Lie Algebras and Representation Theory

Spring Term 2003

Andreas Čap

Institut für Mathematik, Universität Wien, Strudlhofgasse 4, A–1090 Wien

E-mail address: Andreas.Cap@esi.ac.at

Contents

Preface

The aim of this course is to develop the basic general theory of Lie algebras to give a first insight into the basics of the structure theory and representation theory of semisimple Lie algebras.

A problem one meets right in the beginning of such a course is to motivate the notion of a Lie algebra and to indicate the importance of representation theory. The simplest possible approach would be to require that students have the necessary background from differential geometry, present the correspondence between Lie groups and Lie algebras, and then move to the study of Lie algebras, which are easier to understand than the Lie groups themselves. This is unsatisfactory however, since in the further development the only necessary prerequisite is just a good knowledge of linear algebra, so requiring a background in differential geometry just for understanding the motivation seems rather strange.

Therefore, I decided to start the course with an informal discussion of the background. The starting point for this introduction is the concept of a group action, which is very intuitive when starting from the idea of a group of symmetries. Group representations then show up naturally as actions by linear maps on vector spaces. In the case of a Lie group (with matrix groups being the main example) rather than a discrete group, one may linearize the concepts to obtain a Lie algebra and representations of this Lie algebra. The last part of the introduction is then a short discussion of the correspondence between Lie groups and Lie algebras, which shows that in spite of the considerable simplification achieved by passing to the Lie algebra, not too much information is lost.

Most of the rest of the course is based on parts of the second chapter of my forthcoming book "Parabolic geometries" (a joint work with J. Slovák from Brno).

Chapter 2 discusses the general theory of Lie algebras. We start by discussing nilpotent and solvable Lie algebras, and prove the fundamental theorems of Engel and Lie. Next we switch to the discussion of semisimple, simple and reductive Lie algebras. We discuss representations and the Killing form and prove Cartan's criteria for solvability and semisimplicity in terms of the Killing form. We give a proof of complete reducibility of representations of semisimple Lie algebras which is independent of the structure theory of such algebras. This is used to prove that any semisimple Lie algebra is a direct sum of simple ideals. Finally, we describe a systematic way to produce examples of reductive and semisimple Lie algebras of matrices. Some background from linear algebra (in particular concerning Jordan decompositions) is reviewed in the text.

Chapter 3 studies the structure theory of complex semisimple Lie algebras, which is also a fundamental ingredient for the study of representations of such algebras. Choosing a Cartan subalgebra, one obtains the root decomposition of the given Lie algebra into simultaneous eigenspaces under the adjoint action of the Cartan subalgebra. General results on Jordan decompositions show that the elements of the Cartan subalgebra are simultaneously diagonalizable in any finite dimensional representation, thus leading to the weight decomposition. The structure of the root decomposition can be analyzed using the representation theory of $\mathfrak{sl}(2, \mathbb{C})$, which is easy to describe. In that way, one associates with any complex semisimple Lie algebra an abstract root system, which is simply a nice set of vectors in a finite dimensional inner product space. We conclude the chapter by briefly discussing the classification of irreducible root systems, and how this can be used to give a complete classification of complex simple Lie algebras.

The basic theory of complex representations of complex semisimple Lie algebras is studied in chapter 4. With the background developed so far, we quickly arrive at a description of the possible weights of finite dimensional representations. Next, we study highest weight vectors and show that in a finite dimensional representation any such vector generates an irreducible subrepresentation. Using this, we arrive quickly at the result that a finite dimensional irreducible representation is determined up to isomorphism by its highest weight, which has to be dominant an algebraically integral. Next, we discuss two approaches to the proof of existence of finite dimensional irreducible representations with any dominant integral highest weight. The first approach is on a case–by–case basis, using fundamental representations and tensor products. We first discuss the necessary background from multilinear algebra, and then describe the fundamental representations (and some basic relations between them) for the classical simple Lie algebras. Secondly, we outline the general proof for existence of irreducible representations via Verma modules. The necessary background on universal enveloping algebras and induced modules is discussed.

The final chapter offers a brief survey various tools that can be used to describe irreducible representations and to split general representations into irreducible pieces. The first part of the chapter deals with tools for general complex semisimple Lie algebras. We discuss the isotypical splitting, the Casimir element, and various formulae for multiplicities of weights and characters. As an important example, we discuss the decomposition of a tensor product of two irreducible representations. The second part is devoted to the relation between representations of $\mathfrak{gl}(n,\mathbb{C})$ and representations of permutation groups. We discuss Young diagrams and Young symmetrizers, and Weyl's construction of irreducible representations of the classical simple Lie groups in terms of Schur functors.

There are several good books on Lie algebras and representation theory available, which usually however are too detailed for serving as a basis for a relatively short course. Two particularly recommendable sources are the books "Lie groups beyond an introduction" by A.W. Knapp (which I will refer to as [**Knapp**]) and "Representation Theory A First Course" by W. Fulton and J. Harris (which I will refer to as [**Fulton-Harris**]). Both these books do not only discuss Lie algebras but also Lie groups, and [**Fulton-Harris**] also discusses representations of finite groups. The two books also complement each other nicely from the approach taken by the authors: [**Fulton-Harris**] emphasizes examples and the concrete description of representations of the classical simple Lie algebras, [**Knapp**] contains a detailed account of the general theory and also discussed real Lie algebras and Lie groups. Two other recommendable texts which only discuss Lie algebras are the books "Introduction to Lie Algebras and Representation Theory" by J.E. Humphreys, and "Notes on Lie algebras" by H. Samelson. A nice short text is the book "Lectures on Lie Groups and Lie Algebras" by R. Carter, G. Segal, and I. Mac Donald. Apart from a brief survey of the theory of complex semisimple Lie algebras, this also offers an introduction to Lie Groups with an emphasis on the compact case, and an introduction to complex algebraic groups.

CHAPTER 1

Background

Large parts of the theory of Lie algebras can be developed with very little background. Indeed, mainly a good knowledge of linear algebra is needed. Apart from that, only a bit of Euclidean geometry shows up, which again can be traced back to linear algebra. On the other hand, Lie algebras usually not mentioned in introductory courses, so most of the students in this course probably have not heard the definition of a Lie algebra before. Moreover, this definition will probably sound rather strange to most beginners, since skew symmetry and the Jacobi identity are much less intuitive than commutativity and associativity.

Hence I have decided not to start with the abstract definition of a Lie algebra and then develop the theory, but rather to indicate first where the concepts come from, and why it may be a good idea to study Lie algebras and their representations. In particular, I want to show how the idea of a group of symmetries leads (via actions and representations of groups) to Lie algebras and their representations. Moreover, I want to point out in this chapter some examples in which thinking in terms of representation theory is very helpful.

Group actions and group representations

1.1. Symmetries and group actions. The idea of a symmetry is probably one of the basic concepts in mathematics. Usually this is understood as having a distinguished set of functions from a set X to itself which may be thought of as preserving some additional structure. The basic features are that the composition of two symmetries is again a symmetry and that any symmetry is a bijection, whose inverse is a symmetry, too. Of course, this implies that the identity map is always a symmetry. This simply means that the symmetries form a subset of the group $\mathrm{Bij}(X)$ of bijective functions from X to itself (with the group multiplication given by composition), which is closed under the group multiplication and inversion and thus a subgroup.

Let us look at a simple example: Consider a regular triangle X in the plane, and let us number its vertices by $\{1, 2, 3\}$. Then there are a few obvious symmetries of this figure: First, we can consider the reflections in the three lines going to one of the vertices and the mid point of the opposite edge. On the other hand, we may rotate the figure by $\pm\frac{2\pi}{3}$. If we look at the action of these symmetries on the vertices, we see that any reflection fixes one vertex and exchanges the other two vertices, while the two rotations cyclically permute the three vertices. Since also the identity is a symmetry, we see that any bijection $\sigma : \{1, 2, 3\} \to \{1, 2, 3\}$ of the set of edges can be extended to exactly one of the symmetries above.

Otherwise put, denoting by $\mathfrak{S}_3$ the permutation group of three letters, we can associate to each $\sigma \in \mathfrak{S}_3$ a bijection $f_\sigma : X \to X$ of X. Since the permutation corresponding to a symmetry is just given by the restriction to the vertices, we see that tis is compatible with multiplication, i.e. $f_{\sigma\tau} = f_\sigma \circ f_\tau$. This simply means that $\sigma \mapsto f_\sigma$ is a group homomorphism $\mathfrak{S}_3 \to \mathrm{Bij}(X)$. Another way to formulate this, is that we associate to a permutation σ and a point $x \in X$ the point $\sigma \cdot x := f_\sigma(x)$. In this picture, the fact that $\sigma \mapsto f_\sigma$ is a group homomorphism reads as $\mathrm{id} \cdot x = x$ and $\sigma \cdot (\tau \cdot x) = (\sigma\tau) \cdot x$, which looks very similar to the definition of a group. This is a special instance of the following general

DEFINITION. Let X be a set and let G be a group with neutral element $e \in G$.
(1) A *left action* of G on X is a map $\varphi : G \times X \to X$ such that $\varphi(e, x) = x$ and $\varphi(g, \varphi(h, x)) = \varphi(gh, x)$ for all $x \in X$ and $g, h \in G$.
(2) Given a left action $\varphi : G \times X \to X$ and a point $x_0 \in X$ we define the *orbit* $G \cdot x_0$ of x_0 to be the subset $\{x : \exists g \in G : \varphi(g, x_0) = x\} \subset X$ and the *isotropy subgroup* G_{x_0} of x_0 to be $\{g \in G : \varphi(g, x_0) = x_0\} \subset G$.

By definition of an action, for $g \in G$ with inverse g^{-1}, we get $\varphi(g^{-1}, \varphi(g, x)) = \varphi(e, x) = x$ for any $x \in X$. Thus for any $g \in G$ the map $x \mapsto \varphi(g, x)$ is a bijection $X \to X$, so we may also view φ as mapping G to the group of bijections of X, and the definition of an action means that this is a group homomorphism.

If there is no risk of confusion, we will write $g \cdot x$ for $\varphi(g, x)$, in which case the defining properties become $e \cdot x = x$ and $g \cdot (h \cdot x) = gh \cdot x$. The concepts in (2) become very natural if one thinks about the action as a way to use elements of g to move the points of X around. Then the orbit of x_0 is simply the set of all points that can be reached from x_0, while G_{x_0} is the set of all elements of G which do not move the given point x_0. If $g \in G_{x_0}$, then $g^{-1} \cdot x_0 = g^{-1} \cdot g \cdot x_0 = e \cdot x_0 = x_0$, while for $g, h \in G_{x_0}$ we get $gh \cdot x_0 = g \cdot (h \cdot x_0) = g \cdot x_0 = x_0$, so G_{x_0} is really a subgroup of G.

It is easy to describe the orbits and isotropy groups in the example of the regular triangle from above. For example, for any of the vertices, the orbit consists of all three vertices, while the isotropy subgroup consists of two elements, the identity and the reflection fixing the given vertex. It is a good exercise to determine the orbits and isotropy groups of other points of the triangle.

While the example of the triangle is simple and instructive, it is not exactly the kind of example that we will be interested in the sequel, since the symmetry group is discrete in that case. We will be interested in the case of continuous symmetry groups. Let us discuss two relevant examples of this situation:

EXAMPLES. (1) consider the unit circle S^1 in $\mathbb{R}^2$. Of course, one may view rotation by any angle as a symmetry of the circle. More conceptually, we will view $\mathbb{R}^2$ as $\mathbb{C}$ and S^1 as $U(1) = \{z \in \mathbb{C} : |z| = 1\}$. This is a group under multiplication, and rotations are just given by the multiplication by some fixed $z_0 \in U(1)$. Thus, we are led to view the multiplication $U(1) \times U(1) \to U(1)$ as a left action of the group $U(1)$ on itself. The two defining properties of an action just boil down to associativity of the multiplication and the definition of the unit element. Note that the orbit of any $z \in U(1)$ under this action is the full group $U(1)$ while its isotropy group consists of the unit element only.

Of course, this can be similarly done for any group.
(2) In many cases, one meets very interesting symmetries that do not have a direct geometric interpretation but come rather from formal considerations. This is not only important in mathematics but also in large parts of theoretical physics.

A simple example of this situation is the group $GL(n, \mathbb{R})$ of invertible $n \times n$–matrices with real coefficients acting on $\mathbb{R}^n$ in the usual way, i.e. the action is given by $(A, x) \mapsto Ax$. An interpretation of this action in terms of symmetry is that the orbits under this action consist exactly of the coordinate expressions of a vector for all possible bases of $\mathbb{R}^n$. Of course, as it stands this is not yet very interesting, since for $x \neq 0$ the orbit is simply $\mathbb{R}^n \setminus \{0\}$, while the orbit of the point 0 is just $\{0\}$. This corresponds to the fact that the only thing one may say about a vector in $\mathbb{R}^n$ in a basis independent way is whether it is zero or not. We shall see immediately that this example nonetheless quickly leads to very interesting and relevant problems.

As an exercise in that direction, the reader should try to describe the orbits of the action of $GL(n,\mathbb{R})$ on $\mathbb{R}^n \times \mathbb{R}^n$ defined by $A \cdot (x,y) := (Ax, Ay)$.

1.2. Group representations. A *representation* of a group G is simply an action of G on a $\mathbb{K}$–vector space V by linear maps. So we need an action $\varphi : G \times V \to V$ such that $\varphi(g, v + rw) = \varphi(g, v) + r\varphi(g, w)$ for all $g \in G$, $v, w \in V$ and $r \in \mathbb{K}$. We will only be interested in the cases $\mathbb{K} = \mathbb{R}$ and $\mathbb{K} = \mathbb{C}$ in the sequel. This condition is also very simple in the alternative point of view of actions as homomorphisms to groups of bijections: A representation of G is a homomorphism from G to the group $GL(V)$ of invertible linear maps $V \to V$. If V is finite dimensional, we may view it as $\mathbb{K}^n$ so we deal with homomorphisms $G \to GL(n,\mathbb{K})$.

By construction, example (2) from 1.1 describes a representation of $GL(n,\mathbb{R})$ on $\mathbb{R}^n$, the *standard representation*.

When dealing with representations rather than general actions, there are several additional natural concepts: Let V be a representation of the group G and $W \subset V$ a linear subspace. Then W is called *invariant* if $g \cdot w \in W$ for all $w \in W$ and all $g \in G$. If $W \subset V$ is an invariant subspace, we may view the map $w \mapsto g \cdot w$ as a representation of G on W. For any representation V, the trivial subspaces $\{0\}$ and V are invariant. If these are the only invariant subspaces, then V is called *irreducible*.

A related concept is the following: Recall that for two vector spaces V and W, the direct sum $V \oplus W$ is the product $V \times W$ with the component wise vector space operations. If we have representations of some group G on V and W, then there is a natural representation on $V \oplus W$ defined by $g \cdot (v, w) := (g \cdot v, g \cdot w)$, called the direct sum of the given representations. Conversely, if we have given a representation V and invariant subspaces $W_1, W_2 \subset V$ such that any $v \in V$ can be uniquely written as a sum of an element of W_1 and an element of W_2 (which is usually stated as $V = W_1 \oplus W_2$) then we may identify the representation of V with the direct sum of the restrictions to W_1 and W_2. If there are non–trivial invariant subspaces with this property, then the representation V is called *decomposable*, otherwise it is called *indecomposable*.

Of course, any irreducible representation is indecomposable, but the converse is not true in general: Consider the set $G := \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} : a, b, c \in \mathbb{R}, a, c \neq 0 \right\}$ of invertible real upper triangular 2×2–matrices. One immediately verifies that G is a group under matrix multiplication. Of course, we have a standard representation of G on $\mathbb{R}^2$. The line spanned by the first unit vector e_1 is an invariant subspace of $\mathbb{R}^2$, but it is easy to see that apart from $\{0\}$ and $\mathbb{R}^2$, this is the only invariant subspace. Hence V has to be indecomposable although it admits a non–trivial invariant subspace.

A basic way how to pass from actions to representations is by looking at functions. Let us start in a slightly more general setting: Assume that we have given actions of a group G on two sets X and Y and consider the set $\mathcal{F}(X, Y)$ of all functions from X to Y. For $\varphi : X \to Y$ and $g \in G$, we define $g \cdot \varphi : X \to Y$ by $(g \cdot \varphi)(x) := g \cdot (\varphi(g^{-1} \cdot x))$. Obviously, $e \cdot \varphi = \varphi$ and

$$(h \cdot (g \cdot \varphi))(x) = h \cdot ((g \cdot \varphi)(h^{-1} \cdot x)) = h \cdot g \cdot (\varphi(g^{-1} \cdot h^{-1} \cdot x)),$$

and since $(hg)^{-1} = g^{-1}h^{-1}$ we see that this equals $((gh) \cdot \varphi)(x)$. Hence we have defined an action of G on the space $\mathcal{F}(X, Y)$. Moreover, if $Y = V$ is a vector space, then $\mathcal{F}(X, V)$ is a vector space under pointwise operations. If we start with a representation on V (and an arbitrary action on X), then

$$(g\cdot(\varphi+r\psi))(x) = g\cdot(\varphi+r\psi)(g^{-1}\cdot x) = g\cdot(\varphi(g^{-1}\cdot x)+r\psi(g^{-1}\cdot x)) = (g\cdot\varphi)(x)+r(g\cdot\psi)(x),$$

and we obtain a representation of G on $\mathcal{F}(X,V)$. In particular, we may always choose $V=\mathbb{K}$ with the trivial representation defined by $g\cdot r = r$ for all $g\in G$ and $r\in\mathbb{K}$ to obtain a representation on $\mathcal{F}(X,\mathbb{K})$.

EXAMPLES. (1) Let us look at $\mathfrak{S}_3$ acting on the set $\{1,2,3\}$ in the obvious way, and the trivial representation on $\mathbb{K}$. Writing x_i for $x(i)$ for a function $x:\{1,2,3\}\to\mathbb{K}$ we can view a function as an element (x_1,x_2,x_3) of $\mathbb{K}^3$, and hence obtain a representation of $\mathfrak{S}_3$ on $\mathbb{K}^3$ defined by $\sigma\cdot(x_1,x_2,x_3) = (x_{\sigma^{-1}(1)},x_{\sigma^{-1}(2)},x_{\sigma^{-1}(3)})$, i.e. the group acts by permuting coordinates. Of course, the constant functions form an invariant subspace (which is a general phenomenon for representations of the form $\mathcal{F}(X,\mathbb{K})$), the subspace $\{(r,r,r): r\in\mathbb{K}\}$. On this subspace, the action of $\mathfrak{S}_3$ is obviously trivial, i.e. $\sigma\cdot(r,r,r) = (r,r,r)$ for any σ and r. There is also an obvious complementary subspace, namely the space V of all (x,y,z) such that $x+y+z=0$. A conceptual way to see this is to note that the standard inner product $\langle\ ,\ \rangle$ on $\mathbb{R}^3$ is $\mathfrak{S}_3$–invariant, i.e. $\langle\sigma\cdot x,\sigma\cdot y\rangle = \langle x,y\rangle$ for all $x,y\in\mathbb{R}^3$ and $\sigma\in\mathfrak{S}_3$. Now if $W\subset\mathbb{R}^3$ is an invariant subspace, consider the orthogonal complement $W^\perp = \{y\in\mathbb{R}^3 : \langle y,w\rangle = 0\quad\forall w\in W\}$. For $y\in W^\perp$, $w\in W$ and $\sigma\in\mathfrak{S}_3$, we then get $\langle\sigma\cdot y,w\rangle = \langle y,\sigma^{-1}\cdot w\rangle$, which vanishes since W is invariant. Obviously, we have $V = \{(r,r,r): r\in\mathbb{K}\}^\perp$. The resulting representation of $\mathfrak{S}_3$ on V is usually called the *standard representation* of $\mathfrak{S}_3$. It is an easy exercise to prove that this representation is irreducible.

One can do a similar construction for all the permutation groups $\mathfrak{S}_k$ for $k\in\mathbb{N}$. There is an obvious representation of $\mathfrak{S}_k$ on $\mathbb{R}^k$, the vectors with all components equal form an invariant subspace on which the group acts trivially, and the orthogonal complement of this is an irreducible representation which is called the standard representation of $\mathfrak{S}_k$.
(2) Let us consider the action on $U(1) = \{z\in\mathbb{C}: |z| = 1\}$ on itself by multiplication, and the corresponding representation of $U(1)$ on $\mathcal{F}(U(1),\mathbb{C})$, i.e. $(z\cdot f)(w) = f(z^{-1}w)$. Given $f: U(1)\to\mathbb{C}$, we may consider $\tilde{f}:\mathbb{R}\to\mathbb{C}$ defined by $\tilde{f}(t) := f(e^{it})$. Then $\tilde{f}(t+2\pi) = \tilde{f}(t)$, so we obtain periodic functions of period 2π in this way. Given $z\in U(1)$, we may choose $t_0\in[0,2\pi)$ such that $z = e^{it_0}$ and then $\widetilde{z\cdot f}(t) = \tilde{f}(t-t_0)$, so the our representation corresponds just to shifting functions. In particular, we see that continuous, differentiable, and smooth (infinitely often differentiable) functions define invariant subspaces. (Since $U(1)$ is a topological space and a smooth manifold, it makes sense to talk about continuous, differentiable, and smooth functions, but alternatively one may define these properties by requiring $\tilde{f}$ to have them.)

To look for invariant subspaces, let us try the simplest possibility of one–dimensional invariant subspaces. A function f spans a one–dimensional invariant subspace if and only if there is a function $\alpha: U(1)\to\mathbb{C}$ such that $z\cdot f = \alpha(z)f$ for all $z\in U(1)$. The equation $zw\cdot f = z\cdot(w\cdot f)$ then forces $\alpha(zw) = \alpha(z)\alpha(w)$, i.e. α has to be a group homomorphism. Moreover, we can immediately say what the function f has to look like: Indeed, $f(z) = (z^{-1}\cdot f)(1) = \alpha(z^{-1})f(1)$. Since multiplying f by a nonzero factor does not change the subspace, we may simply assume that $f(z) = \alpha(z^{-1})$. In particular, if f is continuous, differentiable or smooth, then α has the same property.

The obvious choices for homomorphism $U(1)\to\mathbb{C}$ are the mappings $z\mapsto z^k$ for $k\in\mathbb{Z}$ and one can show that these are the only continuous homomorphisms. Hence for any $k\in\mathbb{Z}$ we have a one–dimensional invariant subspace in $\mathcal{F}(U(1),\mathbb{R})$ spanned by the function $z\mapsto z^k$. It further turns out that any continuous function on $U(1)$ can be approximated by linear combinations of these power functions, and they even span a dense subspace of the space $L^2(U(1))$ of square integrable functions on $U(1)$. This is the basis for the theory of Fourier series. For applications of Fourier series the main point is that the basic functions behave nicely with respect to differentiation. Indeed $f(z) = z^k$ corresponds to $\tilde{f}(t) = e^{ikt}$ and then $\tilde{f}'(t) = ike^{ikt}$, so any such function is an eigenfunction for the derivative. As we shall see below, this is actually a consequence of the representation theoretic origin of the construction.

