 0$ for every $\beta \in \Delta^+$, or there arises a weight vector $e_\beta e_\alpha v$ of type $\lambda + \alpha + \beta$. We iterate this process with two possible results. Either we arrive, at some step, at a highest weight vector $e_\omega \dots e_\beta e_\alpha v$, or we obtain an infinite sequence $v, e_\alpha v, e_\beta e_\alpha v, e_\gamma e_\beta e_\alpha v, \dots$ of weight vectors of types $\lambda, \lambda + \alpha, \lambda + \alpha + \beta, \lambda + \alpha + \beta + \gamma, \dots$. In particular, vectors in our sequence are eigenvectors of $h_0$ with eigenvalues $\lambda(h), \lambda(h) + \alpha(h), \lambda(h) + \alpha(h) + \beta(h), \lambda(h) + \alpha(h) + \beta(h) + \gamma(h), \dots$, and all these eigenvalues are different. This contradicts to the finite-dimensionality of $V$. Thus, our procedure yields a highest weight vector.

PROPOSITION 2. *If the representation $V$ is irreducible, then a highest weight vector in $V$ is unique up to a proportionality.*

*Proof.* Let $v \in V$ be a highest weight vector. Consider all vectors $e_{-\alpha_k} \dots e_{-\alpha_2} e_{-\alpha_1} v$ for all sequences $\alpha_1, \dots, \alpha_k$ of positive roots and for all $k \geq 0$. Let $W \subset V$ be a subspace spanned by all these vectors. We state that *$W$ is a subrepresentation of $V$.*

Indeed, $W$ is invariant with respect to $\mathfrak{h}$ (since $e_{-\alpha_k} \dots e_{-\alpha_2} e_{-\alpha_1} v$ is a weight vector of type $\lambda - \alpha_1 - \dots - \alpha_k$. In is obviously invariant with respect to $\mathfrak{n}_-$. Now let us prove that it is invariant with respect to $\mathfrak{n}_+$. We want to prove that for all $\alpha, \alpha_1, \dots, \alpha_k \in \Delta_+$, $w = e_\alpha(e_{-\alpha_k} \dots e_{-\alpha_1} v) \in W$. We apply induction with respect to $k$: for $k = 0, w = e_\alpha v = 0$, since $v$ is a highest weight vector. Let $[e_\alpha, e_{-\alpha_k}] = a e_{\alpha - \alpha_k}$ (if $\alpha - \alpha_k$ is not a root, then $a = 0$). Then

$$w = e_\alpha(e_{-\alpha_k} \dots e_{-\alpha_1} v) = e_{-\alpha_k} e_\alpha e_{-\alpha_{k-1}} \dots e_{-\alpha_1} v + a e_{\alpha - \alpha_k} e_{-\alpha_{k-1}} \dots e_{-\alpha_1} v,$$

and both summands belong to $W$ by the induction hypothesis and invariance of $W$ with respect to $\mathfrak{n}_-$ and $\mathfrak{h}$ (whether $\alpha - \alpha_k$ is a positive root, a negative root, zero, or not a root at all).

Since $V$ is irreducible, we conclude that $W = V$. This shows that $V$ is a direct sum of weight spaces $V^\mu$, and every $\mu$ has the form $\lambda - \alpha_1 - \dots - \alpha_k$ where $\alpha_i$ is a positive root (we use the fact that the sum of weight spaces of different weights has to be direct). The space $V^\lambda$ is one-dimensional and is generated by $v$; for all other weight spaces $V^\mu$, $\mu(h_0) < \lambda(h_0)$.

Let now $w \in V$ be another highest weight vector. It belongs to a weight space $V^\mu$. If $\mu = \lambda$, then $w$ is proportional to $v$. Otherwise, $\mu(h_0) < \lambda(h_0)$. Apply to $w$ the construction from the beginning of this proof. We will obtain a subrepresentation of $V$ which is spanned by weight vectors of types $\nu$ with $\nu(h_0) \leq \mu(h_0)$. Hence, this subrepresentation does not contain $v$, and it cannot exist, because $V$ is irreducible. This completes our proof.

REMARK. Proposition

PROPOSITION 3. *Every finite-dimensional representation of a finite-dimensional semi-simple complex Lie algebra is decomposed into a direct sum of weight spaces.*

*Proof.* For irreducible representations, this was stated and proved in the previous proof. In general case, it is true, because every finite-dimensional representation is a direct sum of irreducible representations (the Weyl theorem, Section 3.3.1).

Notice that Proposition 2 provides an algorithm for decomposing of a given finite-dimensional representation into the sum of irreducible representation: it is sufficient to find all highest weight vectors and then choose a maximal family of linear independent highest vectors. The vectors of this family will generate the irreducible components. (This decomposition is unique if and only if every pair of non-proportional highest weight vectors have different types.)

**5.1.2. Examples**. Let $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C})$ $(n \geq 2)$, and let $V$ be (the space of) the natural $n$-dimensional representation of $\mathfrak{g}$. Fix a basis $e_1, \ldots, e_n$ in $V$. Then $\mathfrak{n}_-, \mathfrak{h}, \mathfrak{n}_+$ consist, respectively, of strictly low triangular, diagonal, and upper triangular matrices. The algebra $\mathfrak{n}_+$ is generated by $E_i = E_{i,i+1},\ i = 1, \ldots, n-1$; the algebra $\mathfrak{n}_-$ is generated by $F_i = E_{i+1,i},\ i = 1, \ldots, n-1$; the algebra $\mathfrak{h}$ has a basis $H_i = E_{i,i} - E_{i+1,i+1},\ i = 1, \ldots, n-1$. Obviously, $E_i e_k = \delta_{i,k-1} e_{k-1}$, $F_i e_k = \delta_{ik} e_{k+1}$, $H_i e_k = (\delta_{ik} - \delta_{i,k-1}) e_k$.

Consider the representation $V \otimes V$ (as usual, $A(v \otimes w) = Av \otimes w + v \otimes Aw$). The vectors $e_k \otimes e_\ell$ are all weight vectors (the type $\lambda_{k\ell}$ of $e_k \otimes e_\ell$ is $\lambda_{k\ell}(H_i) = \delta_{ik} + \delta_{i\ell} - \delta_{i,k-1} - \delta_{i,\ell-1}$).