(3) Let V and W be two representations of a group G and consider the representation of $\mathcal{F}(V,W)$ as constructed above, i.e. $(g\cdot\varphi)(v) = g\cdot(\varphi(g^{-1}\cdot v))$. Then we claim that the subspace $L(V,W)$ of linear maps from V to W is invariant. Indeed, since we have given a representation of G on V, we have $g\cdot(v_1+tv_2) = g\cdot v_1 + tg\cdot v_2$ for all $g\in G$, $v_1, v_2 \in V$ and $t\in\mathbb{K}$. Applying a linear map φ, we obtain $\varphi(g\cdot v_1) + t\varphi(g\cdot v_2)$ and since G acts linearly on W, we conclude that $g\cdot\varphi$ is linear, too. Thus, from representations on V and W, we naturally get a representation on $L(V,W)$. Notice in particular that a function φ is invariant under this action if and only if $\varphi(v) = g\cdot(\varphi(g^{-1}\cdot v))$ and applying this to $g\cdot v$, we see that invariance is equivalent to $\varphi(g\cdot v) = g\cdot(\varphi(v))$. A function having this property is called *G–equivariant* or a *homomorphism of representations.*

Let us apply this construction to the case $G = GL(n,\mathbb{K})$ and $V = W = \mathbb{K}^n$, the standard representation. Then we may identify $L(\mathbb{K}^n,\mathbb{K}^n)$ with the space $M_n(\mathbb{K})$ of $n\times n$–matrices with entries from $\mathbb{K}$ via the usual action of matrices on vectors. For $A\in GL(n,\mathbb{K})$ and $B\in M_n(\mathbb{K})$ we then obtain $(A\cdot B)(x) = ABA^{-1}x$, so our representation is just by conjugation. Otherwise put, the orbit of B under this action consists of all matrix representations of the linear map $x\mapsto Bx$ with respect to arbitrary bases of $\mathbb{K}^n$. Finding a particularly nice matrix in the orbit of B is thus just the problem of finding a normal form for B. Hence describing the set of all orbits under this action is one of the main problems of linear algebra, which is solved (for $\mathbb{K} = \mathbb{C}$ and $\mathbb{R}$) by the Jordan normal form.

Passing to the Lie algebra

1.3. The Lie algebra of $GL(n,\mathbb{K})$. Let us start by looking at the group $GL(n,\mathbb{K})$. This is the group of all $n\times n$–matrices A with entries in $\mathbb{K}$ such that $\det(A)\neq 0$. Now we may view $M_n(\mathbb{K})$ as $\mathbb{K}^{n^2}$, so since the determinant function $\det:\mathbb{K}^{n^2}\to\mathbb{K}$ is continuous, the group $GL(n,\mathbb{K})$ is an open subset of $\mathbb{K}^{n^2}$. In particular, it is no problem to talk about continuity, differentiability or smoothness of functions $GL(n,\mathbb{K})\to\mathbb{K}^m$. In particular, given $A\in GL(n,\mathbb{K})$ and $B\in M_n(\mathbb{K})$ we can look at the line $t\mapsto A+tB$ for $t\in\mathbb{R}$, which lies in $GL(n,\mathbb{K})$ for small enough t. Taking $A=\mathbb{I}$, the unit matrix, we may view $M_n(\mathbb{K})$ as the space of possible derivatives at $t=0$ of smooth curves $c:\mathbb{R}\to GL(n,\mathbb{K})$ such that $c(0)=\mathbb{I}$ (the tangent space at $\mathbb{I}$).

The matrix exponential leads to a distinguished curve for each initial direction. Recall that the exponential for matrices can be defined by the usual power series $e^X = \sum_{k=0}^\infty \frac{1}{k!}X^k$, and this series converges absolutely and uniformly on compact subsets of $M_n(\mathbb{K})$ for $\mathbb{K}=\mathbb{R}$ or $\mathbb{C}$. Further, denoting by $\operatorname{tr}(X)$ the trace of $X\in M_n(\mathbb{K})$ one has $\det(e^X) = e^{\operatorname{tr}(X)}\neq 0$. This can be seen by viewing real matrices as complex matrices and using the fact that for any complex matrix X there is an invertible matrix A such that AXA^{-1} is upper triangular. Then $(AXA^{-1})^k = AX^kA^{-1}$ for all $k\in\mathbb{N}$ and thus $e^{AXA^{-1}} = Ae^XA^{-1}$. Since both the determinant and the trace are invariant under conjugation, we conclude that it suffices to show $\det(e^X) = e^{\operatorname{tr}(X)}$ if X is upper triangular. But in this case X^k is upper triangular and the entries on the main diagonal are just the kth powers of the corresponding entries of X, and thus e^X is upper triangular and the entries on the main diagonal are just the exponentials of the corresponding entries of X. Denoting the entries of X on the main diagonal by $x_1, \dots, x_n$, the trace of X is just $x_1 + \dots + x_n$, while the determinant of e^X is $e^{x_1}e^{x_2}\cdots e^{x_n} = e^{x_1+\dots+x_n}$ and the claim follows.

Consequently, we may view $X \mapsto e^X$ as a function $M_n(\mathbb{K}) \to GL(n,\mathbb{K})$. In particular, fixing $X \in M_n(\mathbb{K})$ we have the smooth curve $c_X : \mathbb{R} \to GL(n,\mathbb{K})$ defined by $c(t) = e^{tX}$. From the definition it follows easily that $c_X(0) = \mathbb{I}$ and $c'_X(0) = X$. But even more nicely, the curve c_X can be characterized as the solution of a differential equation: While $e^{X+Y} \neq e^X e^Y$ in general, it is still true that $e^{(t+s)X} = e^{tX}e^{sX}$ for $t, s \in \mathbb{R}$ and $X \in M_n(\mathbb{K})$. Using this, we may write $c'_X(t)$ as the derivative with respect to s at $s = 0$ of $s \mapsto c_X(s+t) = c_X(s)c_X(t)$, which clearly equals $Xc_X(t)$. Since the solution of a smooth first order ordinary differential equation is determined by its value in one point, we see that the curve c_X is characterized by the facts that $c'_X(t) = Xc_X(t)$ and $c_X(0) = \mathbb{I}$.

Now assume that $\varphi : GL(n,\mathbb{K}) \to GL(m,\mathbb{K})$ is a homomorphism (i.e. φ is a representation of $GL(n,\mathbb{K})$ on $\mathbb{K}^m$) which is differentiable. Take $X \in M_n(\mathbb{K})$ and consider the curve $\alpha(t) := \varphi(c_X(t)) = \varphi(e^{tX})$ in $M_m(\mathbb{K})$. Then $\alpha(s+t) = \varphi(e^{(s+t)X}) = \varphi(e^{sX}e^{tX}) = \varphi(e^{sX})\varphi(e^{tX}) = \alpha(s)\alpha(t)$. As above, we may compute the derivative $\alpha'(t)$ as $\frac{d}{ds}|_{s+0}\alpha(s+t)$ and this equals $\alpha'(0)\alpha(t)$. Now $\alpha'(0) = D\varphi(\mathbb{I})(X)$ and denoting this by $\varphi'(X)$, we see that $\alpha'(t) = \varphi'(X)\alpha(t)$ and since $\alpha(0) = \varphi(\mathbb{I}) = \mathbb{I}$ we see that $\varphi(e^{tX}) = e^{t\varphi'(X)}$ for all $t \in \mathbb{R}$ and all $X \in M_n(\mathbb{K})$.

Now the map $X \mapsto e^X$ is not surjective, but one can show that for $\mathbb{K} = \mathbb{C}$ any invertible matrix can be written as a product of finitely many matrices of the form e^X, while for $\mathbb{K} = \mathbb{R}$ the same is true for any matrix of positive determinant. Hence we conclude that for $\mathbb{K} = \mathbb{C}$ the homomorphism φ is uniquely determined by the linear map $\varphi' = D\varphi(\mathbb{I}) : M_n(\mathbb{K}) \to M_m(\mathbb{K})$ while for $\mathbb{K} = \mathbb{R}$ the same is true for the restriction of φ to the subgroup $GL^+(n,\mathbb{R})$ consisting of all matrices of positive determinant.

The remaining thing to do is to construct some structure on $M_n(\mathbb{K})$ out of the multiplication on $GL(n,\mathbb{K})$ which is preserved by the derivative of any homomorphism. The idea to do this, is to take the derivative at $\mathbb{I}$ of maps constructed from the multiplication and have the property that they fix $\mathbb{I}$. The first step is to look at the conjugation by $A \in GL(n,\mathbb{K})$, i.e. the mapping $GL(n,\mathbb{K}) \to GL(n,\mathbb{K})$ defined by $B \mapsto ABA^{-1}$. As we have seen above, $Ae^{tX}A^{-1} = e^{tAXA^{-1}}$ for all $X \in M_n(\mathbb{K})$, and applying a differentiable homomorphism φ to this equation we conclude that

$$e^{t\varphi'(AXA^{-1})} = \varphi(e^{tAXA^{-1}}) = \varphi(Ae^{tX}A^{-1}) = \varphi(A)e^{t\varphi'(X)}\varphi(A)^{-1}.$$

Differentiating at $t = 0$ gives us $\varphi'(AXA^{-1}) = \varphi(A)\varphi'(X)\varphi(A)^{-1}$.

Now we are ready to get a structure with the required properties. Namely, for $X, Y \in M_n(\mathbb{K})$ we define the *Lie bracket* $[X, Y]$ to be the derivative at 0 of the curve $t \mapsto e^{tX}Y(e^{tX})^{-1}$. From above, we see that for a differentiable homomorphism φ we get

$$\varphi'(e^{tX}Y(e^{tX})^{-1}) = \varphi(e^{tX})\varphi'(Y)\varphi(e^{tX})^{-1} = e^{t\varphi'(X)}\varphi'(Y)(e^{t\varphi'(X)})^{-1}.$$

Taking derivatives at zero the left hand side simply gives $\varphi'([X,Y])$ since φ' is linear, while on the right hand side we get $[\varphi'(X), \varphi'(Y)]$ by definition. While this definition may look a bit complicated, the bracket $[X, Y]$ is actually a very simple object. To see this, we only have to note that the derivative at $t = 0$ of the curve $e^{tX}Y(e^{tX})^{-1}$ depends only on the derivatives at $t = 0$ of e^{tX} and $(e^{tX})^{-1}$. For the first curve, this is just X, while for the second one we note that $e^{tX}e^{-tX} = e^0 = \mathbb{I}$ implies that $(e^{tX})^{-1} = e^{-tX}$, and hence has derivative $-X$ in zero. Hence we may replace the curve $e^{tX}Y(e^{tX})^{-1}$

by $(\mathbb{I} + tX)Y(\mathbb{I} - tX) = Y + tXY - tYX + t^2XYX$, from which we conclude that $[X, Y] = XY - YX$, so the bracket is simply the commutator of matrices.

From this last formula, one easily reads off three properties of the bracket, which we will use as the defining properties of a Lie algebra. First the bracket $[\ ,\ ]$ is defined on the vector space $M_n(\mathbb{K})$ and obviously $[Y, X] = -[X, Y]$, so it is skew symmetric. Further

$$\begin{aligned}[X_1 + tX_2, Y] &= (X_1 + tX_2)Y - Y(X_1 + tX_2) = X_1Y + tX_2Y - YX_1 - tYX_2 \\ &= [X_1, Y] + t[X_2, Y],\end{aligned}$$

so the bracket is linear in the first variable, and hence by skew symmetry it is bilinear. The last property, called the Jacobi–identity is a bit less obvious:

$$\begin{aligned}[X, [Y, Z]] &= [X, YZ - ZY] = XYZ - XZY - YZX + ZYX \\ &= [X, Y]Z + YXZ - [X, Z]Y - ZXY + Y[X, Z] - YXZ - Z[X, Y] + ZXY \\ &= [[X, Y], Z] + [Y, [X, Z]].\end{aligned}$$

The space $M_n(\mathbb{K})$ together with the bracket $[X, Y] = XY - YX$ is called the Lie algebra of $GL(n, \mathbb{K})$ and it is denoted by $\mathfrak{gl}(n, \mathbb{K})$.

1.4. Matrix groups and their Lie algebras. What we have done for $GL(n, \mathbb{K})$ above generalizes to appropriate subgroups. Let $G \subset GL(n, \mathbb{K})$ be a subgroup which at the same time is a closed subspace of the topological space $GL(n, \mathbb{K})$. Then one may look at smooth curves $c : \mathbb{R} \to GL(n, \mathbb{K})$ which have values in G and satisfy $c(0) = \mathbb{I}$, and define $\mathfrak{g} \subset M_n(\mathbb{K})$ to be the set of all derivatives at $t = 0$ of such curves. It turns out that $\mathfrak{g}$ is a linear subspace of $M_n(\mathbb{K})$ and a matrix X lies in $\mathfrak{g}$ if and only if $e^{tX} \in G$ for all $t \in \mathbb{R}$. Now for $A \in G$ and $X \in \mathfrak{g}$ the curve $Ae^{tX}A^{-1}$ has values in G, so its derivative in zero lies in $\mathfrak{g}$, i.e. $AXA^{-1} \in \mathfrak{g}$ for all $A \in G$ and $X \in \mathfrak{g}$. Consequently, for $X, Y \in \mathfrak{g}$, the curve $e^{tX}Ye^{-tX}$ has values in $\mathfrak{g}$, and differentiating in $t = 0$, we conclude that $[X, Y] \in \mathfrak{g}$ for all $X, Y \in \mathfrak{g}$, i.e. $\mathfrak{g}$ is a *Lie subalgebra* of $M_n(\mathbb{R})$.

Now we can proceed very similarly as in 1.3 above. Suppose that $G \subset GL(n, \mathbb{K})$ and $H \subset GL(m, \mathbb{K})$ are closed subgroups and $\varphi : G \to H$ is a differentiable homomorphism. (The assumptions imply that G and H are smooth submanifolds of $\mathbb{K}^{n^2}$ respectively $\mathbb{K}^{m^2}$, so there is a well defined notion of differentiability. However, one may simply think of differentiability being defined by the fact that for any $X \in \mathfrak{g}$ the curve $t \mapsto \varphi(e^{tX})$ is a differentiable curve in $\mathbb{K}^{m^2}$.) Since φ maps $\mathbb{I} \in G$ to $\mathbb{I} \in H$ we can form the derivative $\varphi' := D\varphi(\mathbb{I})$, which one may also think about being defined by $\varphi'(X) = \frac{d}{dt}|_{t=0}\varphi(e^{tX})$. As before, one concludes that $\varphi(e^{tX}) = e^{t\varphi'(X)}$ and further that $\varphi'([X, Y]) = [\varphi'(X), \varphi'(Y)]$, i.e. φ' is a *Lie algebra homomorphism.* If G is connected (as a topological space), one can show that any element of G can be written as a finite product of exponentials of elements of $\mathfrak{g}$, hence φ is uniquely determined by φ' provided that G is connected.

Examples. (1) Consider $SL(n, \mathbb{K}) := \{A \in GL(n, \mathbb{K}) : \det(A) = 1\}$. Since the determinant function is continuous, this is a closed subset of $GL(n, \mathbb{K})$ and since $\det(AB) = \det(A)\det(B)$ it is a subgroup. From above we know that $\det(e^A) = e^{\mathrm{tr}(A)}$, so the Lie algebra of this group is $\mathfrak{sl}(n, \mathbb{K}) = \{X \in M_n(\mathbb{K}) : \mathrm{tr}(X) = 0\}$, the subspace of tracefree matrices. The fact that this space is closed under the commutator (which we know in general) here can be easily seen directly since $\mathrm{tr}(XY - YX) = \mathrm{tr}(XY) - \mathrm{tr}(YX)$ and $\mathrm{tr}(YX) = \mathrm{tr}(XY)$ holds in general. So we even see that the bracket of any two matrices lies in $\mathfrak{sl}(n, \mathbb{K})$.

(2) Consider $U(1) \subset GL(1,\mathbb{C}) = \mathbb{C} \setminus \{0\}$. For $z = x + iy \in M_1(\mathbb{C}) = \mathbb{C}$, one has $e^z = e^x e^{iy}$, so $|e^z| = e^x$ and thus the Lie algebra $\mathfrak{u}(1)$ is given by $i\mathbb{R} \subset \mathbb{C}$. Hence the map $t \mapsto e^{it}$ that we have used to convert functions f on $U(1)$ to functions $\tilde{f}$ on $\mathbb{R}$ comes from the Lie algebra approach. Now we can see the representation theory interpretation of the fact that differentiable functions $f : U(1) \to \mathbb{C}$ such that $z \cdot f = \alpha(z)f$ for all $z \in U(1)$ must be eigenfunctions for the derivative: Taking the basic element $i \in i\mathbb{R}$ we are led to looking at the curve $t \mapsto e^{it}$ in $U(1)$. If $f(z^{-1}w) = \alpha(z)f(w)$, then $f(e^{-it}w) = \alpha(e^{it})f(w)$. In the notation used in example (2) of 1.2, this can be written as $\tilde{f}(s-t) = \tilde{\alpha}(t)\tilde{f}(s)$, and differentiating with respect to t we obtain $-\tilde{f}'(s) = \alpha'(0)\tilde{f}(s)$ (provided that α is differentiable at 0).

(3) Consider the orthogonal group $O(n) = \{A \in M_n(\mathbb{R}) : A^T = A^{-1}\}$. This is a closed subset of $GL(n,\mathbb{R})$ since transposition and inversion of matrices are continuous mappings, and it is a subgroup since $(AB)^T = B^T A^T$ and $(AB)^{-1} = B^{-1}A^{-1}$. To obtain the Lie algebra $\mathfrak{o}(n)$ we have to look at the equation $(e^{sX})^T = e^{-sX}$. Since transposition is linear, the derivative of the left hand side at $s = 0$ is simply X^T, so we obtain $X^T = -X$, i.e. the matrix X has to be skew symmetric. One can show that this is the only condition i.e. $\mathfrak{o}(n) = \{X \in M_n(\mathbb{R}) : X^T = -X\}$. It is a nice simple exercise to verify explicitly that $\mathfrak{o}(n)$ is closed under the commutator of matrices.

From the defining equation of $O(n)$ it follows immediately that any orthogonal matrix has determinant ± 1 and there are orthogonal maps having determinant -1, for example the reflection in a hyperplane. Putting $SO(n) := \{A \in O(n) : \det(A) = 1\}$ we see that this is a closed subgroup of $GL(n,\mathbb{R})$, since it is the intersection of the closed subgroups $O(n)$ and $SL(n,\mathbb{R})$. Moreover, since the determinant of e^{tX} is always positive, we see that the Lie algebra $\mathfrak{so}(n)$ coincides with $\mathfrak{o}(n)$.

(4) Let $B(n,\mathbb{K})$ be the set of upper triangular invertible matrices. We can write $B(n,\mathbb{K}) = \{A = (a_{ij}) \in GL(n,\mathbb{K}) : a_{ij} = 0 \quad \forall j < i\}$, which shows that $B(n,\mathbb{K})$ is a closed subset of $GL(n,\mathbb{K})$ and it clearly is a subgroup. We have already noticed in 1.3 that for an upper triangular matrix X also e^X is upper triangular, so the Lie algebra $\mathfrak{b}(n,\mathbb{K})$ contains all upper triangular matrices. On the other hand, suppose that X is any matrix such that e^{tX} is upper triangular for all $t \in \mathbb{R}$. Denoting by $\{e_1, \dots, e_n\}$ the standard basis of $\mathbb{K}^n$, we see that $e^{tX}(e_1)$ must be some multiple of e_1 for all t, i.e. $e^{tX}(e_1) = a(t)e_1$. Differentiating at $t = 0$ we see that $X(e_1) = a'(0)e_1$, so the only nonzero element in the first column of X is x_{11}. Next, $e^{tX}(e_2) = a(t)e_1 + b(t)e_2$, and differentiating we see that in the second column of X only the two topmost entries maybe non–zero. Iterating this argument we conclude that X is upper triangular and thus $\mathfrak{b}(n,\mathbb{K}) = \{X = (x_{ij}) \in M_n(\mathbb{K}) : x_{ij} = 0 \quad \forall j < i\}$.

Similarly, we may consider the subgroup $N(n,\mathbb{K})$ of $B(n,\mathbb{K})$ consisting of those upper triangular matrices whose diagonal entries are equal to one. Similarly as above one verifies that the Lie algebra of this group is the algebra $\mathfrak{n}(n,\mathbb{K}) = \{X = (x_{ij}) \in M_n(\mathbb{K}) : x_{ij} = 0 \quad \forall j \leq i\}$ of strictly upper triangular matrices.

A primer on the Lie group – Lie algebra correspondence

1.5. General Lie groups and their Lie algebras. We next quickly review some facts on general Lie groups. Unexplained terms (like “left invariant vector fields”) are only intended for those who have already heard about general Lie groups, to relate what they now to the developments sketched here. More detailed information on these subjects and the background from differential geometry can be found in the book [**Kolař-Michor-Slovák**] or in the lecture notes [**Michor**].

Extending analysis to more complicated objects than just open subsets in $\mathbb{R}^n$ one is lead to the notion of a smooth manifold. Essentially, a smooth manifold M is a topological space which locally around each point looks like an open subset of $\mathbb{R}^n$ for some fixed n (called the dimension of the manifold). Since differentiation is a local concept, one can define the notion of differentiable or smooth mappings between smooth manifolds. Suppose that G is a smooth manifold which carries a group structure. Then we can consider the multiplication as a map $G \times G \to G$ and the inversion as a map $G \to G$. The product of two smooth manifolds is canonically a smooth manifold, so it makes sense to require that the multiplication is smooth as a map $G \times G \to G$. It can then be shown that also the inversion must be smooth and G is called a *Lie group*.

To any point x in an n–dimensional manifold M one can associate an n–dimensional vector space T_xM, the tangent space of M at x. Derivatives of differentiable mappings are defined on these tangent spaces, i.e. for a smooth map $f : M \to N$ and a point $x \in M$ one has the derivative (or tangent map) of f in x, which is a linear map $T_xf : T_xM \to T_{f(x)}N$. For a Lie group G, the Lie algebra $\mathfrak{g}$ is defined to be the tangent space T_eG of G at the neutral element e of the group G. Usually, one then defines a Lie bracket on $\mathfrak{g}$ directly by showing that $\mathfrak{g}$ may be identified with the space of left invariant vector fields on G and use the restriction of the Lie bracket of vector fields. It is however also possible to follow the construction we have used in the case of matrix groups, an we briefly sketch this:

One can define the exponential map $\exp : \mathfrak{g} \to G$ via flow lines through e of left invariant vector fields, which means that for $X \in \mathfrak{g}$ the curve $t \mapsto \exp(tX)$ is characterized by the analog of the differential equation for the matrix exponential that we have derived in 1.3. Having done this, one can associate to $g \in G$ a linear map $\mathrm{Ad}(g) : \mathfrak{g} \to \mathfrak{g}$ by defining $\mathrm{Ad}(g)(X)$ to be the derivative at $t = 0$ of the curve $t \mapsto g\exp(tX)g^{-1}$. This is called the adjoint action of g and one shows that this defines a smooth homomorphism $\mathrm{Ad} : G \to GL(\mathfrak{g})$, i.e. a representation of G on $\mathfrak{g}$, called the *adjoint representation*. Then one can proceed as in 1.3 by defining $[X, Y]$ to be the derivative at $t = 0$ of the curve $\mathrm{Ad}(\exp(tX))(Y)$. One shows that this operation makes $\mathfrak{g}$ into a Lie algebra, i.e. it is bilinear, skew symmetric and satisfies the Jacobi identity.

Now assume that G and H are Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ and that $\varphi : G \to H$ is a differentiable homomorphism. Then φ maps the neutral element $e \in G$ to $e \in H$, so we can interpret the derivative of φ in e as a linear map $T_e\varphi =: \varphi' : \mathfrak{g} \to \mathfrak{h}$. Using the characterization of the exponential maps via differential equations one concludes similarly as in 1.3 that $\varphi \circ \exp^G = \exp^H \circ \varphi'$ which then implies $\varphi'(\mathrm{Ad}(g)(X)) = \mathrm{Ad}(\varphi(g))(\varphi'(X))$ and further $\varphi'([X, Y]) = [\varphi'(X), \varphi'(Y)]$. Thus the derivatives of differentiable group homomorphisms are Lie algebra homomorphisms. If the Lie group G is connected, then one shows that any element of G can be written as a finite product of exponentials, which shows that φ is uniquely determined by φ' provided that G is connected. Using the compatibility of homomorphisms with the exponential map, one also shows that a differentiable homomorphism of Lie groups is automatically smooth, so there is no need to distinguish between the two cases.

A (finite dimensional) representation of a Lie group G is defined as a smooth homomorphism $\varphi : G \to GL(V)$, where V is a finite dimensional vector space. Equivalently, one may view this as an action $\varphi : G \times V \to V$ which is a smooth map and linear in the second variable. Choosing a basis in V, we may identify it with $\mathbb{K}^n$ and view φ as a homomorphism to $GL(n, \mathbb{K})$.