Exercises. 1. Prove that $e_1 \otimes e_1$ and $e_1 \otimes e_2 - e_2 \otimes e_1$ are highest weight vectors in $V \otimes V$, and every highest weight vector in $V \otimes V$ is proportional to one of these two.

2. Prove that the irreducible components of $V \otimes V$ corresponding to the two highest vectors of Exercise 1 are $S^2 V$ and $\Lambda^2 V$.

Consider now tensor products of more than two copies of $V$.

Exercises 3. Prove that all the highest weight vectors in $V \otimes V \otimes V$ are $a e_1 \otimes e_1 \otimes e_1$ $(a \neq 0)$; $b_1 e_2 \otimes e_1 \otimes e_1 + b_2 e_1 \otimes e_2 \otimes e_1 + b_3 e_1 \otimes e_1 \otimes e_2$ $(b_1 + b_2 + b_3 = 0, (b_1, b_2, b_3) \neq (0,0,0)$; $c(e_1 \otimes e_2 \otimes e_3 - e_1 \otimes e_3 \otimes e_2 - e_2 \otimes e_1 \otimes e_3 + e_2 \otimes e_3 \otimes e_1 + e_3 \otimes e_1 \otimes e_2 - e_3 \otimes e_2 \otimes e_1)$ $(c \neq 0)$. The first and the last vectors generate, respectively, the subrepresentations $S^3 V \subset V \otimes V \otimes V$ and $\Lambda^3 V \subset V \otimes V \otimes V$; vectors from the second family generate isomorphic representations of $\mathfrak{g}$. If we choose two linearly independent vectors from this family, for example, $e_2 \otimes e_1 \otimes e_1 - e_1 \otimes e_2 \otimes e_1$ and $e_1 \otimes e_2 \otimes e_1 - e_1 \otimes e_1 \otimes e_2$, then we obtain two isomorphic subrepresentations of $V \otimes V \otimes V$, and $V \otimes V \otimes V$ becomes the sum of four irreducible representations: these two and also $S^3 V$ and $\Lambda^3 V$. The dimensions of the four components are:

$$\frac{n(n+1)(n+2)}{6}, \ \frac{(n-1)n(n+1)}{3}, \ \frac{(n-1)n(n+1)}{3}, \ \frac{(n-2)(n-1)n}{6}.$$

(Certainly, we assume that $n \geq 3$; if $n = 2$, then the last component disappears.)

4. In general, $\underbrace{V \otimes \ldots \otimes V}_{k}$ is a representation of the symmetric group$S_k$. It is naturally decomposed into the sum of isotypic components corresponding to the types of irreducible representations of $S_k$. Prove that every isotypic component is also a subrepresentation of $\mathfrak{g}$ and is a sum of isomorphic irreducible representations of $\mathfrak{g}$ in the number equal to the dimension of the corresponding irreducible representation of $S_k$.

**5.2. Representations with a given highest weight.** The main goal of this Section is to construct, for every linear function $\lambda\colon \mathfrak{h} \to \mathbb{C}$ an irreducible, possibly infinite-dimensional, representation of $\mathfrak{g}$ with a highest weight $\lambda$ and to show that this representation is unique up to isomorphism. We will also discuss the conditions which ensure the finiteness of dimension, and will briefly discuss a generalization of the whole theory to a class of infinite-dimensional Lie algebras.

**5.2.1. Algebraic digression: universal enveloping algebras and induced representations.** Let $A$ be a complex vector space(actually, the ground field may be arbitrary). Consider the *tensor albebra*

$$T(A) = \bigoplus_{k=0}^{\infty} \left(\otimes^k A\right),$$

that is,

$$T(A) = \mathbb{C} \oplus A \oplus (A \otimes A) \oplus (A \otimes A \otimes A) \oplus \ldots;$$

the multiplication is given by the formula $(a_1' \otimes \ldots \otimes a_p') \cdot (a_1'' \otimes \ldots \otimes a_q'') = a_1' \otimes \ldots \otimes a_p' \otimes a_1'' \otimes \ldots \otimes a_q''$. This is a non-commutative (when $A \neq 0$) graded associative unitary algebra (with the identity element $1 \in \mathbb{C} \subset T(A)$).

Let now $A$ be a complex Lie algebra $\mathfrak{g}$. In the tensor algebra $T(\mathfrak{g})$, consider the two-sided ideal $I(\mathfrak{g})$ generated by all elements of the form

$$[\xi, \eta] - (\xi \otimes \eta - \eta \otimes \xi) \in \mathfrak{g} \oplus (\mathfrak{g} \otimes \mathfrak{g}) \subset T(\mathfrak{g}),\ \xi, \eta \in \mathfrak{g}.$$

The associative unitary algebra $U(\mathfrak{g}) = T(\mathfrak{g})/I(\mathfrak{g})$ is called the *universal enveloping algebra* of $\mathfrak{g}$. This algebra is not graded any more, but is filtered: $F_pU(\mathfrak{g})$ is the image in $U(\mathfrak{g})$ of $\bigoplus_{k=0}^{p} \left(\otimes^k \mathfrak{g}\right)$. We will denote the image of $\xi_1 \otimes \ldots \otimes \xi_k \in T(\mathfrak{g})$ in $U(\mathfrak{g})$ as $\xi_1 \ldots \xi_k$.

The importance of the universal enveloping algebra $U(\mathfrak{g})$ stems mainly from the tautological fact that a representation of a Lie algebra $\mathfrak{g}$ is the same as a representation of $U(\mathfrak{g})$, or, in a more traditional algebraic language, a $U(\mathfrak{g})$-module.

For every $\xi_1, \ldots, \xi_p \in \mathfrak{g}$, the natural projection $\otimes^p \mathfrak{g} \subset \bigoplus_{k=0}^{p} \left(\otimes^k \mathfrak{g}\right) \to F_pU(\mathfrak{g})$ maps $\xi_1 \otimes \ldots \otimes (\xi_{i-1} \otimes \xi_i) \otimes \ldots \otimes \xi_p - \xi_1 \otimes \ldots \otimes (\xi_i \otimes \xi_{i-1}) \otimes \ldots \otimes \xi_p$ into the same element as $\xi_1 \otimes \ldots \otimes [\xi_{i-1} \otimes \xi_i] \otimes \ldots \otimes \xi_p$, that is, into an element of $F_{p-1}U(\mathfrak{g})$. Thus, there arises a linear map

$$S^p\mathfrak{g} \to F_pU(\mathfrak{g})/F_{p-1}U(\mathfrak{g}). \qquad (*)$$

Theorem (Poincaré-Birkhoff-Witt) *The map* $(*)$ *is an isomorphism.*

It is very easy to prove that $(*)$ is onto; the proof of vanishing the kernel is more involved (although quite elementary). We will not prove it here; the proof is contained in many books (see, for example, V. Varadarajan, "*Lie groups, Lie algebras, and their representations*," Springer Verlag Graduate Texts in Mathematics, Vol. 102, Section 3.2).

The Poincaré-Birkhoff-Witt theorem yields a construction of a basis in $U(\mathfrak{g})$.

Corollary. *Let the Lie algebra* $\mathfrak{g}$ *have a finite or countable basis* $g_1, g_2, g_3, \ldots$. *Then the monomials* $g_{i_1} g_{i_2} \ldots g_{i_k}$ *with* $i_1 \geq i_2 \geq \ldots \geq i_k$ *form a basis in* $U(\mathfrak{g})$. *In particular,* *such monomials with $k \leq p$ form a basis in $F_pU(\mathfrak{g})$. Also in particular, $\mathfrak{g}$ is naturally embedded into $U(\mathfrak{g})$.*

The notion of the universal enveloping algebra is also used in the description of a very important operation of extending representations of Lie subalgebras to representations of Lie algebras. The extended representation is called the *induced* representation. Namely, if $V$ is a representation of $\mathfrak{h} \subset \mathfrak{g}$, that is, a $U(\mathfrak{h})$-module, then we define the induced $U(\mathfrak{g})$-module as

$$\operatorname{Ind}_{\mathfrak{h}}^{\mathfrak{g}} V = U(\mathfrak{g}) \otimes_{U(\mathfrak{h})} V$$

(we use in this construction the fact that $U(\mathfrak{g})$ is a two-sided $U(\mathfrak{h})$-module). Informally speaking, for a $g \in \mathfrak{g} - \mathfrak{h}$ and $v \in V$, we add to $V$ an element $gv$, and similarly add all $g_1 \dots g_k v$, taking into account, however, all relations; for example, $g_1 g_2 v$ must be the same as $g_2 g_1 v + [g_1, g_2]v$.

Notice that both universal enveloping algebras and indused representations may be described in the language of adjoint functors.

Exercises (for those who are familiar with the notion of adjoint functors). 5. Prove that the functor $U\colon \mathcal{L}ie \to \mathcal{A}sso$ (where $\mathcal{L}ie$ is the category of (say, complex) Lie algebras and $\mathcal{A}sso$ is the category of associative algebras) is a left adjoint to the functor $L\colon \mathcal{A}sso \to \mathcal{L}ie$ which assigns to an assotiative algebra $A$ the Lie algebra $A$ with the commutator $[a, b] = ab - ba$.

6. Prove that the functor $\operatorname{Ind}_{\mathfrak{h}}^{\mathfrak{g}}\colon \mathfrak{h}\text{-}\mathcal{M}od \to \mathfrak{g}\text{-}\mathcal{M}od$ is a left adjoint to the restriction functor $R\colon \mathfrak{g}\text{-}\mathcal{M}od \to \mathfrak{h}\text{-}\mathcal{M}od$.

**5.2.2. Verma modules.** Let $\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+$ be a complex semisimple Lie algebra, and let $\lambda\colon \mathfrak{h} \to \mathbb{C}$ be a linear function. The function $\lambda$ gives rise to a one-dimensional representation $\mathbb{C}_\lambda$ of $\mathfrak{h}$: $\mathfrak{h}\mathfrak{z} = \lambda(\mathfrak{h}) \cdot \mathfrak{z}$. The projection $\mathfrak{h} \oplus \mathfrak{n}_+ \to \mathfrak{h}$ makes $\mathbb{C}_\lambda$ a (still one-dimensional) $(\mathfrak{h} \oplus \mathfrak{n}_+)$-module. The $\mathfrak{g}$-module

$$M(\lambda) = \operatorname{Ind}_{\mathfrak{h}\oplus\mathfrak{n}_+}^{\mathfrak{g}} \mathbb{C}_\lambda$$

is called the *Verma module* of type $\lambda$[1]). As an $\mathfrak{n}_-$-module, the Verma module is a free $U(\mathfrak{n}_-)$-module with one generator; thus, if $g_1, g_2, \dots$ is a basis in $\mathfrak{n}_-$ (one can use the basis $\{e_{-\alpha}\}$ with $\alpha \in \Delta^+$), then a basis in $M(\lambda)$ consists of $g_{i_1} \dots g_{i_k} v$ where $v = v_\lambda$ is the generator (called also the *vacuum vector*) and $k \geq 0, i_1 \geq \dots \geq i_k$. This description may be converted into an axiomatic definition of the Verma module independent of the inducing operation: it is a module with a basis as above with an additional property: $v$ is a highest weight vector of type $\lambda$ (the action of $\mathfrak{g}$ in $M(\lambda)$ described in this way is reconstructed from the data in the spirit of the Proof of Proposition 2 in Section 5.1.1).

Remark. A description of a Verma module over $\mathfrak{sl}(2, \mathbb{C})$ may be derived from Exercise 13 in Section 3.5, if an integer $n$ is replaced by an arbitrary complex number $\lambda$.

---

[1]) J. Dixmier, in his book "Universal enveloping algebras" writes that it would be more fair to call these modules Bernstein-Gelfand-Gelfand modules, but this term would be too long. One can notice that, precisely as Poincaré-Birkhoff-Witt is commonly abbreviated to *PBW*, Bernstein-Gelfand-Gelfand is commonly abbreviated to $BG^2$.

The most important (for us now) property of Verma modules lies in their "couniversality." Let $V$ be an arbitrary (not necessarily finite-dimensional) representation of $\mathfrak{g}$, and let $v \in V$ be a highest weight vector of type $\lambda$.

PROPOSITION. *There exists a unique* $\mathfrak{g}$*-homomorphism* $\varphi\colon M(\lambda) \to V$ *which takes* $v_\lambda$ *into* $v$.

*Proof.* This homomorphism $\varphi$ must take $g_{i_1} \dots g_{i_k} v_\lambda$ into $g_{i_1} \dots g_{i_k} v$, which proves both its existence and uniqueness.