Similarly, a finite dimensional representation of a Lie algebra $\mathfrak{g}$ is defined as a Lie algebra homomorphism $\psi : \mathfrak{g} \to \mathfrak{gl}(V)$. As above, we may restrict to the case $V = \mathbb{R}^n$

and thus $\mathfrak{gl}(V) = M_n(\mathbb{K})$. The condition that ψ is a Lie algebra homomorphism means that $\psi([X, Y]) = \psi(X)\psi(Y) - \psi(Y)\psi(X)$ since the bracket on $M_n(\mathbb{R})$ is the commutator of matrices. In the picture analogous to an action, this means that we need a bilinear map $\mathfrak{g} \times V \to V$ which we write as $(X, v) \mapsto X \cdot v$ such that $[X, Y] \cdot v = X \cdot (Y \cdot v) - Y \cdot (X \cdot v)$.

Now the above results imply that for a representation $\varphi : G \to GL(n, \mathbb{K})$ the derivative $\varphi' : \mathfrak{g} \to \mathfrak{gl}(n, \mathbb{K})$ is a representation of $\mathfrak{g}$. If there is no risk of confusion, we will denote both the actions of G and of $\mathfrak{g}$ on $V = \mathbb{R}^n$ simply by dots. Also we conclude from above that if G is connected, then φ is determined uniquely by φ'. But indeed we can do quite a bit more: compatibility of φ with the exponential mappings implies that $\varphi(\exp(X)) = e^{\varphi'(X)}$, which means that $\exp(X) \cdot v = v + X \cdot v + \frac{1}{2} X \cdot X \cdot v + \dots$. In particular, let us assume that $W \subset V$ is a linear subspace. For $X \in \mathfrak{g}$, $t \in \mathbb{R}$ and $w \in W$ we then have

$$\exp(tX) \cdot w = w + t(X \cdot w) + t^2(X \cdot X \cdot w) + \dots .$$

If W is G–invariant, then the right hand side lies in W for all t and differentiating the left hand side at $t = 0$, we see that $X \cdot w \in W$, for all $X \in \mathfrak{g}$, and thus W is $\mathfrak{g}$–invariant. Conversely, if W is $\mathfrak{g}$–invariant then $X \cdot w \in W$ and thus $X \cdot X \cdot w$ in W and so on, so we conclude that for each $t \in \mathbb{R}$ the right hand side of the above equation lies in W. (This uses that linear subspaces of finite dimensional vector spaces are automatically closed.) Hence $\exp(X) \cdot w \in W$ for all $X \in \mathfrak{g}$ and if G is connected, this implies that W is G–invariant. Thus we see that for a representation of connected group G on V the question of existence of invariant subspaces (and thus the questions of irreducibility and indecomposability) can be studied by looking at the corresponding Lie algebra representation only.

1.6. A few more facts from Lie theory. To conclude this introduction, we collect some more results on the relation between Lie groups and Lie algebras. The proofs of several of these results are a bit more involved and need more input from differential geometry, so we do not even sketch them. The main reason for including these results is to show that many questions on Lie groups can be reduced to questions on Lie algebras, and that any Lie algebra shows up as the Lie algebra of a Lie group.

We have seen that the Lie group – Lie algebra correspondence works best for connected groups. For most problems, connectedness of the group is however not a serious restriction. If G is any Lie group, then it is easy to show that the connected component G_0 of the neutral element is a normal subgroup in G, so the quotient G/G_0 is a discrete group. In many applications this quotients are very small and it is often easy to pass from information on G_0 to information on G.

The passage from $GL(n, \mathbb{K})$ to arbitrary matrix groups in 1.4 has an analog for general Lie groups: If G is a Lie group and $H \subset G$ is a subgroup which is a closed subset of G, then H turns out to be a smooth submanifold and hence a Lie group. The Lie algebra $\mathfrak{h} = T_eH$ of H is naturally included into $\mathfrak{g} = T_eG$ as a Lie subalgebra. It can be explicitly described as the set of derivatives of smooth curves that have values in H or as the set of all $X \in \mathfrak{g}$ such that $\exp(tX) \in H$ for all $t \in \mathbb{R}$.

Conversely, let us assume that G is a Lie group with Lie algebra $\mathfrak{g}$ and that $\mathfrak{h} \subset \mathfrak{g}$ is a Lie subalgebra (i.e. a linear subspace such that $[X, Y] \in \mathfrak{h}$ for all $X, Y \in \mathfrak{h}$). Then one shows that there is a connected Lie group H with Lie algebra $\mathfrak{h}$ and an injective smooth homomorphism $i : H \to G$ such that $i' : \mathfrak{h} \to \mathfrak{g}$ is the inclusion. The homomorphism i has injective differential in each point of H, and if $i(H)$ is a closed subset of G, then $i : H \to i(H)$ is an isomorphism of Lie groups (i.e. a smooth bijective homomorphism, whose inverse is smooth, too). Hence understanding connected subgroups of G is equivalent to understanding Lie subalgebras of $\mathfrak{g}$.

Next, it turns out that any finite dimensional (abstract) Lie algebra $\mathfrak{g}$ can be viewed as the Lie algebra of a Lie group, which may even be chosen to be a subgroup of $GL(n,\mathbb{R})$ (for sufficiently large n). This is done by showing that any finite dimensional Lie algebra is isomorphic to a Lie algebra of matrices, i.e. a Lie subalgebra of $\mathfrak{gl}(n,\mathbb{R})$ for sufficiently large n. Then the result above implies the existence of a Lie group G with Lie algebra $\mathfrak{g}$ together with an injective homomorphism $i : G \to GL(n,\mathbb{R})$. Therefore, to study the Lie algebras of Lie groups, one has to study all finite dimensional Lie algebras.

Finally, we want to discuss the question of existence of homomorphisms between Lie groups. There is a topological restriction to be taken into account. Suppose that G and H are connected Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ and that $f : \mathfrak{g} \to \mathfrak{h}$ is a homomorphism of Lie algebras. Then one shows that there is a smooth map $\varphi : U \to H$, where $U \subset G$ is an open subset containing e such that $\varphi(e) = e$ and $\varphi(g_1g_2) = \varphi(g_1)\varphi(g_2)$ whenever g_1, g_2 and g_1g_2 all lie in U with derivative $\varphi' = T_e\varphi = f$. If G is simply connected (a topological property) then there even exists a smooth homomorphism $\varphi : G \to H$ with $\varphi' = f$. If G is not simply connected, then one can construct a Lie group $\tilde{G}$ (called the universal covering of G) which also has Lie algebra $\mathfrak{g}$ together with a surjective homomorphism $p : \tilde{G} \to G$ such that p' is the identity. The kernel of p is a discrete normal subgroup of $\tilde{G}$. Hence one obtains a homomorphism $\tilde{\varphi} : \tilde{G} \to H$ such that $\tilde{\varphi}' = f$. To check whether there is a homomorphism $\varphi : G \to H$ with $\varphi' = f$ one then only has to check whether $\tilde{\varphi}$ vanishes on the kernel of p.

Applying the last result to the case $H = GL(n,\mathbb{K})$ we see in particular that for a connected and simply connected group G with Lie algebra $\mathfrak{g}$ representations of G are in bijective correspondence with representations of $\mathfrak{g}$. This applies for example to $G = GL(n,\mathbb{C})$ and $G = SL(n,\mathbb{C})$.

Finally, I want to mention a result that shows that the multiplication on a connected Lie group G is encoded in the Lie bracket of the Lie algebra $\mathfrak{g}$: There is a universal expression CH, called the Campbell–Hausdorff series, which only uses iterated Lie brackets, such that $\exp(X)\exp(Y) = \exp(CH(X,Y))$ for X, Y close enough to zero. There is a (fairly complicated) explicit formula for this series, the first few terms are:

$$CH(X,Y) = X + Y + \tfrac{1}{2}[X,Y] + \tfrac{1}{12}([X,[X,Y]] - [Y,[Y,X]]) + \dots$$

CHAPTER 2

General theory of Lie algebras

Having clarified why it is useful to study Lie algebras and their representations, we can now start to develop the general theory from scratch. The first step is to identify several basic classes of Lie algebras, the two most important of which are solvable and semisimple Lie algebras.

Basic classes of Lie algebras

2.1. Solvable and nilpotent Lie algebras. While it is possible to define and study Lie algebras over arbitrary fields, we restrict to the cases of real and complex Lie algebras.

Definition. (1) A *Lie algebra* over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ is a $\mathbb{K}$–vector space $\mathfrak{g}$ together with a bilinear map $[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, called the *Lie bracket* of $\mathfrak{g}$, which is skew symmetric, i.e. $[Y, X] = -[X, Y]$ for all $X, Y \in \mathfrak{g}$ and satisfies the Jacobi identity, i.e. $[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]]$ for all $X, Y, Z \in \mathfrak{g}$.
(2) Let $(\mathfrak{g}, [\ ,\ ])$ be a Lie algebra. A *Lie subalgebra* of $\mathfrak{g}$ is a linear subspace $\mathfrak{h} \subset \mathfrak{g}$ which is closed under the Lie bracket, i.e. such that $[X, Y] \in \mathfrak{h}$ for all $X, Y \in \mathfrak{h}$. We write $\mathfrak{h} \leq \mathfrak{g}$ is $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$. Of course, in this case $(\mathfrak{h}, [\ ,\ ])$ is a Lie algebra, too.
(3) If $\mathfrak{g}$ and $\mathfrak{h}$ are Lie algebras then a homomorphism $\varphi : \mathfrak{g} \to \mathfrak{h}$ of Lie algebras is a linear mapping which is compatible with the brackets, i.e. such that $[\varphi(X), \varphi(Y)] = \varphi([X, Y])$ for all $X, Y \in \mathfrak{g}$.
(4) An *isomorphism* of Lie algebras is a bijective homomorphism. It is an easy exercise to show that the inverse of such a homomorphism is a homomorphism, too. If there is an isomorphism $\varphi : \mathfrak{g} \to \mathfrak{h}$, then $\mathfrak{g}$ and $\mathfrak{h}$ are called *isomorphic* and we write $\mathfrak{g} \cong \mathfrak{h}$.

Examples. (0) If V is any $\mathbb{K}$–vector space, then we can define the bracket to be identically zero, thus making V into a Lie algebra. Such Lie algebras are called *Abelian.*

(1) From 1.3 we know the example of the Lie algebra $\mathfrak{gl}(n, \mathbb{K}) = M_n(\mathbb{K})$ of $n \times n$–matrices with entries from $\mathbb{K}$ with the commutator of matrices as the Lie bracket. In 1.4 we have seen several examples of Lie subalgebras in $\mathfrak{gl}(n, \mathbb{K})$, like $\mathfrak{sl}(n, \mathbb{K})$ and $\mathfrak{so}(n) \subset \mathfrak{gl}(n, \mathbb{R})$.

As a slight variation, we may look at an arbitrary $\mathbb{K}$–vector space V and define $\mathfrak{gl}(V)$ as the space $L(V, V)$ of linear maps from V to itself with the pointwise vector space operations and the bracket defined by the commutator of linear maps, i.e. $[\varphi, \psi] = \varphi \circ \psi - \psi \circ \varphi$. The proof that this satisfies the Jacobi identity is exactly as in the case of matrices in 1.3
(2) Direct sums: If $\mathfrak{g}$ and $\mathfrak{h}$ are Lie algebras, then the direct sum $\mathfrak{g} \oplus \mathfrak{h}$ is just the vector space direct sum (i.e. $\mathfrak{g} \times \mathfrak{h}$ with the component wise vector space operations) together with the component wise Lie bracket, i.e. $[(X_1, Y_1), (X_2, Y_2)] := ([X_1, X_2], [Y_1, Y_2])$. Verifying that this is again a Lie algebra is a trivial exercise.
(3) Low dimensional examples: If $\dim(\mathfrak{g}) = 1$, then choosing a nonzero element $X \in \mathfrak{g}$, we can write any element of $\mathfrak{g}$ as aX for some $a \in \mathbb{K}$. Then $[aX, bX] = ab[X, X]$ by

bilinearity of the bracket and $[X, X] = -[X, X]$ and thus $[X, X] = 0$ by skew symmetry. Hence any one–dimensional Lie algebra is Abelian.

For $\dim(\mathfrak{g}) = 2$, there is the possibility to have a non–trivial Lie bracket. Consider $\mathfrak{a}(1, \mathbb{K}) := \left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \right\} \subset \mathfrak{gl}(2, \mathbb{K})$. One immediately verifies that

$$\left[\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} a' & b' \\ 0 & 0 \end{pmatrix} \right] = \begin{pmatrix} 0 & ab' - a'b \\ 0 & 0 \end{pmatrix},$$

so this is a Lie subalgebra with non–trivial bracket. Taking the basis consisting of $X, Y \in \mathfrak{a}(1, \mathbb{K})$, where X corresponds to $a = 1$ and $b = 0$, while Y corresponds to $a = 0$ and $b = 1$, we see that $[X, Y] = Y$.

Indeed, we can easily verify that any two dimensional $\mathbb{K}$–Lie algebra $\mathfrak{g}$ with nonzero bracket is isomorphic to $\mathfrak{a}(1, \mathbb{K})$: Taking any basis $\{v, w\}$ for $\mathfrak{g}$ we see that

$$[av + bw, cv + dw] = ac[v, v] + bc[w, v] + ad[v, w] + bd[w, w]$$

by bilinearity. Skew symmetry gives $[v, v] = [w, w] = 0$ and $[w, v] = -[v, w]$, so this reduces to $(ad - bc)[v, w]$. Since the bracket is nonzero we see that $Y' := [v, w] \neq 0$. Choosing a linearly independent vector X' we see that $[X', Y']$ must be a nonzero multiple of Y', so replacing X' by an appropriate multiple we get $[X', Y'] = Y'$. Mapping X' to X and Y' to Y defines a linear isomorphism $\varphi : \mathfrak{g} \to \mathfrak{a}(1, \mathbb{K})$. From above we however conclude that $[aX' + bY', cX' + dY'] = (ad - bc)Y'$ and the same equation with X and Y instead of X' and Y', so φ is a Lie algebra homomorphism.

If $(\mathfrak{g}, [\ ,\ ])$ is a Lie algebra and $A, B \subset \mathfrak{g}$ are nonempty subsets, then we denote by $[A, B]$ the vector space generated by all all elements of the form $[a, b]$ with $a \in A$ and $b \in B$. In this notation, the definition of a Lie subalgebra simply reads as a subspace $\mathfrak{h} \subset \mathfrak{g}$ such that $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$. Clearly, the intersection of an arbitrary family of subalgebras of $\mathfrak{g}$ is again a subalgebra, so for any subset $A \subset \mathfrak{g}$, there is a smallest subalgebra of $\mathfrak{g}$ which contains A, called the *subalgebra generated by* A.

To form quotients of Lie algebras, one needs a strengthening of the notion of a subalgebra. We say that a subalgebra $\mathfrak{h} \leq \mathfrak{g}$ is an *ideal* in $\mathfrak{g}$ and write $\mathfrak{h} \triangleleft \mathfrak{g}$ if $[\mathfrak{g}, \mathfrak{h}] \subset \mathfrak{h}$, i.e. $[X, Y] \in \mathfrak{h}$ for $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$. If $\mathfrak{h}$ is an ideal in $\mathfrak{g}$, we can look at the quotient space $\mathfrak{g}/\mathfrak{h} = \{X + \mathfrak{h} : X \in \mathfrak{g}\}$. For $X, Y \in \mathfrak{g}$ and $H, K \in \mathfrak{h}$ we have $[X + H, Y + K] = [X, Y] + [X, K] + [H, Y] + [H, K]$, and since $\mathfrak{h}$ is an ideal, the last three summands lie in $\mathfrak{h}$. Hence we see that $[X + \mathfrak{h}, Y + \mathfrak{h}] := [X, Y] + \mathfrak{h}$ is a well defined skew symmetric bilinear map $(\mathfrak{g}/\mathfrak{h}) \times (\mathfrak{g}/\mathfrak{h}) \to (\mathfrak{g}/\mathfrak{h})$, and it is clear that this also satisfies the Jacobi–identity. This is called the quotient of $\mathfrak{g}$ by the ideal $\mathfrak{h}$. By construction, the natural map $\pi : \mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$ defined by $\pi(X) := X + \mathfrak{h}$ is a homomorphism of Lie algebras. As for subalgebras, the intersection of any family of ideals in $\mathfrak{g}$ is again an ideal in $\mathfrak{g}$, for any subset $A \subset \mathfrak{g}$, there is the *ideal generated by* A.

Let $\mathfrak{g}$ and $\mathfrak{h}$ be Lie algebras and $\varphi : \mathfrak{g} \to \mathfrak{h}$ a homomorphism. Then the image $\mathrm{im}(\varphi)$ is a linear subspace of $\mathfrak{h}$, while the kernel $\ker(\varphi)$ is a linear subspace of $\mathfrak{g}$. Now the equation $[\varphi(X), \varphi(Y)] = \varphi([X, Y])$ implies that $\mathrm{im}(\varphi)$ is closed under the bracket and thus a subalgebra of $\mathfrak{h}$, and on the other hand $\varphi([X, Y]) = 0$ for $X \in \ker(\varphi)$ and $Y \in \mathfrak{g}$, so $\ker(\varphi)$ is even an ideal in $\mathfrak{g}$.

Next, for an arbitrary Lie algebra $\mathfrak{g}$ consider the subspace $[\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{g}$. Since for $X, Y \in \mathfrak{g}$ we by definition have $[X, Y] \in [\mathfrak{g}, \mathfrak{g}]$ this is an ideal in $\mathfrak{g}$ and the induced bracket on $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is identically zero. Indeed, this is the largest Abelian quotient of $\mathfrak{g}$, since any homomorphism from $\mathfrak{g}$ to an Abelian Lie algebra must by definition vanish on $[\mathfrak{g},\mathfrak{g}]$ and thus factor over this quotient. The ideal $[\mathfrak{g},\mathfrak{g}]\lhd\mathfrak{g}$ is called the *commutator ideal*. This idea can now be extended to two sequences of ideals in $\mathfrak{g}$.

Before we do this, let us note one more fact: Suppose that $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are two ideals in $\mathfrak{g}$. For $X\in\mathfrak{g}$, $H_1\in\mathfrak{h}_1$ and $H_2\in\mathfrak{h}_2$ we have $[X,[H_1,H_2]]=[[X,H_1],H_2]+[H_1,[X,H_2]]$ by the Jacobi identity. Since $\mathfrak{h}_1$ is an ideal we have $[X,H_1]\in\mathfrak{h}_1$ and thus the first summand lies in $[\mathfrak{h}_1,\mathfrak{h}_2]$. Similarly, the second summand lies in $[\mathfrak{h}_1,\mathfrak{h}_2]$, which implies that $[\mathfrak{h}_1,\mathfrak{h}_2]$ is an ideal in $\mathfrak{g}$, which by construction is contained in $\mathfrak{h}_1\cap\mathfrak{h}_2$.

Now let us define $\mathfrak{g}^1=\mathfrak{g}$, $\mathfrak{g}^2=[\mathfrak{g},\mathfrak{g}]$ and inductively $\mathfrak{g}^{k+1}=[\mathfrak{g},\mathfrak{g}^k]$. Inductively, we see that each $\mathfrak{g}^k$ is an ideal in $\mathfrak{g}$ and $\mathfrak{g}^{k+1}\subset\mathfrak{g}^k$. The sequence $\mathfrak{g}\supset\mathfrak{g}^2\supset\dots\supset\mathfrak{g}^k\supset\mathfrak{g}^{k+1}\supset\dots$ is called the *lower central series* of $\mathfrak{g}$. The Lie algebra $\mathfrak{g}$ is called *nilpotent* if $\mathfrak{g}^k=0$ for some $k\in\mathbb{N}$. Nilpotency simply means that applying sufficiently many brackets one always ends up with zero.

On the other hand, we define $\mathfrak{g}^{(1)}=\mathfrak{g}$ and inductively, $\mathfrak{g}^{(k+1)}:=[\mathfrak{g}^{(k)},\mathfrak{g}^{(k)}]$. From above, we again see inductively that each $\mathfrak{g}^{(k)}$ is an ideal in $\mathfrak{g}$, and clearly we have $\mathfrak{g}^{(k+1)}\subset\mathfrak{g}^{(k)}$. Hence we get another decreasing sequence $\mathfrak{g}\supset\mathfrak{g}^{(2)}\supset\dots\supset\mathfrak{g}^{(k)}\supset\mathfrak{g}^{(k+1)}\supset\dots$, which is called the *derived series* of $\mathfrak{g}$. The Lie algebra $\mathfrak{g}$ is called *solvable* if $\mathfrak{g}^{(k)}=0$ for some $k\in\mathbb{N}$. By construction, we have $\mathfrak{g}^2=\mathfrak{g}^{(2)}\subset\mathfrak{g}^1=\mathfrak{g}^{(1)}$, which inductively implies $\mathfrak{g}^{(k)}\subset\mathfrak{g}^k$. In particular, any nilpotent Lie algebra is solvable.

Suppose that $\mathfrak{h}\le\mathfrak{g}$ is a subalgebra. Then clearly $\mathfrak{h}^k\subset\mathfrak{g}^k$ and $\mathfrak{h}^{(k)}\subset\mathfrak{g}^{(k)}$, so if $\mathfrak{g}$ is nilpotent (respectively solvable), then also $\mathfrak{h}$ is nilpotent (respectively solvable). Similarly, if $\mathfrak{h}\lhd\mathfrak{g}$ is any ideal, then the canonical homomorphism $\mathfrak{g}\to\mathfrak{g}/\mathfrak{h}$ induces surjections $\mathfrak{g}^k\to(\mathfrak{g}/\mathfrak{h})^k$ and $\mathfrak{g}^{(k)}\to(\mathfrak{g}/\mathfrak{h})^{(k)}$. Consequently, quotients of nilpotent (solvable) Lie algebras are nilpotent (solvable).

There is a nice characterization of solvable Lie algebras in terms of extensions which also implies a converse to these results. Suppose that $\mathfrak{g}$ is a Lie algebra and we have a finite sequence $\mathfrak{g}\supset\mathfrak{g}_1\supset\dots\supset\mathfrak{g}_{k-1}\supset\mathfrak{g}_k=\{0\}$ such that each $\mathfrak{g}_{j+1}$ is an ideal in $\mathfrak{g}_j$ such that the quotient $\mathfrak{g}_j/\mathfrak{g}_{j+1}$ is Abelian. By the universal property of the commutator ideal, since $\mathfrak{g}/\mathfrak{g}_1$ is Abelian we must have $[\mathfrak{g},\mathfrak{g}]\subset\mathfrak{g}_1$, and inductively it follows that $\mathfrak{g}^{(j)}\subset\mathfrak{g}_j$. In particular $\mathfrak{g}^{(k)}=0$, so $\mathfrak{g}$ is solvable. Conversely, the derived series of a solvable Lie algebra is a sequence of this type, so this is a characterization of solvability. From this characterization one immediately concludes that if $\mathfrak{g}$ is any Lie algebra which has a solvable ideal $\mathfrak{h}\lhd\mathfrak{g}$ such that the quotient $\mathfrak{g}/\mathfrak{h}$ is solvable, then also $\mathfrak{g}$ itself is solvable.

Remark. While the name "nilpotent" seems pretty self–explanatory, it seems worthwhile to make a comment on the origin of the term "solvable". The name comes from the analogous notion for finite groups and arose in the theory of the Galois group associated to a polynomial in the early 19th century. As in the case of Lie algebras, a finite group is solvable if and only if it can be built up step by step from Abelian groups. The Abelian Galois group $\mathbb{Z}_n=\mathbb{Z}/n\mathbb{Z}$ corresponds to polynomials of the form x^n-a. Of course, the solutions of $x^n-a=0$ are exactly the nth roots of a. If the Galois group G of a polynomial p is solvable, then the step by step construction of G from Abelian groups corresponds to constructing a solution of the equation $p(x)=0$ in terms of iterated radicals. Indeed, the existence of such a solution is equivalent to the Galois group being solvable. The Galois group of a polynomial of degree n is a subgroup of the permutation group $\mathfrak{S}_n$ of n letters (which permutes the n solutions of $p(x)=0$). The fact that polynomials of degree ≤ 4 can be solved using radicals comes from the fact that $\mathfrak{S}_n$ (and hence any subgroup of $\mathfrak{S}_n$) is solvable for $n\le 4$. The group $\mathfrak{S}_5$ is not solvable, (and there are polynomials having $\mathfrak{S}_5$ as their Galois group), which implies that polynomial equations of degree ≥ 5 can not be solved using radicals in general.