**5.2.3. Irreducible representations with a given highest weight.** The kernel $\operatorname{Ker}\varphi$ of the homomorphism $\mu$ of the last Proposition must be a submodule of $M(\lambda)$ not containing $v_\lambda$ (and thus not equal to $M(\lambda)$). Let us discuss submodules of $M(\lambda)$. First of all, $M(\lambda)$ is, by construction, the direct sum of weight spaces $M(\lambda)^\mu$ where every $\mu$ has the form $\lambda - \alpha_1 - \dots - \alpha_k$ where $k \geq 0, \alpha_{i_1}, \dots, \alpha_{i_k} \in \Delta^+$. The space $M(\lambda)^\lambda$ is one-dimensional and is generated by $v_\lambda$. Here are two major properties of submodules of $M(\lambda)$.

PROPOSITION 1. *If* $L$ *is a submodule of* $M(\lambda)$*, then* $L = \bigoplus_\mu L \cap M(\lambda)^\mu$ *(in other words,* $N$ *has a basis consisting of weight vectors).*

*Proof.* Indeed, $L$ is also an $\mathfrak{h}$-submodule of the $\mathfrak{h}$-module $M(\lambda)$.

PROPOSITION 2. *There exists a unique submodule* $L(\lambda)$ *of* $M(\lambda)$ *such that*

(1) $v_\lambda \notin L(\lambda)$,

(2) *If* $L$ *is a submodule of* $M(\lambda)$ *not equal to* $M(\lambda)$*, then* $L \subset L(\lambda)$.

*Proof.* $L(\lambda)$ is the sum of all submodules of $M(\lambda)$ not equal to $M(\lambda)$. Notice that a submodule of $M(\lambda)$ is not equal to $M(\lambda)$, if and only if it does not contain $v_\lambda$, or, equivalently, has zero intersection with $M(\lambda)^\lambda$. Hence, the sum of such submodules also has zero intersection with $M(\lambda)^\lambda$.

The submodule $L(\lambda)$ of $M(\lambda)$ from Proposition 2 is called the *maximal submodule* of $M(\lambda)$.

COROLLARY OF PROPOSITION 2. *The Verma module* $M(\lambda)$ *is not irreducible if and only if it contains a highest weight vector not proportional to* $v_\lambda$*, or, equivalently, of type not equal to* $\lambda$.

THEOREM. (1) $M(\lambda)/L(\lambda)$ *is an irreducible representation of* $\mathfrak{g}$ *with the highest weight* $\lambda$.

(2) *Every irreducible representation of* $\mathfrak{g}$ *with a highest weight* $\lambda$ *is isomorphic to* $M(\lambda)L(\lambda)$*; in particular, such a representation is unique up to isomorphism.*

*Proof.* Part (1) is obvious, Part (2) follows from Proposition in Section 5.2.2 and Propositions 1 and 2 above. Indeed, let $V$ be an irreducible representation of $\mathfrak{g}$ with a highest weight vector $v$ of type $\lambda$. Proposition of Section 5.2.2 provides a homomorphism $\varphi\colon M(\lambda) \to V$ with $\varphi(v_\lambda) = v$ and it must be onto, since its image is a subrepresentation of $V$ which contains $v$ and hence is not zero. Hence $V \cong M(\lambda)/\operatorname{Ker}\varphi$. But $\operatorname{Ker}\varphi \not\ni v_\lambda$, it is a submodule of $M(\lambda)$ and it has to be maximal, since if $\operatorname{Ker}\lambda \underset{\neq}{\subset} L \not\ni v_\lambda$, then $V \underset{\neq}{\supset} \varphi(L) \neq 0$, so $V$ is not irreducible. Thus, $\operatorname{Ker}\varphi = L(\lambda)$ and $V \cong M(\lambda/L(\lambda)$.

This result demonstrates a great importance of two problems concerning Verma modules. First, this is a problem of reducibility/irreducibility. This problem has been fully solved for finite-dimensional semisimple Lie algebras and also for some important infinite-dimensional Lie algebras. Regarding that, I can mention two works: V.G. Kac and D.A. Kazhdan, "Structure of representations with highest weight of infinite-dimensional Lie algebras," Adv. Math., **34** (1978), 97–108, and Feigin B.L. and Fuchs, D.B., "Skew-symmetric invariant differential operators and Verma modules over the Virasoro algebra," Funct. Anal. Appl., **16** (1982), 114–126.

The second problem is to determine whether the quotient $M(\lambda)/L(\lambda)$ is finite-dimensional. We will discuss this in the next section.

**5.2.4. Finite-dimensional irreducible representations.** Let again $\mathfrak{g}$ be a finite-dimensional semisimple Lie algebra. Recall that for every positive root $\alpha$ there arises a subalgebra $\mathfrak{s}_\alpha = \mathfrak{g}^\alpha \oplus \mathbb{C}h_\alpha \oplus \mathfrak{g}^{-\alpha}$ of $\mathfrak{g}$ isomorphic to $\mathfrak{sl}(2, \mathbb{C})$ (se Section 3.4). To make this isomorphism more explicit, we can choose in it a basis $E_\alpha \in \mathfrak{g}^\alpha, H_\alpha \in \mathfrak{h}, F_\alpha \in \mathfrak{g}^{-\alpha}$ which satisfies the commutator relations of $\mathfrak{sl}(2)$: $[E_\alpha, F_\alpha] = H_\alpha$, $[H_\alpha, E_\alpha] = 2E_\alpha$, $[H_\alpha, F_\alpha] = -2F_\alpha$. For the moment, it is important for us that $H_\alpha = \dfrac{2}{\langle\alpha, \alpha\rangle} h_\alpha$ (see Section 3.6). We choose simple positive roots $\alpha_i$ and abbreviate the notation $H_{\alpha_i}$ to $H_i$ (as well as $E_{\alpha_i}$ and $F_{\alpha_i}$ to $E_i$ and $F_i$; these last notations will be used in the next Section). Thus, $H_1, \dots, H_r$ is a basis in $\mathfrak{h}$.

A weight $\lambda\colon \mathfrak{h} \to \mathbb{C}$ is called *integral dominant*, if every $\lambda(H_i)$ is a non-negative integer.

THEOREM. *The irreducible representation of $\mathfrak{g}$ with the highest weight $\lambda$ is finite-dimensional if and only if the weight $\lambda$ is integral dominant.*

We will not give a full proof of this theorem; we restrict ourselves to the remark that in one direction it is known to us. If the representation is finite-dimensional, then it is also a finite-dimensional representation of every $\mathfrak{s}_\alpha$. We already know (from Section 3.5) that the eigenvalues of the operator $h$ in every finite-dimensional representation of $\mathfrak{sl}(2)$ are integers, and the maximal of them must be non-negative. This means precisely that the highest weight is integral dominant. It remains to prove that if the highest weight is integral dominant, then the representation is finite-dimensional. We will not do it here, although it can be deduced, without much efforts, from our previous results.

EXERCISE 7. Determine the highest weights of the irreducible components of representations $V \otimes V$ and $V \otimes V \otimes V$ as described in Exercises 1, 2, and 3. Try to generalize this to the representations of Exercise 4.

**5.3. Generalizations: Kac-Moody Lie algebras.** In Section 5.2.4, we described elements $H_i, E_i, F_i,\ i = 1, \dots, r$ (where $r = \dim \mathfrak{h}$) of a complex semisimple Lie algebra $\mathfrak{g}$. We distinguished the relations $E_i, F_i] = H_i$, $[H_i, E_i] = 2E_i$, $[H_i, F_i] = -2F_i$. Also, we know that $[H_i, H_j] = 0$. There are some other relations which are described in terms of the so called *Cartan matrix*. This is an $r \times r$ matrix $A = \|a_{ij}\|$ determined by the formula

$$a_{ij} = \frac{2\langle\alpha_i, \alpha_j\rangle}{\langle\alpha_i, \alpha_i\rangle}$$

where $\alpha_1, \dots, \alpha_r$ are simple positive roots. The results of Section 3.6 imply that $a_{ii} = 2$ and if $i \neq j$, then $a_{ij}$ is a non-positive integer. It is also clear that the Cartan matrix is, in general, not symmetric, but is *symmetrizable*, that is the matrix $DA$ is symmetric, where

$D$ is the diagonal matrix with the diagonal entries $\langle \alpha_i, \alpha_i \rangle$. The Cartan matrix of a simple Lie algebra is immediately reconstructed from the Dynkin diagram: the diagonal entries are all 2, and the entry $a_{ij}$ is 0, if the verices number $i$ and $j$ are not connected by edges, are $-2$ or $-3$, if they are connected by a double or triple edge with the arrow directed to the vertex number $i$, and $-1$ in all other cases. Thus, the Cartan matrices of the Lie algebras of types $A_r, B_r, C_r, D_r$ are, respectively,

$$\begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & \ddots & & \text{zeroes} & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & 2 & -1 & \\ \text{zeroes} & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & \ddots & & \text{zeroes} & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & 2 & -1 & \\ \text{zeroes} & & & -1 & 2 & -1 \\ & & & & -2 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & \ddots & & \text{zeroes} & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & 2 & -1 & \\ \text{zeroes} & & & -1 & 2 & -2 \\ & & & & -1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & \ddots & & \text{zeroes} & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & 2 & -1 & -1 \\ \text{zeroes} & & & -1 & 2 & 0 \\ & & & -1 & 0 & 2 \end{bmatrix};$$

the Cartan matrices of the exceptional Lie algebras of types $G_2, F_4, E_6, E_7, E_8$ are, respectively,

$$\begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -2 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & -1 \\ 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & -1 & 0 & 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & -1 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 2 \end{bmatrix}.$$

The results of Section 3.6 imply that the generators $H_i, E_i, F_i$ of $\mathfrak{g}$ satisly some additional relations.

EXERCISE 8. Prove that the following relations hold:

$$[H_i, E_j] = a_{ij} E_j, \ (\operatorname{ad} E_i)^{-a_{ij}+1} E_j = 0,$$
$$[H_i, F_j] = -a_{ij} F_j, \ (\operatorname{ad} F_i)^{-a_{ij}+1} F_j = 0.$$

Together with relations listed before, these relations compose the following system of *Cartan-Kac-Moody relations* between the generators $H_i, E_i, F_i$:

$$\begin{aligned} [H_i, H_j] &= 0, & [E_i, F_j] &= \delta_{ij} H_i, \\ [H_i, E_j] &= a_{ij} E_j, & (\operatorname{ad} E_i)^{-a_{ij}+1} E_j &= 0, \\ [H_i, F_j] &= -a_{ij} F_j, & (\operatorname{ad} F_i)^{-a_{ij}+1} F_j &= 0. \end{aligned} \qquad (CKM)$$

Let now $A = \|a_{ij}\|$ be an integral $r \times r$ symmetrizable matrix (that is, there exists an integral diagonal matrix $D$ with positive diagonal entries such that the matrix $DA$ is symmetric) with $a_{ii} = 2$ and $a_{ij} \leq 0$ when $i \neq j$. Notice that the symmetrizability condition implies that $a_{ij} = 0$ if and only if $a_{ji} = 0$. We say that $A$ is *reducible*, if the set $\{1,\ldots,r\}$ splits into the union $I \cup J$ of two non-empty subsets such that $a_{ij} = 0$ for $i \in I, j \in J$, and is *irreducible*, if it is not reducible.

Definition. The Lie algebra $\mathfrak{g}(A)$ with $3r$ generators $H_i, E_i, F_i,\ i = 1,\ldots,r$ and relations $(GKM)$ is called a *Kac-Moody Lie algebra* with the Cartan matrix $A$. The number $r$ is called the *rank* of the Lie algebra $\mathfrak{g}(A)$.

Notice that there exist several, not fully equivalent, definitions of Kac-Moody Lie algebras. In the most general definition, $A$ is an arbitrary symmetrizable complex $r \times r$ matrix; however, in this case the last two relations must be changed.

Proofs of the most part of results formulated below can be found in the book by V.G. Kac, "Infinite-dimensional Lie algebras," Cambridge University Press, 1990.

Theorem 1. *The class of complex semisimple Lie algebras coincides with the class of Kac-Moody algebras $\mathfrak{g}(A)$ with positive definite symmetrized Cartan matrix $DA$.*

For non-negative integers $k_1,\ldots,k_r$, not all of which are zeroes, denote by $\mathfrak{g}(A)^{k_1,\ldots,k_r}$ the subspace of $\mathfrak{g}(A)$ spanned by all commutator monomials $[E_{i_1},[E_{i_2},\ldots,[E_{i_{N-1}},E_{i_N}]\ldots]]$ such that precisely $k_j$ of the indices $i_s$ is equal to $j$ (thus, $k_1+\ldots+k_r = N$). The space $\mathfrak{g}(A)^{-k_1,\ldots,-k_r}$ is defined in the same way with $F$ instead of $E$. The space spanned by $H_1,\ldots,H_r$ is denotes by $\mathfrak{h}$.