2.2. The theorems of Engel and Lie. In example (4) of 1.4 we have introduced the Lie algebras $\mathfrak{b}(n,\mathbb{K})$ and $\mathfrak{n}(n,\mathbb{K})$ of upper triangular and strictly upper triangular $n\times n$–matrices with entries from $\mathbb{K}$. If $X,Y\in\mathfrak{b}(n,\mathbb{K})$ are upper triangular, then the product XY is upper triangular, and its entries on the main diagonal are exactly the products $x_{ii}y_{ii}$ of the corresponding entries of X and Y. Consequently, the commutator $[X,Y]$ is not only upper triangular, but also has zeros in the main diagonal, so we see that $[\mathfrak{b}(n,\mathbb{K}),\mathfrak{b}(n,\mathbb{K})]\subset\mathfrak{n}(n,\mathbb{K})$. It is easy to see that these two subspaces are actually equal, but this is not important for the following arguments. Now assume that $X,Y\in\mathfrak{n}(n,\mathbb{K})$ are strictly upper triangular. Then $x_{ij}=0$ for $j<i+1$ and similarly for Y. Now the (i,j)–component of XY is given as $\sum_{k=1}^n x_{ik}y_{kj}$. Now $x_{ik}=0$ for $k<i+1$ while $y_{kj}=0$ for $j<k+1$. In particular, there is no nonzero summand if $j<i+2$, and the same holds for the commutator $[X,Y]$. Hence $\mathfrak{n}(n,\mathbb{K})^2$ is contained in (and actually equal to) the space of those upper triangular matrices which have zeros also in the first diagonal above the main diagonal. Similarly, one shows that if $y_{ij}=0$ for $j<i+\ell$ and $X\in\mathfrak{n}(n,\mathbb{K})$ is arbitrary, then the (i,j)–components of XY and $[X,Y]$ vanish for $j<i+\ell+1$. Hence the (i,j)–components of any element of $\mathfrak{n}(n,\mathbb{K})^3$ vanish for $j<i+3$, and inductively we see that $\mathfrak{n}(n,\mathbb{K})^n=\{0\}$. Hence the Lie algebra $\mathfrak{n}(n,\mathbb{K})$ is nilpotent.

On the other hand, we have seen that $[\mathfrak{b}(n,\mathbb{K}),\mathfrak{b}(n,\mathbb{K})]$ is contained in $\mathfrak{n}(n,\mathbb{K})$, so it is nilpotent and thus solvable. On the other hand, the quotient $\mathfrak{b}(n,\mathbb{K})/[\mathfrak{b}(n,\mathbb{K}),\mathfrak{b}(n,\mathbb{K})]$ is Abelian, and hence also solvable. In the end of 2.1 we have observed that this implies that $\mathfrak{b}(n,\mathbb{K})$ is solvable. The Lie algebra $\mathfrak{b}(n,\mathbb{K})$ is however not nilpotent (for $n>1$). Essentially, we have seen this in example (3) of 2.1 already: Since any Lie subalgebra of a nilpotent Lie algebra is nilpotent, nilpotency of $\mathfrak{b}(n,\mathbb{K})$ would imply nilpotency of $\mathfrak{b}(2,\mathbb{K})$, which in turn would imply nilpotency of the Lie algebra $\mathfrak{a}(1,\mathbb{K})$ from that example. However, we have seen that there are elements $X,Y\in\mathfrak{a}(1,\mathbb{K})$ such that $[X,Y]=Y$, and thus $[X,[X,Y]]=Y$, and so on, so we get arbitrarily long nonzero brackets.

The theorems of Engel and Lie state that these examples are typical. Before we move to these theorems, let us recall some facts from linear algebra as a warm up:

Proposition. *(1) Let V be a finite dimensional $\mathbb{K}$–vector space and $\varphi:V\to V$ a linear map. Then the following conditions are equivalent:*

- (i) *φ is nilpotent, i.e. there is an $N\in\mathbb{N}$ such that $\varphi^N=0$.*
- (ii) *There is a sequence $\{0\}=V_0\subset V_1\subset\dots\subset V_{\ell-1}\subset V_\ell=V$ of subspaces such that $\varphi(V_i)\subset V_{i-1}$ for all $i=1,\dots,\ell$.*
- (iii) *The matrix representation of φ with respect to an appropriate basis of V is strictly upper triangular.*

(2) Let V be an n–dimensional complex vector space and $\varphi:V\to V$ a complex linear map. Then there is a sequence $V_1\subset V_2\subset\dots\subset V_{n-1}\subset V$ of subspaces such that $\dim(V_i)=i$ and $\varphi(V_i)\subset V_i$ for all $i=1,\dots n-1$. Moreover, the matrix representation of φ with respect to an appropriate basis of V is upper triangular.

Proof. (1) (i) $\implies$ (ii): Define $V_i:=\ker(\varphi^i)$ for $i>0$. Then $\varphi^{i+1}(v)=\varphi(\varphi^i(v))$ implies that $V_i\subset V_{i+1}$ and if $\varphi^N=0$, then $V_N=V$. On the other hand, $\varphi^{i-1}(\varphi(X))=\varphi^i(X)$ implies that $\varphi(V_i)\subset V_{i-1}$.
(ii) $\implies$ (iii): Let i_j denote the dimension of V_j. Choosing a basis $\{v_1,\dots,v_{i_1}\}$ for V_1, extending it to a basis of V_2, then to a basis of V_3, and so on, we obtain a basis $\{v_1,\dots,v_{i_\ell}\}$ of V such that $\{v_1,\dots,v_{i_j}\}$ is a basis of V_j for all $j=1,\dots,\ell$. Now $\varphi(V_1)\subset\{0\}$ implies that $\varphi(v_1)=\dots=\varphi(v_{i_1})=0$, so the first i_1 columns of the matrix

representation of φ consist of zeros only. Next, $\varphi(V_2) \subset V_1$ implies that $\varphi(v_k)$ can be written as a linear combination of $v_1, \dots, v_{i_1}$ for $k = i_1 + 1, \dots, i_2$. In particular, only basis vectors coming before v_k show up in this linear combination, which implies that all entries in the columns $i_1 + 1, \dots, i_2$ lie strictly above the main diagonal. Iterating this argument, the result follows.
(iii) $\implies$ (i): We have seen above, that sufficiently large powers of strictly upper triangular matrices are zero.
(2) By induction on $n = \dim(V)$. If $n = 1$, there is nothing to prove. For $n > 1$, since $\mathbb{C}$ is algebraically closed, there is an eigenvector v_1 for φ, and we define $V_1 \subset V$ to be the one–dimensional subspace spanned by v_1. The fact that v_1 is an eigenvector means that $\varphi(V_1) \subset V_1$. Now we define $W := V/V_1$ and denote by $\pi : V \to W$ the canonical map. Since $\varphi(V_1) \subset V_1$, we see that $\tilde{\varphi}(v + V_1) := \varphi(v) + V_1$ is a well defined $\mathbb{C}$–linear map $W \to W$. By induction, there is a sequence $W_1 \subset \cdots \subset W_{n-2} \subset W$ such that $\dim(W_i) = i$ and $\tilde{\varphi}(W_i) \subset W_i$ for all i. Define $V_i := \pi^{-1}(W_{i-1})$ for $i = 2, \dots, n-1$. Then of course $\dim(V_i) = i$, and the definition of $\tilde{\varphi}$ reads as $\pi \circ \varphi = \tilde{\varphi} \circ \pi$. Hence for $v \in V_i$, we have $\pi(v) \in W_{i-1}$ and thus $\tilde{\varphi}(\pi(v)) = \pi(\varphi(v)) \in W_{i-1}$. But means $\varphi(v) \in V_i$ by construction, so we have constructed an appropriate sequence of subspaces. The second part of the claim is proved exactly like (ii) $\implies$ (iii) above. □

The theorems of Engel and Lie generalize these facts from single linear maps to whole Lie algebras. Therefore, we will have to deal with simultaneous eigenvectors. Suppose that V is a vector space, $A \subset L(V, V)$ a linear subspace, and that $v \in V$ is a nonzero vector which is an eigenvector for any map contained in A. Then for any $\varphi \in A$, there is a number $\lambda(\varphi) \in \mathbb{K}$ such that $\varphi(v) = \lambda(\varphi)v$. Of course, $(\varphi + r\psi)(v) = \varphi(v) + r\psi(v)$, which immediately implies that $\lambda(\varphi + r\psi) = \lambda(\varphi) + r\lambda(\psi)$, so $\varphi \mapsto \lambda(\varphi)$ is actually a linear map $A \to \mathbb{K}$, i.e. a linear functional on A. As in the case of a single linear map, the set $\{v \in V : \varphi(v) = \lambda(\varphi)v \quad \forall \varphi \in A\}$ is obviously a linear subspace of V, which is called the λ–eigenspace of A.

Lemma. *Let V be a $\mathbb{K}$–vector space, $\mathfrak{g}$ a Lie subalgebra of $L(V, V)$, $\mathfrak{h}$ an ideal in $\mathfrak{g}$ and $\lambda : \mathfrak{h} \to \mathbb{K}$ a linear functional. Then the subspace $W := \{v \in V : H(v) = \lambda(H)v \quad \forall H \in \mathfrak{h}\}$ is $\mathfrak{g}$–invariant, i.e. $X(w) \in W$ for all $X \in \mathfrak{g}$ and $w \in W$.*

Proof. Take $w \in W$, $X \in \mathfrak{g}$, and $H \in \mathfrak{h}$. Then by definition of the commutator we get $H(X(w)) = X(H(w)) + [H, X](w)$. Since X is linear, the first term on the right hand side gives $\lambda(H)X(w)$ (which is what we want), and since $\mathfrak{h}$ is an ideal we have $[H, X] \in \mathfrak{h}$, so the second term gives $\lambda([H, X])w$. Hence we can complete the proof by showing that $\lambda([H, X]) = 0$ for all $X \in \mathfrak{g}$ and $H \in \mathfrak{h}$.

Fix $X \in \mathfrak{g}$ and $0 \neq w \in W$ and consider the largest number $k \in \mathbb{N}$ such that $\{w, X(w), X^2(w) = X(X(w)), \dots, X^k(w)\}$ are linearly independent. Let $U \subset V$ be the subspace spanned by these elements. We claim that for $u \in U$ we also have $X(u) \in U$. Of course, it suffices to verify this for the basis elements $w, \dots, X^k(w)$, and for all but the last basis elements it is obvious. However, by construction, the set $\{w, X(w), \dots, X^k(w), X^{k+1}(w)\}$ is linearly dependent, which implies that $X^{k+1}(w)$ can be written as a linear combination of $w, \dots, X^k(w)$ and thus the claim. Let us denote by $\rho_X : U \to U$ the linear map $u \mapsto X(u)$.

Next, consider an arbitrary element $H \in \mathfrak{h}$. Then we know that $H(w) = \lambda(H)w$ and we have seen above that $H(X(w)) = \lambda(H)X(w) + \lambda([H, X])w$. We next prove by induction that $H(X^i(w))$ equals the sum of $\lambda(H)X^i(w)$ and a linear combination of $X^j(w)$ for $j < i$. Since we have just verified the case $i = 1$, let us assume that $i > 1$

and we have proved the statement for $X^{i-1}(w)$. As above, we have

$$HX^i(w) = [H,X]X^{i-1}(w) + XHX^{i-1}(w).$$

Since $[H,X] \in \mathfrak{h}$ the first summand in the right hand side produces only a linear combination of terms of the form $X^j(w)$ for $j < i$. Again by induction hypothesis, we second term can be written as $X(\lambda(H)X^{i-1}(w) + Y)$ where Y is a linear combination of $X^j(w)$ for $j < i-1$, and the claim follows immediately.

This claim shows that $H(u) \in U$ for all $H \in \mathfrak{h}$ and $u \in U$, so we also have the map $\rho_H : U \to U$ for all $H \in \mathfrak{h}$. But we see much more, namely that the matrix representation of ρ_H in the basis $\{w, X(w), \dots, X^k(w)\}$ is upper triangular, with all entries on the main diagonal equal to $\lambda(H)$. Hence we conclude that $\lambda(H) = \frac{1}{k+1}\operatorname{tr}(\rho_H)$. But now for arbitrary $H \in \mathfrak{h}$, we have $[H,X](u) = (HX - XH)(u)$, i.e. $\rho_{[H,X]} = \rho_H \circ \rho_X - \rho_X \circ \rho_H$. But this implies that

$$\lambda([H,X]) = \tfrac{1}{k+1}\operatorname{tr}(\rho_{[H,X]}) = \tfrac{1}{k+1}(\operatorname{tr}(\rho_H \circ \rho_X) - \operatorname{tr}(\rho_X \circ \rho_H)) = 0,$$

since the trace is linear and independent of the ordering of the factors in a product. □

We are now ready to formulate the theorems of Lie and Engel:

Theorem. *(1) [Engel] Let V be a $\mathbb{K}$–vector space and $\mathfrak{g} \subset L(V,V)$ a Lie subalgebra such that the linear map $X : V \to V$ is nilpotent for any $X \in \mathfrak{g}$. Then there is a nonzero vector $v \in V$ such that $X(v) = 0$ for all $X \in \mathfrak{g}$.*
(2) [Lie] Let V be a complex vector space and $\mathfrak{g}$ a Lie algebra of complex linear maps $V \to V$ (but not necessarily a complex subspace of $L_{\mathbb{C}}(V,V)$). If $\mathfrak{g}$ is solvable, then there is a nonzero vector $v \in V$ which is a common eigenvector for all $X \in \mathfrak{g}$, i.e. there is a linear functional $\lambda : \mathfrak{g} \to \mathbb{C}$ such that $X(v) = \lambda(X)v$ for all $X \in \mathfrak{g}$.

Proof. (1) By induction on $n := \dim(\mathfrak{g})$. For $n = 1$, we have $\mathfrak{g} = \{tX : t \in \mathbb{K}\}$, where $X : V \to V$ is nilpotent, and we see from part (1) of the above proposition that X has non–trivial kernel. Of course, if $X(v) = 0$ then also $tX(v) = 0$ for $t \in \mathbb{K}$.

Assuming $n > 1$, the main step in the proof is to show that there is an ideal $\mathfrak{h} \lhd \mathfrak{g}$ such that $\dim(\mathfrak{h}) = \dim(\mathfrak{g}) - 1$. To do this consider a maximal proper Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$. (Since any one–dimensional subspace of $\mathfrak{g}$ is a Lie subalgebra, there exists a maximal subalgebra $\mathfrak{h} \neq \mathfrak{g}$.) For $X \in \mathfrak{h}$ consider the map $\operatorname{ad}_X : \mathfrak{g} \to \mathfrak{g}$ defined by $\operatorname{ad}_X(Y) = [X,Y]$. Since $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$, we have $\operatorname{ad}_X(Y) \in \mathfrak{h}$ for all $Y \in \mathfrak{h}$, so we get a well defined map $\rho_X \in L(\mathfrak{g}/\mathfrak{h}, \mathfrak{g}/\mathfrak{h})$ by putting $\rho_X(Y + \mathfrak{h}) = [X,Y] + \mathfrak{h}$.

Now we obviously have $\rho_{X_1+tX_2} = \rho_{X_1} + t\rho_{X_2}$, so $\{\rho_X : X \in \mathfrak{h}\}$ is a linear subspace of $L(\mathfrak{g}/\mathfrak{h}, \mathfrak{g}/\mathfrak{h})$. The Jacobi identity $[X,[Y,Z]] = [[X,Y],Z] + [Y,[X,Z]]$ can be interpreted as $\operatorname{ad}_X \circ \operatorname{ad}_Y = \operatorname{ad}_{[X,Y]} + \operatorname{ad}_Y \circ \operatorname{ad}_X$, i.e. $\operatorname{ad}_{[X,Y]}$ is the commutator of ad_X and ad_Y. Consequently, we also have $\rho_{[X,Y]} = [\rho_X, \rho_Y]$ for all $X, Y \in \mathfrak{h}$, so $\{\rho_X : X \in \mathfrak{h}\}$ is a Lie subalgebra of $L(\mathfrak{g}/\mathfrak{h}, \mathfrak{g}/\mathfrak{h})$. The dimension of this space is of course $\leq \dim(\mathfrak{h})$ and in particular smaller than n, so we can apply the induction hypothesis to it, once we have shown that each ρ_X is a nilpotent map. Of course, it suffices to prove that ad_X is nilpotent for each X.

By definition, $\operatorname{ad}_X(Y) = XY - YX$ and consequently $\operatorname{ad}(X)^2(Y) = X^2Y - 2XYX + YX^2$. Inductively, we see that $\operatorname{ad}(X)^m(Y)$ is a linear combination of expressions of the form X^iYX^{m-i} for $0 \leq i \leq m$. Since X is nilpotent, we know that $X^N = 0$ for some $N \in \mathbb{N}$. But then if $m \geq 2N$ then in each of the summands X^iYX^{m-i} either $X^i = 0$ or $X^{m-i} = 0$, so we conclude that $\operatorname{ad}(X)^{2N} = 0$.

By induction hypothesis we now find a element $Y + \mathfrak{h} \in \mathfrak{g}/\mathfrak{h}$, which is nonzero (i.e. $Y \notin \mathfrak{h}$) such that $\rho_X(Y) = 0$ for all $X \in \mathfrak{h}$. This means that $[X,Y] \in \mathfrak{h}$ for all

$X \in \mathfrak{h}$. Thus denoting by $\mathfrak{h}'$ the subspace of $\mathfrak{g}$ spanned by $\mathfrak{h}$ and Y, we see that $\mathfrak{h}'$ is a Lie subalgebra of $\mathfrak{g}$ and $\mathfrak{h}$ is an ideal in $\mathfrak{h}'$. Since $\mathfrak{h}$ was assumed to be a maximal proper subalgebra, we must have $\mathfrak{h}' = \mathfrak{g}$, which implies that $\mathfrak{h}$ is an ideal in $\mathfrak{g}$ and $\dim(\mathfrak{h}) = \dim(\mathfrak{g}) - 1$.

Applying the induction hypothesis to $\mathfrak{h}$ we find a nonzero element $v \in V$ such that $X(v) = 0$ for all $X \in \mathfrak{h}$. This means that the space $W := \{v \in V : X(v) = 0 \quad \forall X \in \mathfrak{h}\}$ is nontrivial. By the Lemma, W is $\mathfrak{g}$ invariant, so $Y(w) \in W$ for all $w \in W$. But since Y is nilpotent as a map $V \to V$ it is also nilpotent as a map $W \to W$, so $Y : W \to W$ has nontrivial kernel. Hence there is an element $0 \neq w \in W$ such that $Y(w) = 0$. But $w \in W$ also means $X(w) = 0$ for all $X \in \mathfrak{h}$, so we see that all elements of $\mathfrak{g}$ map w to zero.
(2) Again, we proceed by induction on $n = \dim(\mathfrak{g})$. The case $n = 1$ reduces to the fact that any complex linear map admits an eigenvector, so we assume $n > 1$ and we first need an ideal $\mathfrak{h}$ in $\mathfrak{g}$ such that $\dim(\mathfrak{h}) = \dim(\mathfrak{g}) - 1$. Indeed, since $\mathfrak{g}$ is solvable, $[\mathfrak{g}, \mathfrak{g}]$ must be a proper subspace of $\mathfrak{g}$. (Otherwise, we would have $\mathfrak{g}^{(2)} = \mathfrak{g}$ and hence $\mathfrak{g}^{(k)} = \mathfrak{g}$ for all k.) Consequently, the Abelian Lie algebra $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ has dimension at least one. Choose a nonzero real linear map $\alpha : \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}] \to \mathbb{R}$ and let $\pi : \mathfrak{g} \to \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ be the natural map. Then both π and α are homomorphisms of Lie algebras, so $\alpha \circ \pi$ is a homomorphism and thus $\mathfrak{h} := \ker(\alpha \circ \pi)$ is an ideal in $\mathfrak{g}$. By construction, $\dim(\mathfrak{h}) = \dim(\mathfrak{g}) - 1$ and $\mathfrak{h}$ is solvable as a Lie subalgebra of the solvable Lie algebra $\mathfrak{g}$.

By induction hypothesis, there is a nonzero vector $v_0 \in V$ and a real linear functional $\lambda : \mathfrak{h} \to \mathbb{C}$ such that $X(v_0) = \lambda(X)v_0$ for all $X \in \mathfrak{h}$. Hence the space $W := \{v \in V : X(v) = \lambda(X)v \quad \forall X \in \mathfrak{h}\}$ is a nontrivial complex subspace of V and it is $\mathfrak{g}$–invariant by the Lemma. Choosing a nonzero element $Y \in \mathfrak{g} \setminus \mathfrak{h}$, we see that $\mathfrak{g}$ is spanned by Y and $\mathfrak{h}$, and we consider Y as a linear map $W \to W$. Since Y is complex linear, it has an eigenvector $w \in W$. Denoting by $a \in \mathbb{C}$ the corresponding eigenvalue, we let $\tilde{\lambda} : \mathfrak{g} \to \mathbb{C}$ be the unique real linear map such that $\tilde{\lambda}|_{\mathfrak{h}} = \lambda$ and $\tilde{\lambda}(Y) = a$. Then by construction $X(w) = \tilde{\lambda}(X)w$ for all $X \in \mathfrak{g}$. □

As in the case of a single linear map discussed in the proof of the proposition, one easily derives from this the following consequence. Filling in the details of the proof is a highly recommended exercise.

Corollary. *(1) Let $\mathfrak{g} \subset L(V, V)$ be a Lie subalgebra which satisfies the assumptions of Engel's theorem. Then there is a basis of V with respect to which any element $X \in \mathfrak{g}$ is represented by a strictly upper triangular matrix. In particular, the Lie algebra $\mathfrak{g}$ is isomorphic to a Lie subalgebra of some $\mathfrak{n}(n, \mathbb{K})$ and thus nilpotent.*
(2) Let V be a complex vector space and $\mathfrak{g} \subset L_{\mathbb{C}}(V, V)$ be a solvable (but not necessarily complex) Lie subalgebra. Then there is a basis of V with respect to which any element $X \in \mathfrak{g}$ is represented by an upper triangular matrix. In particular, $\mathfrak{g}$ is isomorphic to a Lie subalgebra of some $\mathfrak{b}(n, \mathbb{C})$.

It should be noted that part (2) actually provides an equivalent condition for simultaneous triangulability of families of linear mappings on complex vector spaces. Namely, if V is a finite dimensional complex vector space and $A \subset L_{\mathbb{C}}(V, V)$ is any set of linear maps, then there is a basis of V with respect to which all elements of A are represented by upper triangular matrices, if and only if the real Lie algebra $\mathfrak{g}$ generated by A is solvable. Indeed, if $\mathfrak{g}$ is solvable, then part (2) of the corollary provides such a basis. Conversely, given such a basis, the set A becomes a subset of $\mathfrak{b}(n, \mathbb{C})$ where $n = \dim(V)$, and since this is a Lie algebra containing A, it also contains $\mathfrak{g}$. Since $\mathfrak{b}(n, \mathbb{C})$ is solvable, we conclude that $\mathfrak{g}$ is solvable, too.

2.3. Semisimple, simple, and reductive Lie algebras.

Definition. Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.
(1) $\mathfrak{g}$ is called *semisimple* if it has no nonzero solvable ideal.
(2) $\mathfrak{g}$ is called *simple* if $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ and the only ideals in $\mathfrak{g}$ are $\{0\}$ and $\mathfrak{g}$.
(3) $\mathfrak{g}$ is called *reductive*, if any solvable ideal of $\mathfrak{g}$ is contained in the *center* $\mathfrak{z}(\mathfrak{g}) := \{X \in \mathfrak{g} : [X, Y] = 0 \quad \forall Y \in \mathfrak{g}\}$ of $\mathfrak{g}$.

Remarks. (1) Obviously, the class of solvable Lie algebras (which in particular includes nilpotent and Abelian Lie algebras) is disjoint from the class of semisimple Lie algebras. We shall see later that any Lie algebra splits into a solvable and a semisimple part.
(2) There is a small choice in the notion of a simple Lie algebra, which is whether one–dimensional Lie algebras (which are automatically Abelian, see 2.1) should be considered as simple or not. I have chosen not to do so, which is the reason for putting the condition that $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ in the definition of a simple Lie algebra. This has the advantage that it ensures that a simple Lie algebra cannot be solvable, since $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ implies $\mathfrak{g} = \mathfrak{g}^{(k)}$ for all $k \in \mathbb{N}$. In particular this implies that any simple Lie algebra is semisimple.
(3) By construction, the center $\mathfrak{z}(\mathfrak{g})$ of a Lie algebra $\mathfrak{g}$ is an Abelian ideal in $\mathfrak{g}$. This immediately implies that any semisimple Lie algebra has trivial center, and that the solvable ideals of a reductive Lie algebra are exactly the subspaces of its center.
(4) We have noted in 2.2 that if $\mathfrak{h} \lhd \mathfrak{g}$ is an ideal, then also $[\mathfrak{h}, \mathfrak{h}]$ is an ideal in $\mathfrak{g}$. Inductively, we see that any term $\mathfrak{h}^{(k)}$ in the derived series of $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. In particular, if $\mathfrak{h}$ is solvable, then the last nontrivial term in the derived series is an Abelian ideal in $\mathfrak{g}$. Thus, a Lie algebra is semisimple if and only if it does not contain a nonzero Abelian ideal.

Examples. (1) The fundamental example of a simple Lie algebra is provided by the algebra $\mathfrak{sl}(n, \mathbb{K})$ of all tracefree $n \times n$–matrices with entries in $\mathbb{K}$ for $n \geq 2$. Let E_{ij} be the *elementary matrix*, which has a one in the j–th column of the i–th row and all other entries equal to zero. The multiplication rule for elementary matrices is simply $E_{ij}E_{k\ell} = \delta_{jk}E_{i\ell}$, where δ_{jk} is the Kronecker delta, i.e. zero for $j \neq k$ and one for $j = k$. Clearly, the elements of the forms E_{ij} and $E_{ii} - E_{jj}$ for $i \neq j$ span the vector space $\mathfrak{g}$. Now for $i \neq j$ we have $[E_{ii} - E_{jj}, E_{ij}] = 2E_{ij}$ and $[E_{ij}, E_{ji}] = E_{ii} - E_{jj}$, which shows that $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$.

To prove simplicity, let $\mathfrak{h} \lhd \mathfrak{g}$ be a nonzero ideal in $\mathfrak{g}$. One immediately verifies that for $A \in \mathfrak{g}$ and any i, j the commutator $[E_{ij}, A]$ is obtained by taking the matrix whose i–th row is the j–th row of A while all other rows are zero, and subtracting from it the matrix whose j–th column is the i–th column of A while all other columns are zero. Now we first observe that this implies that $\mathfrak{h}$ contains a matrix which has a nonzero off–diagonal entry. Indeed, if $A = (a_{ij}) \in \mathfrak{h}$ is diagonal, then since A is tracefree, we find indices $i \neq j$ such that $a_{ii} - a_{jj} \neq 0$. But then $[E_{ij}, A]$ has in the i–th row of the j–th column the entry $a_{jj} - a_{ii}$ and thus a nonzero off–diagonal entry. But now suppose that $A = (a_{ij}) \in \mathfrak{h}$ is such that $a_{ji} \neq 0$ for some fixed $i \neq j$. Then using our description one immediately verifies that $[E_{ij}, [E_{ij}, A]] = -2a_{ji}E_{ij}$, which implies $E_{ij} \in \mathfrak{h}$. Thus, also $[E_{ij}, E_{ji}] = E_{ii} - E_{jj} \in \mathfrak{h}$, which in turn implies E_{ji} in $\mathfrak{h}$. If $n = 2$, we are finished here, otherwise we next observe that for $k \neq i, j$ we have $[E_{ij}, E_{jk}] = E_{ik} \in \mathfrak{h}$ and hence $E_{\ell k} = [E_{\ell i}, E_{ik}] \in \mathfrak{h}$ for $\ell \neq i, k$. Since similar as above we can also get the diagonal matrices contained in $\mathfrak{g}$, this implies $\mathfrak{h} = \mathfrak{g}$.
(2) From simplicity of $\mathfrak{sl}(n, \mathbb{K})$, we can immediately conclude that $\mathfrak{g} := \mathfrak{gl}(n, \mathbb{K})$ is a reductive Lie algebra. Let us first describe the center of $\mathfrak{g}$. Denoting by $\mathbb{I}$ the $n \times n$–unit

matrix, we clearly have $[\mathbb{I}, A] = 0$ for all $A \in \mathfrak{g}$, so $\mathbb{K} \cdot \mathbb{I} := \{t\mathbb{I} : t \in \mathbb{K}\} \subset \mathfrak{z}(\mathfrak{g})$. On the other hand, any $A \in \mathfrak{g}$ can be written (uniquely) as $A = \frac{\operatorname{tr}(A)}{n}\mathbb{I} + (A - \frac{\operatorname{tr}(A)}{n}\mathbb{I})$ and the second summand is obviously tracefree, which implies that $\mathfrak{g} = \mathbb{K} \cdot \mathbb{I} \oplus \mathfrak{sl}(n, \mathbb{K})$ as a vector space.

Since the commutator of any two matrices is tracefree, we see that $[\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{sl}(n, \mathbb{K})$ and from (1) we see that we must have equality. For a solvable ideal $\mathfrak{h}$ in $\mathfrak{g}$ the intersection of $\mathfrak{h}$ with $\mathfrak{sl}(n, \mathbb{C})$ is a solvable ideal in that Lie algebra and thus has to vanish. But clearly $[\mathfrak{g}, \mathfrak{h}] \subset \mathfrak{h} \cap [\mathfrak{g}, \mathfrak{g}] = \{0\}$, so $\mathfrak{h} \subset \mathfrak{z}(\mathfrak{g})$ and hence $\mathfrak{g}$ is reductive. In particular, $\mathfrak{z}(\mathfrak{g}) \cap [\mathfrak{g}, \mathfrak{g}] = \{0\}$, and since $\mathbb{K} \cdot \mathbb{I} \subset \mathfrak{z}(\mathfrak{g})$ we conclude that $\mathfrak{z}(\mathfrak{g}) = \mathbb{K} \cdot \mathbb{I}$. Hence we see that $\mathfrak{g} = \mathfrak{z}(\mathfrak{g}) \oplus [\mathfrak{g}, \mathfrak{g}]$ as a Lie algebra. We will see later that this is always the case for reductive Lie algebras.

(3) In example (2) of 2.1 we have met the direct sum $\mathfrak{g} \oplus \mathfrak{h}$ of two Lie algebras. By construction, $\mathfrak{g} = \{(X, 0) : X \in \mathfrak{g}\} \subset \mathfrak{g} \oplus \mathfrak{h}$ is an ideal and similarly for $\mathfrak{h}$. By definition of the bracket on $\mathfrak{g} \oplus \mathfrak{h}$ the projection onto the first factor is a surjective Lie algebra homomorphism $\pi_{\mathfrak{g}} : \mathfrak{g} \oplus \mathfrak{h} \to \mathfrak{g}$, and similarly for $\mathfrak{h}$. Now if $\mathfrak{a} \subset \mathfrak{g} \oplus \mathfrak{h}$ is an ideal, then $\pi_{\mathfrak{g}}(\mathfrak{a}) \subset \mathfrak{g}$ is immediately seen to be an ideal in $\mathfrak{g}$. Assuming that $\mathfrak{a}$ is solvable and $\mathfrak{g}$ is semisimple, we see that $\pi_{\mathfrak{g}}(\mathfrak{a}) = 0$, since this is also solvable, and thus $\mathfrak{a} \subset \mathfrak{h}$. If $\mathfrak{h}$ is semisimple, too, then this implies $\mathfrak{a} = \{0\}$, so we conclude that the direct sum of two semisimple Lie algebras is semisimple.

Clearly, we can also form direct sums of finitely many factors, and in particular we see that a (finite) direct sum of simple Lie algebras is semisimple. We shall see below that also the converse is true, i.e. any semisimple Lie algebra is a direct sum of simple ideals.

Representations and the Killing Form

Having introduced the basic subclasses of Lie algebras, we next characterize theses classes in terms of a canonical invariant bilinear form, called the Killing form. This form is a special case of a more general construction of invariant bilinear forms from representations of a Lie algebra.

2.4. Representations. We have already met the definition of a representation of a Lie algebra in 1.5:

Definition. A (finite–dimensional) *representation* of a Lie algebra $\mathfrak{g}$ on a (finite–dimensional) $\mathbb{K}$–vector space V is a homomorphism $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ of Lie algebras.

In the finite dimensional case, we may always choose a basis to identify V with $\mathbb{K}^n$ for $n = \dim(V)$ and thus view representations as homomorphisms to $\mathfrak{gl}(n, \mathbb{K})$, but it will often be more convenient to work with an abstract vector space. We have also noted already that one may equivalently view a representation as being given by a bilinear map $\rho : \mathfrak{g} \times V \to V$ such that $\rho([X, Y], v) = \rho(X, \rho(Y, v)) - \rho(Y, \rho(X, v))$. If the representation under consideration is clear from the context, then we will often simply write $X \cdot v$ or Xv for $\rho(X)(v)$.

If $\mathfrak{g}$ is real, then by a *complex representation* we will mean a representation by complex linear maps on a complex vector space. In the picture of bilinear maps, this simply means that $\rho : \mathfrak{g} \times V \to V$ is complex linear in the second variable. If $\mathfrak{g}$ is itself complex, then "complex representation" will also mean that the homomorphism defining the representation is complex linear, or in the bilinear picture that $\rho : \mathfrak{g} \times V \to V$ is complex bilinear.

If $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ and $\rho' : \mathfrak{g} \to \mathfrak{gl}(V')$ are representations, then a *homomorphism* from ρ to ρ' is a linear map $\varphi : V \to V'$ which is compatible with the $\mathfrak{g}$ actions, i.e. such that $\varphi(\rho(X)(v)) = \rho'(X)(\varphi(v))$ or in the simple notation $\varphi(X \cdot v) = X \cdot \varphi(v)$. Homomorphisms of representations are also called *intertwining operators* or *equivariant maps.* An *isomorphism* of representations is a bijective homomorphism. If $\varphi : V \to W$ is an isomorphism, then the inverse map $\varphi^{-1} : W \to V$ is automatically a homomorphism, too.

A representation is called *trivial* if any $X \in \mathfrak{g}$ acts by the zero map on V.

An important example of a representation of a Lie algebra is the *adjoint representation*, $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ defined by $\mathrm{ad}(X)(Y) := [X, Y]$. As we have seen in the proof of theorem 2.2, the Jacobi identity for the bracket exactly states that this is a representation of $\mathfrak{g}$.

A representation $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is called *faithful* iff the map ρ is injective. If this is not the case, then the kernel of ρ is an ideal in $\mathfrak{g}$. In particular, considering the adjoint representation we see that the kernel by definition is the set of all $X \in \mathfrak{g}$ such that $[X, Y] = 0$ for all $Y \in \mathfrak{g}$, so this is exactly the center of $\mathfrak{g}$. On the other hand, by definition any non–trivial representation of a simple Lie algebra is automatically faithful.

Suppose that $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is a representation. A *subrepresentation* or an *invariant subspace* of V is a linear subspace $W \subset V$ such that $\rho(X)(w) \in W$ for all $X \in \mathfrak{g}$ and all $w \in W$. A representation V is called *irreducible* iff $\{0\}$ and V are the only invariant subspaces.

Lie's theorem from 2.2 above immediately gives us some information on complex representations of solvable Lie algebras. Indeed, the statement immediately implies that any such representation contains an invariant subspace of dimension one, which in particular implies that irreducible complex representations are automatically one–dimensional and trivial. More precisely, any such representation is given by a complex valued linear functional on the Abelian Lie algebra $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$.

Note that for a homomorphism $\varphi : V \to W$ between two representations, the kernel $\ker(\varphi)$ is a subrepresentation of V and the image $\varphi(V)$ is a subrepresentation of W. In particular, this implies that a nonzero homomorphism with irreducible source is automatically injective while for irreducible target it is automatically surjective. Hence, a homomorphism between irreducible representations is either zero or an isomorphism. A simple consequence of this is called Schur's lemma:

LEMMA (Schur). *Let V be a complex irreducible representation of a Lie algebra $\mathfrak{g}$. Then any homomorphism $\varphi : V \to V$ is a complex multiple of the identity map.*

PROOF. If $\varphi : V \to V$ is a nonzero homomorphism, then for each $\lambda \in \mathbb{C}$ also $\varphi - \lambda\,\mathrm{id}$ is a homomorphism. Since φ must have an eigenvalue λ_0, we conclude that $\varphi - \lambda_0\,\mathrm{id}$ has nontrivial kernel, so from above we know that it is identically zero. □

Many natural constructions for vector spaces can be naturally extended to representations of Lie algebras. Given two representations of a Lie algebra $\mathfrak{g}$ on vector spaces V and W, there is an obvious representation on the direct sum $V \oplus W$, defined by $X \cdot (v, w) := (X \cdot v, X \cdot w)$. This is called the *direct sum* of the representations V and W. Similarly, given representations of $\mathfrak{g}$ on V and W, we can construct a natural representation on the space $L(V, W)$ of all linear maps from V to W. Namely, for $\varphi : V \to W$ and $X \in \mathfrak{g}$ we define $(X \cdot \varphi)(v) := X \cdot (\varphi(v)) - \varphi(X \cdot v)$. This definition may look a bit strange at the first glance, but it is exactly the differentiated version of the group action on $L(V, W)$ from 1.2.

Let us verify that this indeed defines a representation: By definition, we compute

$$(X\cdot(Y\cdot\varphi))(v) = X\cdot((Y\cdot\varphi)(v)) - (Y\cdot\varphi)(X\cdot v) =$$
$$X\cdot(Y\cdot(\varphi(v))) - X\cdot(\varphi(Y\cdot v)) - Y\cdot(\varphi(X\cdot v)) + \varphi(Y\cdot(X\cdot v)).$$

Subtracting the same term with X and Y exchanged, the two middle summands cancel, and we are left with

$$X\cdot(Y\cdot(\varphi(v))) - Y\cdot(X\cdot(\varphi(v))) - \varphi(X\cdot(Y\cdot v) - Y\cdot(X\cdot v)),$$

which equals $([X,Y]\cdot\varphi)(v)$ since we have representations on V and W.

In particular, we may use for W the trivial representation on $\mathbb{K}$, thus obtaining a natural representation on the dual space $\mathbb{V}^* = L(V,\mathbb{K})$ of V. This is called the *dual representation* or the *contragradient representation* of V. By definition, $(X\cdot\varphi)(v) = -\varphi(X\cdot v)$ in this case. Denoting the given representation by $\rho:\mathfrak{g}\to\mathfrak{gl}(V)$ and the contragradient by $\rho^*:\mathfrak{g}\to\mathfrak{gl}(V^*)$, we hence have $\rho^*(X) = (\rho(-X))^*$ by definition of the dual mapping. We will discuss the important construction of tensor products of representations in chapter 4.

A finite dimensional representation $\rho:\mathfrak{g}\to V$ of a Lie algebra $\mathfrak{g}$ is called *completely reducible* if it can be written as a direct sum of irreducible representations. Finally, the representation is called *indecomposable* if and only if there are no subrepresentations $W_1, W_2 \subset V$, which both are of positive dimension such that $V = W_1 \oplus W_2$.

There is a certain subtlety in the relations between indecomposability and irreducibility. First of all, more or less by definition, any representation can be written as a direct sum of indecomposable representations. On the other hand, complete reducibility is a rather restrictive condition in general. To see this, consider complex representations of a solvable Lie algebra $\mathfrak{g}$. From above, we know that any irreducible complex representation comes from a representation of $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$, so this also holds for any completely reducible representation. Hence on any such representation, the derived algebra $[\mathfrak{g},\mathfrak{g}]$ acts trivially, so they do not seem much of the Lie algebra structure of $\mathfrak{g}$. In particular, the adjoint representation of $\mathfrak{g}$ can not be completely reducible except for some trivial cases.

One reason for the popularity of the notion of irreducibility is that large part of the representation theory of Lie algebras have their origins in the study of *unitary representations* of a Lie group, i.e. representations admitting an invariant positive definite inner product. The corresponding condition on the Lie algebraic level is that there is a positive definite inner product on the representation space V such that any element of $\mathfrak{g}$ acts by a skew symmetric (respectively skew hermitian) operator on V, i.e. $\rho(X)^* = -\rho(X)$ for all $X \in \mathfrak{g}$. Writing the inner product as $\langle\ ,\ \rangle$, the condition means that $\langle X\cdot v, w\rangle = -\langle v, X\cdot w\rangle$. In particular this implies that for a subrepresentation $W \subset V$ the orthogonal complement $W^\perp$ is an invariant subspace, too. Using this, we see that any such representation is completely reducible, and in particular the notions of indecomposability and irreducibility coincide. We shall soon see that any representation of a semisimple Lie algebra is completely reducible.

2.5. Complexification. Before we can continue developing the general theory of Lie algebras, we have to discuss the relation between real and complex Lie algebras and in particular the technique of complexification. Recall that for a real vector space V, the complexification $V_\mathbb{C}$ is $V\times V$ with the component–wise addition and scalar multiplication defined by $(a+ib)(v_1,v_2) := (av_1 - bv_2, bv_1 + av_2)$. One easily verifies directly that this makes $V\times V$ into a complex vector space. We may view V as the real subspace $\{(v,0) : v\in V\}$ of $V_\mathbb{C}$. By definition, $i(v,0) = (0,v)$ which implies that the element (v_1, v_2) can be written as $v_1 + iv_2$. Moreover, any real basis of V can be considered as a basis for the $\mathbb{C}$–vector space $V_{\mathbb{C}}$.

Now assume that $\mathfrak{g}$ is a real Lie algebra and let $\mathfrak{g}_{\mathbb{C}}$ be the complexification of the vector space $\mathfrak{g}$. Define a Lie bracket on $\mathfrak{g}_{\mathbb{C}}$ by

$$[(X_1, Y_1), (X_2, Y_2)] := ([X_1, X_2] - [Y_1, Y_2], [X_1, Y_2] + [Y_1, X_2]).$$

This is visibly skew symmetric and real bilinear. Moreover,

$$[i(X_1, Y_1), (X_2, Y_2)] = [(-Y_1, X_1), (X_2, Y_2)] = (-[Y_1, X_2] - [X_1, Y_2], -[Y_1, Y_2] + [X_1, X_2]),$$

which equals $i[(X_1, Y_1), (X_2, Y_2)]$, so the bracket is complex bilinear. Since any element of $\mathfrak{g}_{\mathbb{C}}$ can be written as a complex linear combination of elements of $\mathfrak{g}$, we conclude from complex bilinearity that to see that the bracket of $\mathfrak{g}_{\mathbb{C}}$ satisfies the Jacobi identity, it suffices to verify it for elements of $\mathfrak{g}$. Hence $\mathfrak{g}_{\mathbb{C}}$ is a complex Lie algebra.

The complexification of a vector space has a universal property. Indeed, consider a real vector space V as a subspace of its complexification $V_{\mathbb{C}}$ and any complex vector space W. If $\varphi : V \to W$ is any real linear map, then we define $\tilde{\varphi} : V_{\mathbb{C}} \to W$ by $\tilde{\varphi}(v_1, v_2) := \varphi(v_1) + i\varphi(v_2)$. Of course, the restriction to V is given by $\tilde{\varphi}|_V = \varphi$. Moreover, $\tilde{\varphi}$ is visibly $\mathbb{R}$–linear and

$$\tilde{\varphi}(i(v_1, v_2)) = \tilde{\varphi}(-v_2, v_1) = -\varphi(v_2) + i\varphi(v_1) = i\tilde{\varphi}(v_1, v_2),$$

which shows that $\tilde{\varphi}$ is complex linear. Of course, any complex linear map is determined by its restriction to V, so we see that any real linear map $\varphi : V \to W$ into a complex vector space extends uniquely to a complex linear map $\tilde{\varphi} : V_{\mathbb{C}} \to W$. In particular, a real linear map $\varphi : V \to V$ may be viewed as having values in the complex vector space $V_{\mathbb{C}}$ and thus extends uniquely to a complex linear map $\varphi_{\mathbb{C}} : V_{\mathbb{C}} \to V_{\mathbb{C}}$. Choosing a basis of V we may identify linear maps $V \to V$ with $n \times n$–matrices, where $n = \dim_{\mathbb{R}}(V)$. As we have noted above, we may also view this basis as a basis of $V_{\mathbb{C}}$ and in this basis $\varphi_{\mathbb{C}}$ has the same matrix representation as φ (but viewing the real matrix as a complex matrix).

Consider the special case of a real Lie algebra $\mathfrak{g}$ with complexification $\mathfrak{g}_{\mathbb{C}}$ and a (real) Lie algebra homomorphism φ into a complex Lie algebra $\mathfrak{h}$. Then the complex linear map $\tilde{\varphi} : \mathfrak{g}_{\mathbb{C}} \to \mathfrak{h}$ is defined by $\tilde{\varphi}(X, Y) = \varphi(X) + i\varphi(Y)$. We claim that $\tilde{\varphi}$ is a homomorphism of Lie Algebras. Since any element of $\mathfrak{g}_{\mathbb{C}}$ can be written as a (complex) linear combination of elements of $\mathfrak{g}$ and $\tilde{\varphi}$ is complex linear, it suffices to verify the equation $\tilde{\varphi}([X, Y]) = [\tilde{\varphi}(X), \tilde{\varphi}(Y)]$ in the case where $X, Y \in \mathfrak{g} \subset \mathfrak{g}_{\mathbb{C}}$. But on such elements $\tilde{\varphi}$ coincides with φ so the result is obvious. Hence we see that any real Lie algebra homomorphism $\varphi : \mathfrak{g} \to \mathfrak{h}$ to a complex Lie algebra $\mathfrak{h}$ uniquely extends to a homomorphism $\tilde{\varphi} : \mathfrak{g}_{\mathbb{C}} \to \mathfrak{h}$ of complex Lie algebras. In particular, any representation of $\mathfrak{g}$ on a complex vector space V extends uniquely to a complex representation of $\mathfrak{g}_{\mathbb{C}}$.

Examples. We next discuss two examples which show that it is often possible to describe complexifications nicely. These examples are also important since they are the simplest examples which show that passing to the complexification leads to a loss of information.

(1) Consider $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{R})$. The inclusion of real matrices into complex matrices gives us a real linear map $\mathfrak{g} \hookrightarrow \mathfrak{sl}(2, \mathbb{C})$, which extends to a homomorphism $\mathfrak{g}_{\mathbb{C}} \to \mathfrak{sl}(2, \mathbb{C})$, which is given by $(A, B) \mapsto A + iB$. Since both A and B have real entries, this homomorphism is visibly injective, and since $\mathfrak{g}$ has real dimension three, the complex dimension of $\mathfrak{g}_{\mathbb{C}}$ is also three, so $\mathfrak{g}_{\mathbb{C}} \cong \mathfrak{sl}(2, \mathbb{C})$.