Theorem 2. *The space $\mathfrak{h}$ is $r$-dimensional (that is, $H_1,\ldots,H_r$ are linearly independent). The vector space $\mathfrak{g}(A)$ is a direct sum*

$$\left[\bigoplus\nolimits_{k_1,\ldots k_r} \mathfrak{g}(A)^{-k_1,\ldots,-k_r}\right] \oplus \mathfrak{h} \oplus \left[\bigoplus\nolimits_{k_1,\ldots k_r} \mathfrak{g}(A)^{k_1,\ldots,k_r}\right],$$

*and this direct sum decomposition determines a $\mathbb{Z}\oplus\ldots\oplus\mathbb{Z}$ grading of $\mathfrak{g}(A)$.*

The first and the last summands of the decomposition in Theorem 2 are subalgebras of $\mathfrak{g}(A)$, which can be denoted, in analogy with the finite-dimansional case, as $\mathfrak{n}_-$ and $\mathfrak{n}_+$. If the matrix $A$ is non-degenerate, then the decomposition is a root decomposition as was defined before (the root $\lambda\colon \mathfrak{h} \to \mathbb{C}$ corresponding to $\mathfrak{g}^{k_1,\ldots,k_n}$ is defined by the formula $\lambda(H_i) = \sum_j a_{ij}k_j$). If the matrix $A$ is degenerate, then $\mathfrak{g}(A)$ has a center (containing in $\mathfrak{h}$), and the weight decomposition is coarser than the decomposition of Theorem 2.

The case closest to the finite-dimensional case is the case when the matrix $DA$ has $r-1$ positive eigenvalues and one zero eigenvalue.

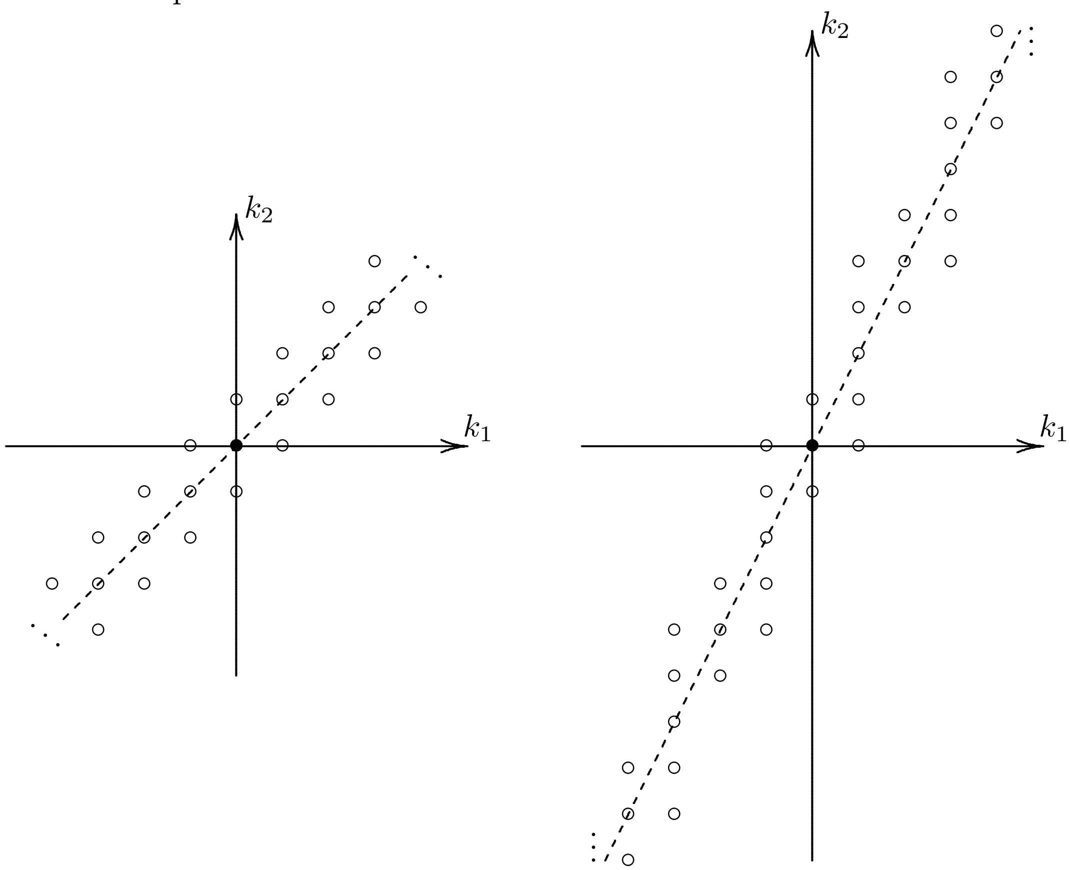
Theorem 3. *The class of Kac-Moody algebras $\mathfrak{g}(A)$ with the eigenvalues of the matrix $DA$ being as described above coincides with the class of infinite-dimensional Kac-Moody algebras with (at most) polynomial growth of $d_k = \dim\bigoplus_{k_1+\ldots+k_r=k}\mathfrak{g}(A)^{k_1,\ldots,k_r}$.*

Kac-Moody algebras of this group are called *affine*. If $r = 2$, there are (up to swapping rows and columns) two matrices $A$ of this class:

$$\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix}.$$

The spaces $\mathfrak{g}(A)^{k_1,k_2}$ are presented, schematically, in the following diagram. Black dot correspond to $\mathfrak{h}$, the dimension is 2; the basis of $\mathfrak{h}$ is $\{H_1, H_2\}$, and the element $H_1 + H_2$

for the left diagram and $H_1 + 2H_2$ for the right diagram is central. White dots correspond to 1-dimensional spaces.




EXERCISES. 8. Prove that after factorizing over the center, the first algebra becomes isomorphic to $\mathfrak{sl}(2, \mathbb{C}) \otimes \mathbb{C}[t, t^{-1}]$. The elements $E_1, E_2, F_1, F_2$ correspond, with respect to this isomorphism, to

$$\begin{bmatrix} 0 & 0 \\ t & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & t^{-1} \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

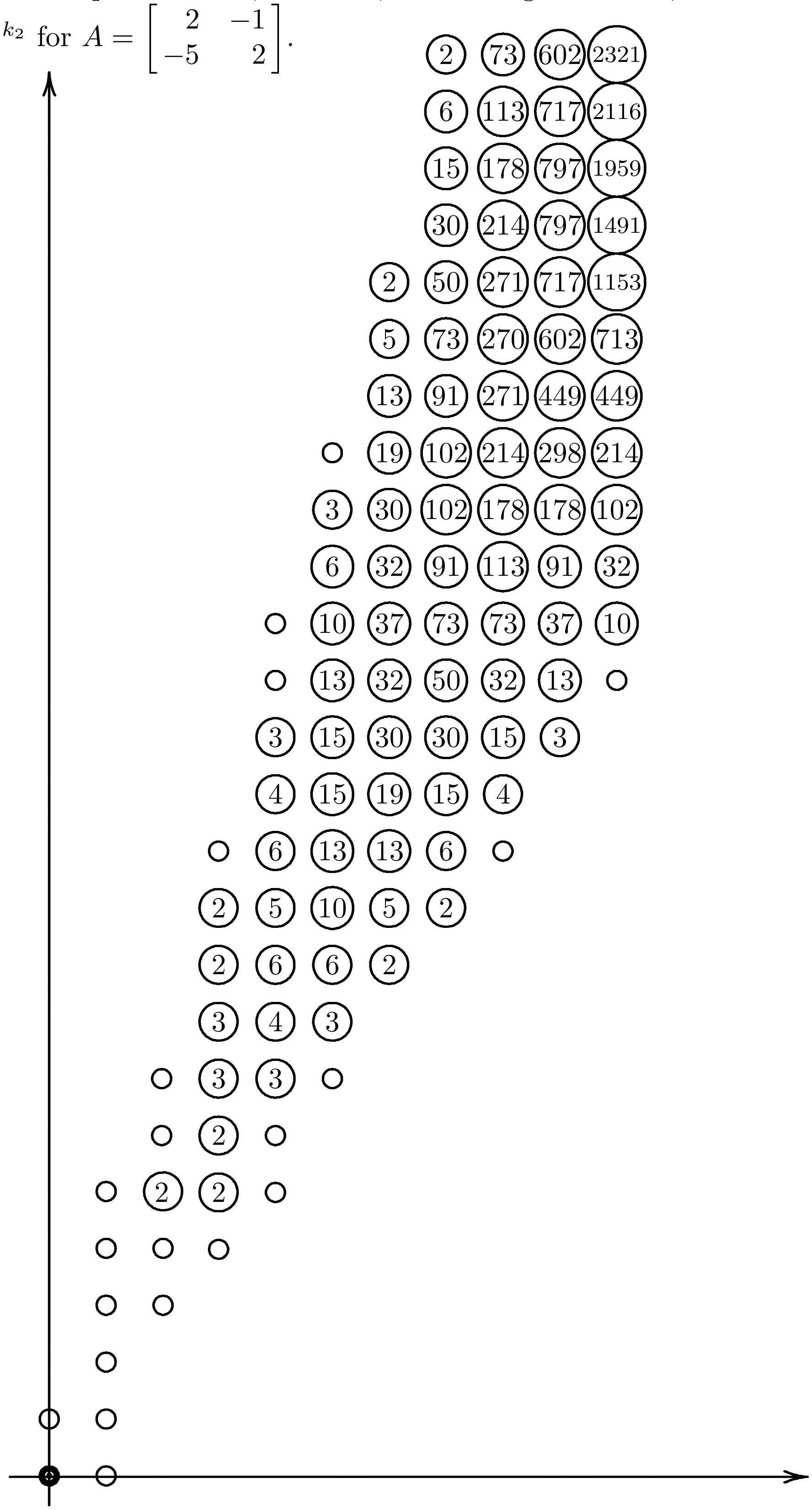
9. Let $A$ and $B$ be the spaces of skew-symmetric and symmetric matrices in $\mathfrak{sl}(3, \mathbb{C})$. Prove that, after factorizing over the center, the second algebra becomes isomorphic to a subalgebra of $\mathfrak{sl}(3, \mathbb{C})$ consisting of elements of the form

$$\ldots + t^{-3}b_{-3} + t^{-2}a_{-2} + t^{-1}b_{-1} + a_0 + tb_1 + t^2a_2 + t^3b_3 + \ldots$$

(finite sum) where $\ldots, b_{-3}, b_{-1}, b_1, b_3, \ldots \in B,\ \ldots, a_{-2}, a_0, a_2, \ldots \in A$.

A full classification of affine Lie algebras is known. There are two kinds of them, both represented for $r = 2$. They are either $\mathfrak{g} \otimes \mathbb{C}[t, t^{-1}]$ where $\mathfrak{g}$ is a finite-dimensional simple Lie algebra, or are constructed from a finite-dimensional simple Lie algebra and a finite order automorphism of its Dynkin diagram (this order is 2, except an automorphism of order 3 of the Dynkin diagram $D_4$).

For other type of Cartan matrices, the growth of the dimensions $d_k$ is exponential. To demonstrate this phenomenon, we show, in the diagram below, the dimensions of the spaces $\mathfrak{g}(A)^{k_1,k_2}$ for $A = \begin{bmatrix} 2 & -1 \\ -5 & 2 \end{bmatrix}$.



If the Cartan matrix is reducible, then the Kac-Moody algebra splits into a direct sum of two Kac-Moody algebras. If it is irreducible, then the Kac-Moody algebra is *graded simple*, that is, it does not have proper ideals $I$ such that $I = \bigoplus \left[I \cap \mathfrak{g}(A)^{k_1,\ldots,k_r}\right]$.

EXERCISE 10. Prove that the Kac-Moody algebras of Exercises 8 and 9 are graded simple, but not simple.

The theory of Sections 5.2.2 and 5.2.3 can be extended to Kac-Moody algebras without essential changes. The same is true regarding the definitions and formulas of the next Section. See details in the book of Kac cited above.

**5.4. Formulas for dimensions.** We return to the representations of a finite-dimensional complex semisimple Lie algebra $\mathfrak{g}$.

**5.4.1. The Weyl dimension formula.** Let $\rho$ be one half of the sum of all positive roots of $\mathfrak{g}$ and let $V(\lambda)$ be the space of the irreducible finite-dimension representation of $\mathfrak{g}$ with th highest weight $\lambda$.