(2) Consider the subspace $\mathfrak{h} := \mathfrak{su}(2) \subset \mathfrak{sl}(2, \mathbb{C})$ of tracefree complex matrices A such that $A^* = -A$, where A^* is the conjugate transpose of A. Explicitly, this means that

$A = \begin{pmatrix} ia & z \\ -\bar{z} & -ia \end{pmatrix}$ for some $a \in \mathbb{R}$ and $z \in \mathbb{C}$, so we see that $\mathfrak{h}$ is a real three–dimensional subspace of $\mathfrak{sl}(2, \mathbb{C})$. From the fact that $(AB)^* = B^*A^*$ one immediately concludes that $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{sl}(2, \mathbb{C})$. As above, the inclusion $\mathfrak{h} \hookrightarrow \mathfrak{sl}(2, \mathbb{C})$ extends to a homomorphism $\mathfrak{h}_{\mathbb{C}} \to \mathfrak{sl}(2, \mathbb{C})$, which is given by $(A, B) \mapsto A + iB$. One immediately verifies that this is also injective, so we conclude that $\mathfrak{h}_{\mathbb{C}} \cong \mathfrak{sl}(2, \mathbb{C})$ and thus $\mathfrak{h}_{\mathbb{C}} \cong \mathfrak{g}_{\mathbb{C}}$. We shall see later, that the Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ themselves are not isomorphic.

In spite of the fact that one looses some information by passing to the complexification of a Lie algebra, many important properties are preserved. Note first that by construction any real Lie algebra $\mathfrak{g}$ is a real Lie subalgebra of its complexification $\mathfrak{g}_{\mathbb{C}}$. From 2.1 we hence conclude that if $\mathfrak{g}_{\mathbb{C}}$ is solvable (nilpotent), then also $\mathfrak{g}$ is solvable (nilpotent). The converse of these assertions is also true: Since $\mathfrak{g}$ is a Lie subalgebra of $\mathfrak{g}_{\mathbb{C}}$ we obviously have $[\mathfrak{g}, \mathfrak{g}] \subset [\mathfrak{g}_{\mathbb{C}}, \mathfrak{g}_{\mathbb{C}}]$, and thus also the complexification $[\mathfrak{g}, \mathfrak{g}]_{\mathbb{C}}$ (which obviously sits inside $\mathfrak{g}_{\mathbb{C}}$) is contained in $[\mathfrak{g}_{\mathbb{C}}, \mathfrak{g}_{\mathbb{C}}]$. On the other hand, from the explicit formula for the bracket on $\mathfrak{g}_{\mathbb{C}}$ we see that for $(X_i, Y_i) \in \mathfrak{g}_{\mathbb{C}}$ both components of $[(X_1, Y_1), (X_2, Y_2)]$ are linear combinations of brackets, so we conclude that $[\mathfrak{g}_{\mathbb{C}}, \mathfrak{g}_{\mathbb{C}}] = [\mathfrak{g}, \mathfrak{g}]_{\mathbb{C}}$. Otherwise put, $[\mathfrak{g}_{\mathbb{C}}, \mathfrak{g}_{\mathbb{C}}]$ is the complex subspace of $\mathfrak{g}_{\mathbb{C}}$ spanned by $[\mathfrak{g}, \mathfrak{g}]$. Inductively, this implies that $(\mathfrak{g}_{\mathbb{C}})^k$ is the complex subspace spanned by $\mathfrak{g}^k$ and $(\mathfrak{g}_{\mathbb{C}})^{(k)}$ is the complex subspace spanned by $\mathfrak{g}^{(k)}$. In particular, this implies that if $\mathfrak{g}$ is solvable (nilpotent) then also $\mathfrak{g}_{\mathbb{C}}$ has this property.

This also leads to preliminary information on semisimplicity: Suppose that $\mathfrak{g}$ is a real Lie algebra, whose complexification $\mathfrak{g}_{\mathbb{C}}$ is semisimple. Then for a solvable ideal $\mathfrak{h} \subset \mathfrak{g}$ we can consider $\mathfrak{h}_{\mathbb{C}} \subset \mathfrak{g}_{\mathbb{C}}$, and we know from above that this is a solvable subalgebra. But the explicit formula for the bracket on $\mathfrak{g}_{\mathbb{C}}$ immediately shows that $\mathfrak{h}_{\mathbb{C}}$ is an ideal in $\mathfrak{g}_{\mathbb{C}}$, so semisimplicity of $\mathfrak{g}_{\mathbb{C}}$ implies $\mathfrak{h}_{\mathbb{C}} = \{0\}$ and thus $\mathfrak{h} = \{0\}$. Hence, $\mathfrak{g}$ is semisimple, too. We will soon see that the converse of this result also holds, i.e. that a real Lie algebra $\mathfrak{g}$ is semisimple if and only if its complexification $\mathfrak{g}_{\mathbb{C}}$ is semisimple.

2.6. The Killing form. Let $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ be a finite dimensional representation of a Lie algebra $\mathfrak{g}$ on a $\mathbb{K}$–vector space V. Then we define a bilinear form $B_\rho = B_V : \mathfrak{g} \times \mathfrak{g} \to \mathbb{K}$ by $B_\rho(X, Y) := \operatorname{tr}(\rho(X) \circ \rho(Y))$. This form is obviously symmetric since $\operatorname{tr}(\rho(X) \circ \rho(Y)) = \operatorname{tr}(\rho(Y) \circ \rho(X))$. Moreover we may compute

$$\begin{aligned} B_\rho([X, Y], Z) = {} & \operatorname{tr}((\rho(X) \circ \rho(Y) - \rho(Y) \circ \rho(X)) \circ \rho(Z)) = \\ & \operatorname{tr}(\rho(X) \circ \rho(Y) \circ \rho(Z) - \rho(X) \circ \rho(Z) \circ \rho(Y)) = B_\rho(X, [Y, Z]). \end{aligned}$$

This may be equivalently written as $B(\operatorname{ad}_Y(X), Z) = -B(X, \operatorname{ad}_Y(Z))$. Bilinear forms $\mathfrak{g} \times \mathfrak{g} \to \mathbb{K}$ having this property are called *invariant*. In particular, applying this construction to the adjoint representation, we obtain the *Killing form* $B =: B_{\operatorname{ad}}$ of the Lie algebra $\mathfrak{g}$, which is one of the main tools in for the study of semisimple Lie algebras. The Killing form has stronger invariance properties than the general forms B_ρ. Indeed, let $\varphi : \mathfrak{g} \to \mathfrak{g}$ be any automorphism of the Lie algebra $\mathfrak{g}$, i.e. an isomorphism from $\mathfrak{g}$ to itself. Then the equation $\varphi([X, Y]) = [\varphi(X), \varphi(Y)]$ can be interpreted as $\varphi \circ \operatorname{ad}_X = \operatorname{ad}_{\varphi(X)} \circ \varphi$, which means that $\operatorname{ad}_{\varphi(X)} = \varphi \circ \operatorname{ad}_X \circ \varphi^{-1}$. But this implies $\operatorname{ad}(\varphi(X)) \circ \operatorname{ad}(\varphi(Y)) = \varphi \circ \operatorname{ad}(X) \circ \operatorname{ad}(Y) \circ \varphi^{-1}$, so both sides have the same trace and thus $B(\varphi(X), \varphi(Y)) = B(X, Y)$. Hence the Killing form is invariant under arbitrary automorphisms of $\mathfrak{g}$. For automorphisms of the form $e^{\operatorname{ad}(X)} = \operatorname{Ad}(\exp(X))$ with $X \in \mathfrak{g}$, invariance of any form B_ρ follows from the above infinitesimal invariance property, but in general there are much more automorphisms.

Note that if $\mathfrak{g}$ is a real Lie algebra with complexification $\mathfrak{g}_{\mathbb{C}}$ and V is a complex representation of $\mathfrak{g}$, then we know from 2.5 above, that the representation extends to a representation $\tilde{\rho}$ of $\mathfrak{g}_{\mathbb{C}}$. Hence we get an extension of the (complex valued) trace from B_ρ to a bilinear form $B_{\tilde{\rho}}$ on $\mathfrak{g}_{\mathbb{C}}$. By construction, $\tilde{\rho}(X,Y) = \tilde{\rho}(X+iY) = \rho(X) + i\rho(Y)$ and thus

$$\tilde{\rho}(X_1,Y_1)\circ\tilde{\rho}(X_2,Y_2) = \rho(X_1)\circ\rho(X_2) - \rho(Y_1)\rho(Y_2) + i(\rho(X_1)\circ\rho(Y_2) + \rho(Y_1)\circ\rho(X_2)).$$

Taking the trace we conclude that

$$B_{\tilde{\rho}}(X_1+iY_1, X_2+iY_2) = B_\rho(X_1,X_2) - B_\rho(Y_1,Y_2) + i(B_\rho(X_1,Y_2) + B_\rho(Y_1,X_2)),$$

so this is just the complex bilinear extension of B_ρ.

Starting from a real representation of $\mathfrak{g}$ on V, we may first complexify V to obtain a complex representation of $\mathfrak{g}$. Using a basis of V as a basis of $V_{\mathbb{C}}$ the complex linear extension of a map $V \to V$ is represented by the same matrix, which implies $B_{V_{\mathbb{C}}} = B_V$, where we view B_V as a complex form having only real values. Then we can extend $V_{\mathbb{C}}$ to a representation of $\mathfrak{g}_{\mathbb{C}}$, and the resulting form on $\mathfrak{g}_{\mathbb{C}}$ is again the complex bilinear extension of B_V. In particular, the Killing form $B_{\mathfrak{g}_{\mathbb{C}}}$ of $\mathfrak{g}_{\mathbb{C}}$ is simply the complex bilinear extension of $B_{\mathfrak{g}}$.

Example. Using the Killing form, we can now show that the Lie algebras $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{R})$ and $\mathfrak{h} = \mathfrak{su}(2)$ from examples (1) and (2) of 2.5 are not isomorphic. If $\varphi : \mathfrak{g} \to \mathfrak{h}$ is an isomorphism, then as in the proof of invariance of the Killing form above we get $\mathrm{ad}(\varphi(X)) = \varphi \circ \mathrm{ad}(X) \circ \varphi^{-1}$ and thus $B_{\mathfrak{h}}(\varphi(X),\varphi(Y)) = B_{\mathfrak{g}}(X,Y)$. So we can prove that the two Lie algebras cannot be isomorphic by showing that they have different Killing forms. Moreover, since both $\mathfrak{g}$ and $\mathfrak{h}$ have complexification $\mathfrak{sl}(2,\mathbb{C})$ we see that their Killing forms are simply the restrictions of the Killing form of $\mathfrak{sl}(2,\mathbb{C})$, so we will denote all the Killing forms by B.

To compute this Killing form, we look at a special basis of $\mathfrak{sl}(2,\mathbb{C})$ which will be very important in the sequel. Put $H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Clearly, these elements form a complex basis of $\mathfrak{sl}(2,\mathbb{C})$ (and a real basis of $\mathfrak{sl}(2,\mathbb{R})$), and one immediately computes that $[H,E] = 2E$, $[H,F] = -2F$ and $[E,F] = H$. This easily implies that in the basis $\{E,H,F\}$ we get the matrix representations

$$\mathrm{ad}(H) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} \qquad \mathrm{ad}(E) = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad \mathrm{ad}(F) = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}.$$

From this, one immediately concludes that $B(E,E) = B(F,F) = B(H,E) = B(H,F) = 0$, while $B(H,H) = 8$ and $B(E,F) = 4$. In particular, $0 \neq E \in \mathfrak{g}$ satisfies $B(E,E) = 0$.

On the other hand, the elements iH, $u := E - F$ and $v := i(E+F)$ form a basis of $\mathfrak{h}$. From above, one immediately sees that $B(iH,u) = B(iH,v) = 0$ while we compute

$$B(u,v) = iB(E-F,E+F) = i(B(E,E) + B(E,F) - B(F,E) - B(F,F)) = 0,$$

so these three vectors are orthogonal with respect to B. But $B(iH,iH) = -B(H,H) = -8$ and $B(u,u) = B(v,v) = -2B(E,F) = -8$, so all the basis vectors have negative inner products with themselves. Hence we conclude that for $A = aiH + bu + cv \in \mathfrak{h}$ we get $B(A,A) = -8(a^2+b^2+c^2)$, which is < 0 if $A \neq 0$.

Suppose that $\mathfrak{g}$ is a solvable Lie algebra, V is a finite dimensional complex vector space, and ρ is a representation of $\mathfrak{g}$ on V. By Lie's theorem (see 2.2), there is a basis of V such that any element of X acts by an upper triangular matrix. Since the commutator of two upper triangular matrices is strictly upper triangular, we conclude that any element of $[\mathfrak{g},\mathfrak{g}]$ acts by a strictly upper triangular matrix. Moreover, the product of an upper triangular matrix with a strictly upper triangular matrix is strictly upper triangular, and thus in particular tracefree, we see that $B_V(X,Y)=0$ for any $X\in\mathfrak{g}$ and $Y\in[\mathfrak{g},\mathfrak{g}]$. Passing to a complexification of the representation space, we see that this holds for real representations, too. In particular, we get the result for the Killing form.

Cartan's criterion for solvability states, that this property of the Killing form actually characterizes solvable Lie algebras. This will immediately lead to a nice characterization of semisimple Lie algebras in terms of the Killing form.

2.7. Jordan decomposition. To prove Cartan's criterion for solvability, we need a lemma on Jordan decompositions. These are a weakening of the Jordan normal form but stated in a basis independent way, and we briefly recall the background from linear algebra.

Definition. Let V be a finite dimensional complex vector space and $f:V\to V$ a complex linear map. Then a *Jordan decomposition* of f is an expression $f=f_1+f_2$, where $f_1,f_2:V\to V$ are complex linear maps such that f_1 is diagonalizable and f_2 is nilpotent and such that $f_1\circ f_2=f_2\circ f_1$.

Proposition. *Let V be a finite dimensional complex vector space and $f:V\to V$ a complex linear map. Then there is a unique Jordan decomposition $f=f_S+f_N$ of f.*

Sketch of proof. Existence: Since $\mathbb{C}$ is algebraically closed, the characteristic polynomial p_f of f can be written as $p_f(t)=(\lambda_1-t)^{n_1}\cdots(\lambda_k-t)^{n_k}$, where $\lambda_1,\dots,\lambda_k$ are the different eigenvalues of f and n_i is the algebraic multiplicity of λ_i. The sum $n_1+\dots+n_k$ of these multiplicities equals the dimension of V. Recall that for any complex polynomial $p(t)=a_0+a_1t+\dots+a_Nt^N$, one has the linear map $p(f)=a_0\operatorname{id}+a_1f+\dots+a_Nf^N$. Obviously, if p and q are polynomials then $p(f)\circ q(f)=(pq)(f)$ so in particular the two maps commute. Further, the Cayley–Hamilton theorem says that $p_f(f)=0$.

Now for any eigenvalue λ_i the generalized eigenspace $V_{(\lambda_i)}$ of f with eigenvalue λ_i is defined as the kernel of $(\lambda_i\operatorname{id}-f)^{n_i}$. For $v\in V_{(\lambda_i)}$ we obviously have $\lambda_iv-f(v)\in V_{(\lambda_i)}$ and since λ_iv lies in that subspace anyhow, we conclude that $f(V_{(\lambda_i)})\subset V_{(\lambda_i)}$.

Next, we claim that $V=V_{(\lambda_1)}\oplus\dots\oplus V_{(\lambda_k)}$ and the projection $V\to V_{(\lambda_i)}$ which takes out the component in that space can be written as a polynomial in f. We prove this by induction on the number k of different eigenvalues. For $k=1$, the Cayley–Hamilton theorem implies that $(\lambda_1\operatorname{id}-f)^{n_1}=0$ and thus $V=V_{(\lambda_1)}$ and the projection is the identity map.

So let us assume $k\geq 1$ and that the statement has been proved for $k-1$ different eigenvalues. Define $p_1:=(\lambda_1-t)^{n_1}$ and let p_2 be the product of the other factors in p_f, so $p_f=p_1p_2$ and thus $p_1(f)\circ p_2(f)=0$. Moreover, p_1 and p_2 are relatively prime, i.e. if we have polynomials r,s_1,s_2 such that $p_1=rs_1$ and $p_2=rs_2$, then r is constant: Suppose that $r(z_0)=0$. This implies that $p_1(z_0)=0$ and $p_2(z_0)=0$, which on the one hand implies that $z_0=\lambda_1$ and on the other hand $z_0\in\{\lambda_2,\dots,\lambda_k\}$ which is a contradiction. Hence r has no zeros and thus must be constant.

Elementary theory of polynomials implies that there are polynomials q_1,q_2 such that $p_1q_1+p_2q_2=1$ and thus $p_1q_1(f)+p_2q_2(f)=\operatorname{id}$. Now define $\pi_i:=p_iq_i(f)$ for $i=1,2$, so by definition $\pi_1+\pi_2=\operatorname{id}$, i.e. $v=\pi_1(v)+\pi_2(v)$ for all $v\in V$. Moreover $\pi_1\circ\pi_2=p_1q_1p_2q_2(f)=p_1(f)\circ p_2(f)\circ q_1q_2(f)$, and $p_1(f)\circ p_2(f)=0$, so $\pi_1\circ\pi_2=0$ and also the other composition vanishes. Hence $\pi_1(v) \in \ker(\pi_2)$ for all $v \in V$ and similarly $\pi_2(v) \in \ker(\pi_1)$ and from above we see that $\ker(\pi_1) \cap \ker(\pi_2) = \{0\}$. Hence we see that $V = \ker(\pi_1) \oplus \ker(\pi_2)$, and the projections onto the two factors are just the maps π_i which are polynomials in f. Since $\pi_1 = q_1(f) \circ p_1(f)$ we get $V_{(\lambda_1)} = \ker(p_1(f)) \subset \ker(\pi_1)$. Conversely, if $\pi_1(v) = 0$, then $v = \pi_2(v)$ and then $p_1(f)(v) = p_1(f) \circ p_2(f) \circ q_2(f)(v) = 0$, so we conclude that $\ker(\pi_1) = V_{(\lambda_1)}$. The projection onto this summand is given by π_2, which is a polynomial in f.

Next, if $v \in W := \ker(\pi_2)$, the $\pi_2(f(v)) = f(\pi_2(v)) = 0$ and hence $f(W) \subset W$, so the decomposition $V = V_{(\lambda_1)} \oplus W$ is compatible with f. But this immediately implies that p_f is the product of p_{f_1} and p_{f_2}, where the f_i are the restrictions of f to the two summands. Hence $p_{f_2} = p_2$, so f_2 has eigenvalues $\lambda_2, \dots, \lambda_k$ and it is easy to see that the generalized eigenspace of f_2 are exactly the generalized eigenspace of f with respect to these $k-1$ eigenvalues. Applying the induction hypothesis to f_2, we see that $W = V_{(\lambda_2)} \oplus \cdots \oplus V_{(\lambda_k)}$ and the projections onto the summands are polynomials in f_2, hence polynomials in $f \circ \pi_2$ and the claim follows.

Now denote by π_i the projection onto $V_{(\lambda_i)}$ and define $f_S := \sum_{i=1}^k \lambda_i \pi_i$. Then of course, f_S is diagonalizable with eigenvalues λ_i and eigenspaces $V_{(\lambda_i)}$. Then both f_S and $f_N := f - f_S$ are polynomials in f and thus commute with each other, so it remains to show that f_N is nilpotent. Since any $v \in V$ can be written as a linear combination of elements of the spaces $V_{(\lambda_i)}$, it suffices to verify this for $v \in V_{(\lambda_i)}$. But by construction, for such elements we have $f_S(v) = \lambda_i v$ so $(f - f_S)(v) = (f - \lambda_i \,\mathrm{id})(v)$ and by construction $(f - \lambda_i \,\mathrm{id})^{n_i}(v) = 0$, which completes the proof of existence.

Uniqueness: Suppose that $f = f_1 + f_2$ is any Jordan decomposition. Let V_{μ_i} be the f_1–eigenspace with eigenvalue μ_i. For $v \in V_{\mu_i}$ we have $f_1(f_2(v)) = f_2(f_1(v)) = \mu_i(f_2(v))$, so $f_2(V_{\mu_i}) \subset V_{\mu_i}$. By part (1) of proposition 2.2 we find a basis of each V_{μ_i} such that the restriction of f_2 to V_{μ_i} is represented by a strictly upper triangular matrix. Since f_1 is diagonalizable, the union of the bases obtained in that way is a basis for V. The matrix representation of $f = f_1 + f_2$ with respect is block diagonal, with each block upper triangular with all diagonal entries equal to one μ_i. Consequently, we see that the eigenvalues of f are exactly the μ_i (including the multiplicities) and moreover V_{μ_i} is exactly the generalized eigenspace of f corresponding to that eigenvalue. But this implies that f_1 equals the map f_S from above and thus $f_S + f_N = f = f_S + f_2$ and so $f_2 = f_N$. □

Remarks. (1) The parts f_S and f_N in the Jordan decomposition of f are referred to as the *semisimple* and the *nilpotent* part of f.
(2) The remaining step to get from this result to the Jordan normal form is to find the normal form for nilpotent linear maps, which is a refinement of part (1) of proposition 2.2.

Now the result we will need to prove Cartan's criteria is the following:

Lemma. *Let V be a finite dimensional complex vector space and let $X \in \mathfrak{gl}(V)$ be a linear mapping with Jordan decomposition $X = X_S + X_N$. Then the Jordan decomposition of* $\mathrm{ad}(X) : \mathfrak{gl}(V) \to \mathfrak{gl}(V)$ *is given by* $\mathrm{ad}(X) = \mathrm{ad}(X_S) + \mathrm{ad}(X_N)$.

Proof. For $X_1, X_2 \in \mathfrak{gl}(V)$ we have $[\mathrm{ad}(X_1), \mathrm{ad}(X_2)] = \mathrm{ad}([X_1, X_2])$, so for commuting linear maps also the adjoint actions commute. In particular, $\mathrm{ad}(X_S)$ and $\mathrm{ad}(X_N)$ commute, so it suffices to verify that $\mathrm{ad}(X_S)$ is diagonalizable and $\mathrm{ad}(X_N)$ is nilpotent. The proof that nilpotency of X_N implies nilpotency of $\mathrm{ad}(X_N)$ is exactly as the argument for matrices done in the proof of Engel's theorem in 2.2.

To prove that $\mathrm{ad}(X_S)$ is diagonalizable, we choose a basis of V consisting of eigenvectors for X_S and work with matrix representations in that basis. By construction, the matrix corresponding to X_S is diagonal in that basis, and we denote by λ_i the (not necessarily different) diagonal entries. Then let E_{ij} be the elementary matrix introduced in example (1) of 2.3, i.e. the only nonzero entry of E_{ij} is a one in the jth column of the ith row. Visibly, the product $X_S E_{ij}$ equals $\lambda_i E_{ij}$ while $E_{ij} X_S = \lambda_j E_{ij}$. Thus we see that any elementary matrix is an eigenvalue for $\mathrm{ad}(X_S)$ with eigenvalue $\lambda_i - \lambda_j$, so in particular we have obtained a basis for $\mathfrak{gl}(V)$ consisting of eigenvectors for $\mathrm{ad}(X_S)$. □

Some basic results on semisimple Lie algebras

2.8. Cartan's criteria for solvability and semisimplicity. The characterization of semisimple Lie algebras via their Killing form is an essential tool, which will quickly lead to the proofs of the main general results on semisimple Lie algebras. The term "Cartan's criterion" is used for all three parts of the following theorem.

THEOREM. *(1) Let V be a vector space and $\mathfrak{g} \subset \mathfrak{gl}(V)$ be a Lie subalgebra. If B_V is zero on $\mathfrak{g}$, then $\mathfrak{g}$ is solvable.*
(2) A Lie algebra $\mathfrak{g}$ is solvable if and only if its Killing form satisfies $B(\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}]) = 0$.
(3) A Lie algebra $\mathfrak{g}$ is semisimple if and only if its Killing form is non–degenerate.