THEOREM (Weyl).

$$\dim V(\lambda) = \prod_{\alpha \in \Delta^+} \frac{\langle \lambda + \rho, \alpha \rangle}{\langle \rho, \alpha \rangle}.$$

*Proof* see in the books of Varadarajan (see reference on Page 74) or Onishchik–Vinberg (reference on Page 70). The reader can try to prove it as an exercise.

EXERCISE 11. Let $m_1, \ldots, m_n$ be positive integers. Prove that the dimension of the ireducible representation of $\mathfrak{sl}(n+1, \mathbb{C})$ with the highest weight $\lambda$ defined by the formula $\lambda(H_i) = m_i - 1$ (where $H_i = E_{ii} - E_{i+1,i+1}$) is

$$\prod_{1 \le i \le j \le n} \frac{m_i + m_{i+1} + \ldots + m_j}{j - i + 1}.$$

EXERCISE 12. Deduce from this formula the results of Exercises 1–3 of Section 5.1.2.

**5.4.2. The Freudenthal formula.** The dimension formula of Section 5.4.1 does not solve the problem of computing the dimensions of the weight spaces. Denote by $d_\mu(\lambda)$ the dimension of the weight space $V(\lambda)^\mu$ (if $\mu$ is not a weight os a representation $V(\lambda)$, then $d_\mu(\lambda) = 0$). We know that $d_\lambda(\lambda) = 1$. the following formula provides a recursive procedure of computing dimensions $d_\mu(\lambda)$ when all $d_\nu(\lambda)$ with $\nu = \mu + \alpha_1 + \ldots + \alpha_k$ for $\alpha_1, \ldots, \alpha_k \in \Delta^+$ are known.

THEOREM (Freudenthal).

$$[\langle \lambda + \rho, \lambda + \rho \rangle - \langle \mu + \rho, \mu + \rho \rangle] d_\mu(\lambda) = 2 \sum_{\alpha \in \Delta^+, k > 0} \langle \mu + k\alpha, \alpha \rangle d_{\mu + k\alpha}(\lambda).$$

*Proof*; the same references as for the Weyl theorem in Section 5.4.1.

Notice that the Freudental formula cannot be used for computing $d_\mu(\lambda)$, if $\langle \lambda + \rho, \lambda + \rho \rangle = \langle \mu + \rho, \mu + \rho \rangle$. Notice also that this formula cn be applied for infinite-dimensinal representations (of a finite-dimensional semisimple Lie algebra) with highest weight.

Exercise 13. Let $\mathfrak{g} = \mathfrak{sl}(2)$. Then there is only one positive root, $\alpha$, $\rho = \dfrac{\alpha}{2}$, $\langle \rho, \rho \rangle = \dfrac{\langle \alpha, \alpha \rangle}{4} = \dfrac{1}{2}$. Deduce from the Freudenthal formula that if $\lambda = n\rho$, $\mu = m\rho$, then

$$d_\mu(\lambda) = \begin{cases} 0, & \text{if } n - m \notin 2\mathbb{Z}_{\geq 0}; \\ 1, & \text{if } n - m \in 2\mathbb{Z}_{\geq 0}, \text{ and } n \notin \mathbb{Z}_{\geq 0}; \\ 1, & \text{if } n \in \mathbb{Z}_{\geq 0}, \text{ and } m = n, n-2, \ldots, 2-n, -n. \end{cases}$$

**5.4.3. The Weyl character formula.** We use the notation of Section 5.4.2. The *character* of the representation $V(\lambda)$ with the highest weight $\lambda$ is, by definition,

$$\operatorname{Ch} V(\lambda) = \sum_\mu d_\mu(\lambda) e^\mu$$

where $e^\mu$ is a symbol satisfying the rule $e^\mu e^\nu = e^{\mu+\nu}$. To state the Weyl formula for $\operatorname{Ch} V(\lambda)$ we need the notion of the *Weyl group*.

For a non-zero $\xi \in \mathfrak{h}^*$ we define $s_\xi \colon \mathfrak{h} \to \mathfrak{h}$ as a reflection in the hyperplane $\{\xi(h) = 0\}$; for a non-zero root $\alpha$,

$$s_\alpha(h) = h - \alpha(h) h_\alpha.$$

The Weyl group $W = W(\mathfrak{g})$ of $\mathfrak{g}$ is defined as the subgroup of the group of orthogonal transformations of $\mathfrak{h}$ generated by $s_\alpha, \alpha \in \Delta^+$. It is possible to prove that the Weyl group (for a finite-dimensional semisimple Lie algebra) is always finite.

A better known definition of the Weyl group is given in terms of compact Lie groups. Let $G$ be a compact semisimple Lie group, and let $T$ bi its "maximal torus", that is, a subgroup corresponding to a Cartan subalgebra of $\operatorname{Lie} G$. Then $W(G) = N(T)/T$ where $N(T)$ is the normalizer of $T$.

Exercises. 14. For $\mathfrak{sl}(n, \mathbb{C})$ the Weyl group is isomorphic to the symmetric group $S_n$.

15. For $\mathfrak{o}(2n+1)$ and $\mathfrak{sp}(n)$, the Weyl group is isomorphic to the group of transformations of $\mathbb{C}^n$ which permute the vectors $e_1, \ldots, e_n$ of the standard basis and change signs at some of them.

16. For $\mathfrak{o}(2n)$, the Weyl group is isomorphic to the group of transformations of $\mathbb{C}^n$ which permute the vectors $e_1, \ldots, e_n$ of the standard basis and change signs at an even number of them.

17. Describe the group $W(G_2)$ (it must have the order 12).

There is a homomorphism $\varepsilon \colon W \to \{\pm 1\}$ such that $\varepsilon(s_\alpha) = -1$ for every $\alpha$.

Theorem (Weyl)

$$\operatorname{Ch} V(\lambda) = \frac{\sum_{w \in W} \varepsilon(w) e^{w(\lambda+\rho)}}{e^\rho \prod_{\alpha \in \Delta^+} (1 - e^{-\alpha})}.$$

There exists a generalization of the Weyl group and the Weyl character formula (called the Weyl-Kac character formula) to arbitrary Kac-Moody Lie algebras. This can be found in the Kac book (mentioned in Section 5.3).

Dixi.