PROOF. (1) Complexifying first V, we can view $\mathfrak{g}$ as a subalgebra of the Lie algebra of complex linear maps on $V_{\mathbb{C}}$, and then the complexification $\mathfrak{g}_{\mathbb{C}}$ of $\mathfrak{g}$ is a complex subalgebra in there. Since solvability of $\mathfrak{g}_{\mathbb{C}}$ implies solvability of $\mathfrak{g}$ (see 2.5) we may without loss of generality assume that V is complex $\mathfrak{g}$ is a complex Lie subalgebra of $\mathfrak{gl}(V)$.

Now it suffices to show that $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent, see 2.2, and in view of part (1) of Corollary 2.2 we can prove this by showing that any $X \in [\mathfrak{g}, \mathfrak{g}]$ acts as a nilpotent linear map on V. For $X \in [\mathfrak{g}, \mathfrak{g}]$ let X_S be the semisimple part in the Jordan decomposition. This is diagonalizable, and we define $\bar{X}_S : V \to V$ to be the linear map which has the same eigenspaces but conjugate eigenvalues, i.e. $\bar{X}_S$ acts by multiplication by $\bar{\lambda}$ on the λ–eigenspace of X_S. Choosing a basis of V with respect to which X_S is diagonal and X_N is strictly upper triangular, we see that $\mathrm{tr}(\bar{X}_S \circ X) = \sum_i |\lambda_i|^2$, where the sum is over all (not necessarily different) eigenvalues λ_i of X. Hence we may complete the proof by showing that this trace has to vanish. The tricky point about this is that in general neither X_S nor $\bar{X}_S$ lies in $\mathfrak{g}$.

By lemma 2.7, $\mathrm{ad}(X_S)$ is the semisimple part in the Jordan decomposition of the linear map $\mathrm{ad}(X) : \mathfrak{gl}(V) \to \mathfrak{gl}(V)$. This semisimple part can be written as a polynomial in $\mathrm{ad}(X)$, which implies that $\mathrm{ad}(X_S)(\mathfrak{g}) \subset \mathfrak{g}$. On the other hand, from the proof of Lemma 2.7 we see that starting from a basis of V consisting of eigenvectors of X_S the corresponding elementary matrices are eigenvectors for $\mathrm{ad}(X_S)$ with eigenvalue given as the difference of two eigenvalues of X_S. This implies that $\mathrm{ad}(\bar{X}_S)$ is diagonalizable with the same eigenspaces as $\mathrm{ad}(X_S)$ but conjugate eigenvalues. Since the projections onto the eigenspaces of $\mathrm{ad}(X_S)$ are polynomials in $\mathrm{ad}(X_S)$ we see that also $\mathrm{ad}(\bar{X}_S)$ is a polynomial in $\mathrm{ad}(X_S)$. Hence $\mathrm{ad}(X_S)(\mathfrak{g}) \subset \mathfrak{g}$ implies that $\mathrm{ad}(\bar{X}_S)(\mathfrak{g}) \subset \mathfrak{g}$.

Now since $X \in [\mathfrak{g}, \mathfrak{g}]$, we can write it as a finite sum $X = \sum_i [Y_i, Z_i]$. But then

$$\mathrm{tr}(\bar{X}_S \circ X) = \sum_i \mathrm{tr}(\bar{X}_S \circ [Y_i, Z_i]) = \sum_i \mathrm{tr}([\bar{X}_S, Y_i] \circ Z_i) = \sum_i B_V(\mathrm{ad}(\bar{X}_s)(Y_i), Z_i) = 0.$$

(2) We have seen the necessity of the condition already in the end of 2.6. Conversely, we show that even $B([\mathfrak{g}, \mathfrak{g}], [\mathfrak{g}, \mathfrak{g}]) = 0$ implies solvability of $\mathfrak{g}$. Indeed, by (1) this implies

that the image of $[\mathfrak{g},\mathfrak{g}]$ under the adjoint representation is solvable. Since the kernel of the adjoint representation of $[\mathfrak{g},\mathfrak{g}]$ is simply the center of $[\mathfrak{g},\mathfrak{g}]$, which is an Abelian (and hence solvable) ideal, we conclude that $[\mathfrak{g},\mathfrak{g}]$ is solvable. Since the quotient of $\mathfrak{g}$ by the solvable ideal $[\mathfrak{g},\mathfrak{g}]$ is Abelian, we conclude that $\mathfrak{g}$ is solvable.
(3) For semisimple $\mathfrak{g}$ consider the null space $\mathfrak{h} := \{X \in \mathfrak{g} : B(X,Y) = 0 \quad \forall Y \in \mathfrak{g}\}$ of the Killing form. By invariance of the Killing form, this is an ideal in $\mathfrak{g}$, and by (1) the image $\operatorname{ad}(\mathfrak{h}) \subset \mathfrak{gl}(\mathfrak{g})$ is solvable. Since the kernel of the adjoint map is $\mathfrak{h} \cap \mathfrak{z}(\mathfrak{g})$ and thus an Abelian ideal in $\mathfrak{h}$, we conclude that $\mathfrak{h}$ is solvable, and thus $\mathfrak{h} = \{0\}$.

Conversely, let us assume that B is non–degenerate and that $\mathfrak{h} \subset \mathfrak{g}$ is an Abelian ideal. For $X \in \mathfrak{h}$ and $Y \in \mathfrak{g}$, we see that $\operatorname{ad}(Y) \circ \operatorname{ad}(X)$ maps $\mathfrak{g}$ to $\mathfrak{h}$ and $\mathfrak{h}$ to zero, so by part (1) of Proposition 2.2 this map is nilpotent and thus tracefree. Hence X lies in the null space of B, so $X = 0$. Since $\mathfrak{g}$ has no nontrivial Abelian ideal, we know from 2.3 that $\mathfrak{g}$ is semisimple. □

Next we note several important consequences of this result. In particular, as promised in 2.5 we show that semisimplicity is well behaved with respect to complexification and we prove that the study of semisimple Lie algebras reduces to the study of simple Lie algebras.

Corollary. *(1) If $\mathfrak{g}$ is a semisimple Lie algebra, then there are simple ideals $\mathfrak{g}_1, \dots, \mathfrak{g}_k$ in $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{g}_1 \oplus \dots \oplus \mathfrak{g}_k$ as a Lie algebra. This decomposition is unique and any ideal in $\mathfrak{g}$ is equal to the sum of some of the $\mathfrak{g}_i$. In particular, $\mathfrak{g} = [\mathfrak{g},\mathfrak{g}]$ and any ideal in $\mathfrak{g}$ as well as any image of $\mathfrak{g}$ under a Lie algebra homomorphism is semisimple.*
(2) A real Lie algebra $\mathfrak{g}$ is semisimple if and only if its complexification $\mathfrak{g}_{\mathbb{C}}$ is semisimple.
(3) If $\mathfrak{g}$ is a complex simple Lie algebra and $\Phi : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ is an invariant complex bilinear form, then Φ is a multiple of the Killing form. In particular, if Φ is nonzero, then it is automatically symmetric and non–degenerate.

Proof. (1) If $\mathfrak{h} \subset \mathfrak{g}$ is an ideal, then the annihilator $\mathfrak{h}^\perp := \{Y \in \mathfrak{g} : B(X,Y) = 0 \quad \forall X \in \mathfrak{h}\}$ is an ideal, too. Indeed for $Y \in \mathfrak{h}^\perp$, $Z \in \mathfrak{g}$ and $X \in \mathfrak{h}$ we have $B([Y,Z],X) = B(Y,[Z,X])$ which vanishes since $[Z,X] \in \mathfrak{h}$. Moreover, since the Killing form restricts to zero on the ideal $\mathfrak{h} \cap \mathfrak{h}^\perp$, this ideal is solvable, so $\mathfrak{h} \cap \mathfrak{h}^\perp = \{0\}$, and thus $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^\perp$. In particular, $[\mathfrak{h},\mathfrak{h}^\perp] = \{0\}$, and thus the Killing form of $\mathfrak{h}$ is the restriction of the Killing form of $\mathfrak{g}$ and hence is non–degenerate. Thus $\mathfrak{h}$ is semisimple and the result follows by induction. The fact that $\mathfrak{g} = [\mathfrak{g},\mathfrak{g}]$ follows immediately, since $\mathfrak{g}_i = [\mathfrak{g}_i,\mathfrak{g}_i]$ for any of the simple ideals.

Since we have seen that any ideal in $\mathfrak{g}$ is semisimple, the statement about the ideals follows now if we prove that any simple ideal $\mathfrak{h} \lhd \mathfrak{g}$ equals one of the $\mathfrak{g}_i$. Since $\mathfrak{h} \cap \mathfrak{g}_i$ is an ideal both in $\mathfrak{h}$ and in $\mathfrak{g}_i$ it must either equal $\mathfrak{h}$ and $\mathfrak{g}_i$ or be zero. But if $\mathfrak{h} \cap \mathfrak{g}_i = \{0\}$ for all i, then $[\mathfrak{g},\mathfrak{h}] = \{0\}$, but we know that $[\mathfrak{h},\mathfrak{h}] = \mathfrak{h}$, so $\mathfrak{h} = \{0\}$. Thus the statement about ideals as well as the uniqueness of the decomposition follows. If φ is a homomorphism from $\mathfrak{g}$ to some Lie algebra, then its kernel is an ideal and thus the sum of some of the $\mathfrak{g}_i$. But then the image is isomorphic to the sum of the remaining simple ideals and hence semisimple.
(2) From 2.5 we know that $\mathfrak{g}$ is semisimple if $\mathfrak{g}_{\mathbb{C}}$ is semisimple. But the converse now immediately follows from part (3) of the theorem and the fact that the Killing form of $\mathfrak{g}_{\mathbb{C}}$ is just the complex bilinear extension of the Killing form of $\mathfrak{g}$.
(3) The adjoint representation of $\mathfrak{g}$ is a complex representation which is irreducible, since a $\mathfrak{g}$–invariant subspace in $\mathfrak{g}$ by definition is an ideal in $\mathfrak{g}$. Now a complex bilinear form Φ :

$\mathfrak{g}\times\mathfrak{g} \to \mathbb{C}$ induces a linear map $\Phi^\vee : \mathfrak{g} \to \mathfrak{g}^* = L(\mathfrak{g}, \mathbb{C})$ defined by $\Phi^\vee(X)(Y) := \Phi(X, Y)$. Invariance of Φ reads as $\Phi([Z, X], Y) = -\Phi(X, [Z, Y])$ which means that $\Phi^\vee(\mathrm{ad}_Z(X)) = -(\mathrm{ad}_Z)^* \circ \Phi^\vee(X)$, i.e. that $\Phi^\vee$ is a homomorphism of $\mathfrak{g}$–modules. Non–degeneracy of B implies that $B^\vee$ in injective and thus an isomorphism. Hence $(B^\vee)^{-1} \circ \Phi^\vee$ is a homomorphism from the irreducible representation $\mathfrak{g}$ to itself. Now the result follows from Schur's Lemma in 2.4. □

2.9. Complete reducibility. The next essential application of the Killing form is to prove that any finite dimensional representation of a semisimple Lie algebra is completely reducible, so irreducible representations are the ones of central interest. We have already observed in 2.4 that in order to prove this, it suffices to show that any invariant subspace in a representation admits an invariant complement.

THEOREM. *Let $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ be a finite dimensional representation of a semisimple Lie algebra $\mathfrak{g}$. Then any invariant subspace $W \subset V$ has an invariant complement.*

PROOF. Since by corollary 2.7(1) $\rho(\mathfrak{g}) \subset \mathfrak{gl}(V)$ is semisimple, so we may assume that $\mathfrak{g} \subset \mathfrak{gl}(V)$. As in the proof of part (3) of theorem 2.7, the null space of the form B_V would be a solvable ideal in $\mathfrak{g}$, so B_V is non–degenerate. Now let $\{X_1, \dots, X_n\}$ be a basis of $\mathfrak{g}$, and let $\{Y_1, \dots, Y_n\}$ be the dual basis with respect to B_V, i.e. $B_V(X_i, Y_j) = \delta_{ij}$, the Kronecker delta. By definition, any $X \in \mathfrak{g}$ can be written as $\sum a_i Y_i$ and then $B_V(X, X_j) = a_j$, so we see that $X = \sum B(X, X_i)Y_i$. In the same way, we get $X = \sum B(X, Y_i)X_i$.

Define the *Casimir operator* $C_V : V \to V$ by $C_V := \sum_{i=1}^n Y_i \circ X_i$. For $X \in \mathfrak{g}$, we directly get $X \circ C_V - C_V \circ X = \sum_i([X, Y_i] \circ X_i + Y_i \circ [X, X_i])$. But by construction, $[X, Y_i] = \sum_j B_V([X, Y_i], X_j)Y_j$, which by invariance of B_V equals $-\sum_j B_V(Y_i, [X, X_j])Y_j$, while $[X, X_i] = \sum_j B(Y_j, [X, X_i])X_j$. Inserting these, we get $X \circ C_V - C_V \circ X = 0$, so C_V commutes with the action of any element of $\mathfrak{g}$. Moreover, by construction $\mathrm{tr}(C_V) = \sum_i \mathrm{tr}(Y_i \circ X_i) = \sum_i B_V(Y_i, X_i) = n = \dim(\mathfrak{g})$, and by construction for any $\mathfrak{g}$–invariant subspace $W \subset V$, we have $C_V(W) \subset W$.

Step 1: Assume that W is irreducible and has codimension one, i.e. $\dim(W) = \dim(V) - 1$. In this case, the action of $\mathfrak{g}$ on V/W must be trivial since $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$. So in particular, C_V acts trivially on V/W. But this means that $C_V(V) \subset W$, and thus C_V must have non–trivial kernel. On the other hand, since W is irreducible, Schur's lemma (see 2.4) implies that C_V acts by a scalar on W and this scalar must be nonzero, since otherwise $C_V \circ C_V = 0$ which contradicts $\mathrm{tr}(C_V) \neq 0$. Thus, $\ker(C_V)$ is a complementary subspace to W, and since C_V commutes with the action of any element of $\mathfrak{g}$, this subspace is $\mathfrak{g}$–invariant.

Step 2: Assume that W is not irreducible but still of codimension one. Then we use induction on $\dim(W)$. If $Z \subset W$ is a nontrivial invariant subspace, then either by step 1 (if W/Z is irreducible) or by the induction hypothesis we find a complement of W/Z in V/Z, whose preimage in V we denote by Y. Then $Y \subset V$ is a representation of $\mathfrak{g}$ and $Z \subset Y$ is an invariant subspace of codimension one, and $\dim(Z) < \dim(W)$. Either by step 1 (if Z is irreducible) or by induction hypothesis, we find an invariant subspace U such that $Y = U \oplus Z$ and hence $V = U \oplus W$.

Step 3: Assume that W is irreducible but of arbitrary dimension. Consider the set of all linear maps $\varphi : V \to W$ whose restriction to W is a scalar multiple of the identity. Now $L(V, W)$ is naturally a representation of $\mathfrak{g}$ with the action $(X \cdot \varphi)(v) = X\varphi(v) - \varphi(Xv)$. If φ restricts to a multiple of the identity on W, then obviously $X \cdot \varphi$ restricts to zero on W, so these maps also form a representation of $\mathfrak{g}$. Moreover, the subspace of all φ which restrict to zero on W is visibly invariant, and it clearly has codimension one. From step 2 we know that there is a complementary submodule, which must be trivial by construction. Taking a nonzero element in this complement, its restriction to W is a nonzero multiple of the identity, so dividing by that factor we obtain a map π in the complement such that $\pi|_W = \mathrm{id}_W$. Since the complement is a trivial representation of $\mathfrak{g}$, we see that $X\pi(v) = \pi(Xv)$ for all $X \in \mathfrak{g}$, and in particular $\ker(\pi) \subset V$ is an invariant subspace. Since by construction $\ker(\pi) \cap W = \{0\}$ and $\dim(\ker(\pi)) = \dim(V) - \dim(W)$, we see that $V = \ker(\pi) \oplus W$, so we have found an invariant complement.

The final step to deal with an arbitrary invariant subspace is now done by induction exactly as step 2. □

As a corollary, we can clarify the structure of reductive Lie algebras and prove a result on derivations of semisimple Lie algebras. A *derivation* of a Lie algebra $\mathfrak{g}$ is a linear map $D : \mathfrak{g} \to \mathfrak{g}$ such that $D([X, Y]) = [D(X), Y] + [X, D(Y)]$ for all $X, Y \in \mathfrak{g}$. The Jacobi identity exactly says that for any $X \in \mathfrak{g}$ the map $\mathrm{ad}(X)$ is a derivation of $\mathfrak{g}$. Derivations of this form are called *inner derivations*. One immediately verifies that the commutator of two derivations is again a derivation, so the vector space $\mathfrak{Der}(\mathfrak{g})$ of all derivations of $\mathfrak{g}$ is a Lie subalgebra of $\mathfrak{gl}(\mathfrak{g})$. Note that the derivation property can be rewritten as $[D, \mathrm{ad}(X)] = \mathrm{ad}(D(X))$ (where the bracket is in $\mathfrak{gl}(\mathfrak{g})$), which shows that the inner derivations form an ideal $\mathrm{ad}(\mathfrak{g})$ in $\mathfrak{Der}(\mathfrak{g})$.

One of the reasons for the importance of derivations is their relation to automorphisms of $\mathfrak{g}$, i.e. Lie algebra isomorphisms from $\mathfrak{g}$ to itself. Obviously, the set $\mathrm{Aut}(\mathfrak{g})$ of all automorphisms of $\mathfrak{g}$ is a subgroup of the group $GL(\mathfrak{g})$ of all linear isomorphisms from $\mathfrak{g}$ to itself. By definition, an invertible linear map $\varphi : \mathfrak{g} \to \mathfrak{g}$ is an automorphism if an only if $\varphi([X, Y]) = [\varphi(X), \varphi(Y)]$ for all $X, Y \in \mathfrak{g}$. This obviously implies that $\mathrm{Aut}(\mathfrak{g})$ is a closed subset of $GL(\mathfrak{g})$ and thus a matrix group as defined in 1.4. It turns out that the Lie algebra of this matrix group is exactly $\mathfrak{Der}(\mathfrak{g}) \subset \mathfrak{gl}(\mathfrak{g})$.

COROLLARY. *(1) Let $\mathfrak{g}$ be a reductive Lie algebra. Then $[\mathfrak{g}, \mathfrak{g}]$ is semisimple and $\mathfrak{g} = \mathfrak{z}(\mathfrak{g}) \oplus [\mathfrak{g}, \mathfrak{g}]$ as a Lie algebra.*
(2) For a semisimple Lie algebra $\mathfrak{g}$, the adjoint action is an isomorphism of Lie algebras from $\mathfrak{g}$ onto $\mathfrak{Der}(\mathfrak{g})$. In particular, any derivation of a semisimple Lie algebra is inner.

PROOF. (1) Consider the image $\mathrm{ad}(\mathfrak{g})$ of $\mathfrak{g}$ under the adjoint representation. Since the kernel of the adjoint representation is exactly the center $\mathfrak{z}(\mathfrak{g})$, we see that $\mathrm{ad}(\mathfrak{g}) \cong \mathfrak{g}/\mathfrak{z}(\mathfrak{g})$. Suppose that $\mathfrak{a}$ is a solvable ideal in $\mathrm{ad}(\mathfrak{g})$, and let $\tilde{\mathfrak{a}}$ be its preimage in $\mathfrak{g}$. Since $\mathrm{ad}([X, Y]) = [\mathrm{ad}(X), \mathrm{ad}(Y)]$ this is an ideal in $\mathfrak{g}$ and it is solvable since it contains the Abelian ideal $\mathfrak{z}(\mathfrak{g})$ and $\tilde{\mathfrak{a}}/\mathfrak{z}(\mathfrak{g}) \cong \mathfrak{a}$ is solvable. Since $\mathfrak{g}$ is reductive, we must have $\tilde{\mathfrak{a}} \subset \mathfrak{z}(\mathfrak{g})$, so $\mathfrak{a} = \{0\}$ and thus $\mathrm{ad}(\mathfrak{g})$ is semisimple. Applying the above theorem to the adjoint representation of $\mathfrak{g}$, we obtain a complementary subspace $\mathfrak{h}$ to the invariant subspace $\mathfrak{z}(\mathfrak{g})$. Invariance of $\mathfrak{h}$ means $[\mathfrak{g}, \mathfrak{h}] \subset \mathfrak{h}$, so $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. Moreover $\mathfrak{h} \cong \mathrm{ad}(\mathfrak{g})$ is semisimple and $\mathfrak{g} = \mathfrak{z}(\mathfrak{g}) \oplus \mathfrak{h}$ as a Lie algebra.

By construction we have $[\mathfrak{g}, \mathfrak{g}] = [\mathfrak{g}, \mathfrak{h}] \subset \mathfrak{h}$, but since $\mathfrak{h}$ is semisimple we must even have $[\mathfrak{h}, \mathfrak{h}] = \mathfrak{h}$ and thus also $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{h}$.
(2) Since $\mathfrak{g}$ is semisimple, we have $\mathfrak{z}(\mathfrak{g}) = \{0\}$, so the adjoint representation is faithful, and $\mathfrak{g} \cong \mathrm{ad}(\mathfrak{g})$. From above we know that $\mathrm{ad}(\mathfrak{g}) \subset \mathfrak{Der}(\mathfrak{g}) \subset \mathfrak{gl}(\mathfrak{g})$ are $\mathfrak{g}$–subrepresentations, so by the theorem there exists a $\mathfrak{g}$–invariant complement U to $\mathrm{ad}(\mathfrak{g})$ in $\mathfrak{Der}(\mathfrak{g})$. But from above we know that $\mathrm{ad}(\mathfrak{g})$ is an ideal in $\mathfrak{Der}(\mathfrak{g})$, so this complementary representation must be trivial. But for $D \in U$ this implies $0 = [D, \mathrm{ad}(X)] = \mathrm{ad}(D(X))$

for all $X \in \mathfrak{g}$ and by injectivity of the adjoint action this gives $D(X) = 0$ for all $X \in \mathfrak{g}$, so $D = 0$. □

Remark. There is a weaker analog of part (1) of the corollary for an arbitrary Lie algebra $\mathfrak{g}$, which is called the *Levi decomposition*. The first step towards this is to observe that any Lie algebra $\mathfrak{g}$ contains a maximal solvable ideal $\mathfrak{rad}(\mathfrak{g})$, called the *radical* of $\mathfrak{g}$. To see this, one first observes that if $\mathfrak{a}$ and $\mathfrak{b}$ are ideals in $\mathfrak{g}$, then the subspace $\mathfrak{a} + \mathfrak{b}$ spanned by the two ideals is an ideal in $\mathfrak{g}$, too. Moreover, $\mathfrak{a} \cap \mathfrak{b}$ is an ideal in both $\mathfrak{a}$ and $\mathfrak{b}$, and clearly $(\mathfrak{a}+\mathfrak{b})/\mathfrak{a} \cong \mathfrak{b}/(\mathfrak{a}\cap\mathfrak{b})$. If $\mathfrak{a}$ and $\mathfrak{b}$ are solvable, then so is $\mathfrak{a}\cap\mathfrak{b}$, so we conclude from the above isomorphism that also $\mathfrak{a}+\mathfrak{b}$ is solvable. Iterating this argument, we see that the subspace spanned by all solvable ideals of $\mathfrak{g}$ is itself a solvable ideal, which by construction is maximal.

It is almost a tautology that the quotient $\mathfrak{g}/\mathfrak{rad}(\mathfrak{g})$ is a semisimple Lie algebra. (As in the proof of part (1) of the corollary, the preimage in $\mathfrak{g}$ of a solvable ideal in the quotient is a solvable ideal and thus contained in the radical.) The Levi–decomposition is then achieved by finding a subalgebra $\mathfrak{l} \subset \mathfrak{g}$ which is complementary to the radical, so $\mathfrak{g} = \mathfrak{l} \oplus \mathfrak{rad}(\mathfrak{g})$. The proof of existence of $\mathfrak{l}$ is similar to a combination of part (1) of the corollary and step 3 of the theorem, see [**Fulton-Harris**, Appendix E].

2.10. Examples of reductive and semisimple Lie algebras. We next describe an general result which exhibits certain matrix Lie algebras as being reductive. Since verifying triviality of the center of a Lie algebra is usually rather simple, this result provides an efficient criterion for semisimplicity, too. Let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$. One $\mathbb{R}$ we define the conjugation to be the identity, while on $\mathbb{C}$ we consider the usual conjugation $\overline{a+ib} = a - ib$. In any case, we have $\overline{xy} = \bar{x}\bar{y}$ and mapping (x, y) to the real part of $x\bar{y}$ defines a positive definite real inner product on $\mathbb{K}$. Next, for $n \in \mathbb{N}$ let us consider the space $M_n(\mathbb{K})$ of $n \times n$–matrices with entries from $\mathbb{K}$. For $A \in M_n(\mathbb{K})$ let A^* be the conjugate transpose of A, i.e. $(a_{ij})^* = (\bar{a}_{ji})$. Then $(AB)^* = B^*A^*$ and we may consider the bilinear map $(A, B) \mapsto \mathrm{Re}(\mathrm{tr}(AB^*))$. For $A = (a_{ij})$ and $B = (B_{ij})$ the i–th diagonal entry of AB^* is just $\sum_j a_{ij}\bar{b}_{ij}$, and forming the trace we also have to sum over all i. Hence $(A, B) \mapsto \mathrm{Re}(\mathrm{tr}(AB^*))$ is just the standard extension to $M_n(\mathbb{K}) \cong \mathbb{K}^{n^2}$ of the inner product on $\mathbb{K}$ from above and thus a positive definite inner product on $M_n(\mathbb{K})$.

Proposition. *Let $\mathfrak{g} \subset M_n(\mathbb{K})$ be a Lie subalgebra such that for all $X \in \mathfrak{g}$ also $X^* \in \mathfrak{g}$. Then $\mathfrak{g}$ is reductive.*

Proof. As above $\langle X, Y\rangle := \mathrm{Re}(\mathrm{tr}(XY^*))$ defines a positive definite real inner product on $\mathfrak{g}$. Now suppose that $\mathfrak{h}$ is an ideal in $\mathfrak{g}$, and consider its orthogonal complement $\mathfrak{h}^\perp$. For $X \in \mathfrak{h}^\perp$, $Y \in \mathfrak{g}$ and $Z \in \mathfrak{h}$, we then have $\langle [X, Y], Z\rangle = \mathrm{Re}(\mathrm{tr}(XYZ^* - YXZ^*)) = \mathrm{Re}(\mathrm{tr}(X(YZ^* - Z^*Y)))$, and the last expression can be rewritten as $\langle X, -[Y^*, Z]\rangle$. By assumption, $Y^* \in \mathfrak{g}$, so $[Y^*, Z] \in \mathfrak{h}$, and hence the inner product is zero, which implies that $\mathfrak{h}^\perp$ is an ideal in $\mathfrak{g}$, too.

We now prove that if $\mathfrak{h}$ is solvable, then $\mathfrak{h} \subset \mathfrak{z}(\mathfrak{g})$ by induction on the length k of the derived series of $\mathfrak{h}$. Hence k is the largest number such that $\mathfrak{h}^{(k)} \neq \{0\}$. If k=1, then $\mathfrak{h}$ is Abelian and we get $[\mathfrak{h}, \mathfrak{h}^\perp] \subset \mathfrak{h} \cap \mathfrak{h}^\perp = \{0\}$, while $[\mathfrak{h}, \mathfrak{h}] = \{0\}$ by assumption. Since $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^\perp$ this implies $\mathfrak{h} \subset \mathfrak{z}(\mathfrak{g})$. If $k > 1$, then we may by induction hypothesis assume that $\mathfrak{h}^{(2)} \subset \mathfrak{z}(\mathfrak{g})$. Moreover, since $\mathfrak{h}^{(2)}$ is an ideal in $\mathfrak{g}$, we get $\mathfrak{g} = \mathfrak{h}^{(2)} \oplus (\mathfrak{h}^{(2)})^\perp$, which together with $\mathfrak{h}^{(2)} \subset \mathfrak{h}$ implies that $\mathfrak{h} = \mathfrak{h}^{(2)} \oplus \mathfrak{a}$, where $\mathfrak{a} = \mathfrak{h} \cap (\mathfrak{h}^{(2)})^\perp$. From above we know that $(\mathfrak{h}^{(2)})^\perp$ is an ideal in $\mathfrak{g}$ so also $\mathfrak{a}$ is an ideal in $\mathfrak{g}$. Since $\mathfrak{a}\cap\mathfrak{h}^\perp = \{0\}$ we conclude that $[\mathfrak{a}, \mathfrak{h}^\perp] = \{0\}$. On the other hand, $[\mathfrak{a}, \mathfrak{h}] \subset \mathfrak{a} \cap \mathfrak{h}^{(2)} = \{0\}$, so $\mathfrak{a} \subset \mathfrak{z}(\mathfrak{g})$ and thus $\mathfrak{h} \subset \mathfrak{z}(\mathfrak{g})$. □

Together with Corollary 2.9(1) this implies that a Lie subalgebra of $M_n(\mathbb{K})$ which is closed under conjugate transpose and has trivial center is semisimple. A useful observation in this direction is that one may often use Schur's Lemma 2.4 to conclude triviality of the center. In the case $\mathbb{K} = \mathbb{C}$, one has to show that the obvious representation of $\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{C})$ on $\mathbb{C}^n$ is irreducible to conclude that the center of $\mathfrak{g}$ coincides with the intersection of $\mathfrak{g}$ with complex multiples of the identity matrix, which usually is easy to determine. For $\mathbb{K} = \mathbb{R}$ one can often use the above argument after embedding $\mathfrak{g}$ into $M_n(\mathbb{C})$.

Using the proposition and the observation on irreducibility, we can now list a number of examples of reductive and semisimple Lie algebras. The first obvious class of examples are provided by skew symmetric matrices, i.e. matrices A such that $A^T = -A$. These are Lie subalgebras, since from $(AB)^T = B^T A^T$ one immediately concludes that $[A, B]^T = -[A^T, B^T]$. For $\mathbb{K} = \mathbb{C}$ we further have $A^* = \overline{(A^T)} = -\bar{A} = -(A^*)^T$. Consequently, we see that $\mathfrak{so}(n, \mathbb{K}) = \{X \in M_n(\mathbb{K}) : X^T = -X\}$ is a reductive Lie algebra for $\mathbb{K} = \mathbb{R}$ and $\mathbb{C}$. Obviously, $\mathfrak{so}(1, \mathbb{K}) = \{0\}$ while $\mathfrak{so}(2, \mathbb{K})$ is one–dimensional and thus Abelian.

For $n \geq 3$, the Lie algebras $\mathfrak{so}(n, \mathbb{K})$ are actually semisimple: Let us first discuss the case $\mathbb{K} = \mathbb{C}$. Obviously, any element from the standard basis of $\mathbb{C}^n$ can be mapped to any other element of that basis by a matrix in $\mathfrak{so}(n, \mathbb{C})$. On the other hand, a vector with more than one nonzero components can be mapped first to a linear combination of at most two of the basis vectors and then to a nonzero multiple of a basis vector. Thus we see that $\mathbb{C}^n$ is the only nonzero $\mathfrak{so}(n, \mathbb{C})$–invariant subspace of $\mathbb{C}^n$. Since $\mathfrak{so}(n, \mathbb{C})$ obviously contains no nonzero multiples of the identity, we see that it has trivial center and thus is semisimple. Viewing $\mathfrak{so}(n, \mathbb{R})$ as a subalgebra of $\mathfrak{gl}(n, \mathbb{C})$, we see that any complex subspace of $\mathbb{C}^n$ that is invariant under $\mathfrak{so}(n, \mathbb{R})$ must also be invariant under $\mathfrak{so}(n, \mathbb{C})$ (which visibly is the complexification) so semisimplicity follows.

The other obvious (even simpler) idea is to directly work with skew Hermitian matrices, i.e. matrices A such that $A^* = -A$, in the complex case. As above, $(AB)^* = B^* A^*$ implies that $[A, B]^* = -[A^*, B^*]$, so $\mathfrak{u}(n) := \{X \in M_n(\mathbb{C}) : X^* = -X\}$ is a Lie subalgebra. Since $\mathfrak{u}(n)$ is obviously closed under conjugate transpose, it is reductive. Moreover, $\mathfrak{u}(n)$ obviously contains $\mathfrak{so}(n, \mathbb{R})$, so we conclude that $\mathbb{C}^n$ is an irreducible representation of $\mathfrak{u}(n)$. Hence the center of $\mathfrak{u}(n)$ consists of all purely imaginary multiples of the identity matrix. The subalgebra $\mathfrak{su}(n) := \{X \in \mathfrak{u}(n) : \operatorname{tr}(X) = 0\}$ is a codimension one subspace and clearly $[\mathfrak{u}(n), \mathfrak{u}(n)] \subset \mathfrak{su}(n)$. Since $\mathfrak{z}(\mathfrak{u}(n))$ is one–dimensional, the two spaces have the same dimension ad hence coincide. In particular, the Lie algebra $\mathfrak{su}(n)$ is semisimple.

CHAPTER 3

Structure theory of complex semisimple Lie algebras

In this chapter we discuss the structure theory of complex semisimple Lie algebras, which also forms the basis for the representation theory of these algebras. Most of the theory can be reduced to the representation theory of $\mathfrak{sl}(2,\mathbb{C})$, which therefore is of fundamental importance.

Throughout the chapter we will emphasize those aspects of the theory which are needed to deal with concrete examples and only loosely discuss the questions related to the classification of complex semisimple Lie algebras.

Cartan subalgebras

3.1. Absolute Jordan decomposition. The first key point in the theory of complex semisimple Lie algebras is the notion of semisimple elements. Recall from 2.7 that the Jordan–decomposition of an endomorphism f of a finite dimensional complex vector space V decomposes f into a sum $f_S + f_N$ of a diagonalizable map f_S and a nilpotent map f_N. Both these maps can be expressed as polynomials in f and thus commute with f and with each other, and they are uniquely determined by this property. In terms of a basis of V such that f is in Jordan normal form, f_S is simply the diagonal part and f_N is the off–diagonal part. For semisimple Lie algebras, there is a universal version of the Jordan decomposition, which gives the Jordan decomposition in any finite dimensional representation. Before we can discuss this, we have to prove some background from linear algebra.

Lemma. *Let V be a finite dimensional vector space and $W \subset V$ a linear subspace.*
(1) If $f : V \to V$ is a diagonalizable linear map such that $f(W) \subset W$, then the restriction of f to W is diagonalizable, too.
(2) If $f : V \to V$ is any linear map with Jordan decomposition $f = f_S + f_N$ such that $f(W) \subset W$, then the Jordan decomposition of the restriction $f|_W : W \to W$ is given by $f|_W = f_S|_W + f_N|_W$.

Proof. (1) Let λ_i be the different eigenvalues of f and V_{λ_i} the corresponding eigenspaces. Diagonalizability of f implies that V is the direct sum of these eigenspaces. Thus any element $w \in W$ may be uniquely written as a sum $w = v_1 + \dots + v_k$ such that each v_i is an eigenvector of f with eigenvalue λ_i. In the proof of Proposition 2.7 we have observed that the projection π_i onto the eigenspace V_{λ_i} can be written as a polynomial in f. Thus $f(W) \subset W$ implies $\pi_i(W) \subset W$, so we conclude that $v_i = \pi_i(w)$ lies in W for each $i = 1, \dots, k$. Hence we conclude that W is the direct sum of the subspaces $W \cap V_{\lambda_i}$, which are exactly the eigenspaces of the restriction $f|_W$. But this of course implies that we can find a basis of W consisting of eigenvectors for $f|_W$, so f_W is diagonalizable.
(2) Since f_S and f_N are polynomials in f, they also map the subspace W to itself. Moreover, by part (1), the restriction $f_S|_W$ is diagonalizable, and since f_N is nilpotent, also $f_N|_W$ is nilpotent. Since f_S and f_N commute, also their restrictions to W commute,

so the claim follows from the uniqueness of the Jordan decomposition, see Proposition 2.7. □

THEOREM. *(1) Let V be a complex vector space, and $\mathfrak{g} \subset \mathfrak{gl}(V)$ be a semisimple Lie subalgebra. Then for any $X \in \mathfrak{g}$ the semisimple and nilpotent part of the Jordan decomposition of X (viewed as a linear map $V \to V$) lie in $\mathfrak{g}$.*
(2) [Absolute Jordan decomposition] Let $\mathfrak{g}$ be any semisimple Lie algebra. Then for any $X \in \mathfrak{g}$ there are unique elements $X_S, X_N \in \mathfrak{g}$ such that $\mathrm{ad}(X) = \mathrm{ad}(X_S) + \mathrm{ad}(X_N)$ is the Jordan decomposition of $\mathrm{ad}(X) : \mathfrak{g} \to \mathfrak{g}$.
(3) If $\mathfrak{g}$ is any semisimple Lie algebra and $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is a finite dimensional complex representation, then for any element $X \in \mathfrak{g}$, the Jordan decomposition of $\rho(X) : V \to V$ is given by $\rho(X) = \rho(X_S) + \rho(X_N)$.

PROOF. (1) This is done by writing $\mathfrak{g}$ in a smart way as an intersection of Lie subalgebras of $\mathfrak{gl}(V)$: First define $\mathfrak{n} := \{A \in \mathfrak{gl}(V) : [A, X] \in \mathfrak{g} \quad \forall X \in \mathfrak{g}\}$ (the *normalizer* of $\mathfrak{g}$ in $\mathfrak{gl}(V)$). By the Jacobi identity, $\mathfrak{n}$ is a Lie subalgebra of $\mathfrak{gl}(V)$ and of course, $\mathfrak{g} \subset \mathfrak{n}$. Moreover, if $A \in \mathfrak{n}$ and $A = S + N$ is the Jordan decomposition of the linear map $A : V \to V$, then by Lemma 2.7 $\mathrm{ad}(A) = \mathrm{ad}(S) + \mathrm{ad}(N)$ is the Jordan decomposition of $\mathrm{ad}(A) : \mathfrak{gl}(V) \to \mathfrak{gl}(V)$. In particular, $\mathrm{ad}(S)$ and $\mathrm{ad}(N)$ are polynomials in $\mathrm{ad}(A)$ so $\mathrm{ad}(A)(\mathfrak{g}) \subset \mathfrak{g}$ implies $\mathrm{ad}(S)(\mathfrak{g}) \subset \mathfrak{g}$ and $\mathrm{ad}(N)(\mathfrak{g}) \subset \mathfrak{g}$. Thus we see that $A \in \mathfrak{n}$ implies $S, N \in \mathfrak{n}$.

Next, for any $\mathfrak{g}$–invariant subspace $W \subset V$ define $\mathfrak{s}_W := \{A \in \mathfrak{gl}(V) : A(W) \subset W \text{ and } \mathrm{tr}(A|_W) = 0\}$. Of course, $X(W) \subset W$ for each $X \in \mathfrak{g}$, and since $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ we conclude that $X|_W$ can be written as a sum of restrictions of commutators to W, and these are tracefree. Hence we conclude that $\mathfrak{g} \subset \mathfrak{s}_W$, and obviously $\mathfrak{s}_W$ is a subalgebra of $\mathfrak{gl}(V)$. Moreover, if $A = S + N$ is the Jordan decomposition of $A \in \mathfrak{s}_W$, viewed as a linear map $V \to V$, we have $S(W) \subset W$ and $N(W) \subset W$ since S and N are polynomials in A. Since N is nilpotent also the restriction $N|_W$ is nilpotent and thus tracefree, so $N \in \mathfrak{s}_W$ and thus also $S = A - N \in \mathfrak{s}_W$.

Now let $\mathfrak{g}'$ denote the intersection of $\mathfrak{n}$ and the subalgebras $\mathfrak{s}_W$ for all $\mathfrak{g}$–invariant subspaces $W \subset V$. As an intersection of subalgebras, this is a Lie subalgebra of $\mathfrak{gl}(V)$ and from above we see that $\mathfrak{g} \subset \mathfrak{g}'$ and for $A \in \mathfrak{g}'$ with Jordan decomposition $A = S+N$ we have $S, N \in \mathfrak{g}'$. Thus we may conclude the proof by showing that $\mathfrak{g}' = \mathfrak{g}$. Since $\mathfrak{g}$ is a subalgebra of $\mathfrak{g}'$ we have $\mathrm{ad}(X)(\mathfrak{g}') \subset \mathfrak{g}'$ for all $X \in \mathfrak{g}$, so $\mathfrak{g}'$ is a representation of $\mathfrak{g}$ and $\mathfrak{g} \subset \mathfrak{g}'$ is an invariant subspace. By Theorem 2.9 there is an invariant complement $U \subset \mathfrak{g}'$ to $\mathfrak{g}$. Moreover, since $\mathfrak{g}' \subset \mathfrak{n}$ we see that $[\mathfrak{g}, \mathfrak{g}'] \subset \mathfrak{g}$, so the $\mathfrak{g}$–action on U must be trivial. Let $Y \in U$ be any element. Since V is a direct sum of irreducible $\mathfrak{g}$–representations by Theorem 2.9, it suffices to show that $Y|_W = 0$ for any $\mathfrak{g}$–irreducible subspace $W \subset V$ to conclude that $Y = 0$. But if W is such a subspace, the by construction $Y \in \mathfrak{s}_W$, so $\mathrm{tr}(Y|_W) = 0$. On the other hand, since U is a trivial $\mathfrak{g}$–representation we see that $[X, Y] = 0$ for all $X \in \mathfrak{g}$, so Schur's Lemma from 2.4 implies that $Y|_W$ must be a multiple of the identity. But this together with $\mathrm{tr}(Y|_W) = 0$ implies $Y|_W = 0$, which completes the proof.
(2) Consider the Lie subalgebra $\mathrm{ad}(\mathfrak{g}) \subset \mathfrak{gl}(\mathfrak{g})$. For $X \in \mathfrak{g}$ let $\mathrm{ad}(X) = \mathrm{ad}(X)_S+\mathrm{ad}(X)_N$ be the Jordan decomposition of $\mathrm{ad}(X) : \mathfrak{gl}(V) \to \mathfrak{gl}(V)$. By part (2) of the Lemma, this is also the Jordan decomposition of $\mathrm{ad}(X) : \mathfrak{g} \to \mathfrak{g}$ and by part (1) of the Theorem we have $\mathrm{ad}(X)_S, \mathrm{ad}(X)_N \in \mathrm{ad}(\mathfrak{g})$. From 2.9 we know that $\mathrm{ad} : \mathfrak{g} \to \mathrm{ad}(\mathfrak{g})$ is an isomorphism, whence we obtain unique elements $X_S, X_N \in \mathfrak{g}$ such that $\mathrm{ad}(X_S) = \mathrm{ad}(X)_S$ and $\mathrm{ad}(X_N) = \mathrm{ad}(X)_N$.

Notice that in the special case of a Lie subalgebra $\mathfrak{g} \subset \mathfrak{gl}(V)$ the absolute Jordan decomposition $X = X_S + X_N$ coincides with the Jordan decomposition of the linear map $X : V \to V$. Indeed, if $X = S + N$ is the Jordan decomposition of $X : V \to V$, then by Lemma 2.7 $\mathrm{ad}(X) = \mathrm{ad}(S) + \mathrm{ad}(N)$ is the Jordan decomposition of $\mathrm{ad}(X) : \mathfrak{gl}(V) \to \mathfrak{gl}(V)$. Since $\mathrm{ad}(X)(\mathfrak{g}) \subset \mathfrak{g}$, part (2) of the Lemma implies that this is also the Jordan decomposition of $\mathrm{ad}(X) : \mathfrak{g} \to \mathfrak{g}$ and injectivity of ad implies $S = X_S$ and $N = X_N$.
(3) In view of the last remark it suffices to show that if $\mathfrak{g}$ is a semisimple Lie algebra and $\rho : \mathfrak{g} \to \mathfrak{g}'$ is a surjective homomorphism of Lie algebras (which implies that $\mathfrak{g}'$ is semisimple, too, see part (1) of corollary 2.8), then for $X \in \mathfrak{g}$ the absolute Jordan decomposition of $\rho(X)$ is $\rho(X_S) + \rho(X_N)$.

This is easy in the case that ρ is actually an isomorphism. In that case, $\rho([X,Y]) = [\rho(X), \rho(Y)]$ may be written as $\mathrm{ad}(\rho(X)) = \rho \circ \mathrm{ad}(X) \circ \rho^{-1}$ for each $X \in \mathfrak{g}$. In particular, since $\mathrm{ad}(X_S)$ is diagonalizable, the same is true for $\mathrm{ad}(\rho(X_S))$ and since $\mathrm{ad}(X_N)$ is nilpotent, also $\mathrm{ad}(\rho(X_N))$ is nilpotent. Moreover, since $\mathrm{ad}(X_S)$ and $\mathrm{ad}(X_N)$ commute, we see that $\mathrm{ad}(\rho(X_S))$ and $\mathrm{ad}(\rho(X_N))$ commute. Hence we conclude that $\mathrm{ad}(\rho(X)) = \mathrm{ad}(\rho(X_S)) + \mathrm{ad}(\rho(X_N))$ is the Jordan decomposition of $\mathrm{ad}(\rho(X))$, so $\rho(X)_S = \rho(X_S)$ and $\rho(X)_N = \rho(X_N)$ follow from (2).

On the other hand, we know from Corollary 2.8 that $\mathfrak{g}$ splits into the direct sum $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_k$ of simple ideals and the kernel of ρ (which must be an ideal in $\mathfrak{g}$) equals the sum of some of these simple ideals. Hence ρ may be written as the composition of an isomorphism $\mathfrak{g}_{i_1} \oplus \cdots \oplus \mathfrak{g}_{i_\ell} \to \mathfrak{g}'$ with the obvious projection $X = X^1 + \cdots + X^k \mapsto X^{i_1} + \cdots + X^{i_\ell}$. Thus to conclude the proof it suffices to show that the absolute Jordan decomposition of $X = X^1 + \cdots + X^k$ is given by $X = (X^1_S + \cdots + X^k_S) + (X^1_N + \cdots + X^k_N)$.

By construction $[X^i, X^j] = 0$ since this bracket has to be contained in $\mathfrak{g}_i \cap \mathfrak{g}_j$, so we conclude that $[X^i, X] = 0$, and thus $\mathrm{ad}(X^i)$ commutes with $\mathrm{ad}(X)$. Since $\mathrm{ad}(X^i_S)$ is a polynomial in $\mathrm{ad}(X^i)$, this also commutes with $\mathrm{ad}(X)$ and hence $\mathrm{ad}(X^1_S + \cdots + X^k_S)$ commutes with $\mathrm{ad}(X)$. Similarly, $\mathrm{ad}(X^1_N + \cdots + X^k_N)$ commutes with $\mathrm{ad}(X)$ and $\mathrm{ad}(X^1_S + \cdots + X^k_S)$. For each i choose a basis of $\mathfrak{g}_i$ in which $\mathrm{ad}(X^i_S)$ is diagonal and $\mathrm{ad}(X^i_N)$ is strictly upper triangular, and take the union of theses bases as a basis for $\mathfrak{g}$. Then visibly in this basis $\mathrm{ad}(X^1_S + \cdots + X^k_S)$ is diagonal and $\mathrm{ad}(X^1_N + \cdots + X^k_N)$ is strictly upper triangular, so the result follows. □

3.2. Cartan subalgebras.

Definition. Let $\mathfrak{g}$ be a complex semisimple Lie algebra. An element $X \in \mathfrak{g}$ is called *semisimple* if the linear map $\mathrm{ad}(X) : \mathfrak{g} \to \mathfrak{g}$ is diagonalizable. Otherwise put, $X \in \mathfrak{g}$ is semisimple if there is no nilpotent part in the absolute Jordan decomposition of X.

Any semisimple Lie algebra $\mathfrak{g}$ contains nonzero semisimple elements. In view of the absolute Jordan decomposition, it suffices to show that there exist elements $X \in \mathfrak{g}$ such that $X_S \neq 0$. If this would not be the case, then $\mathrm{ad}(X) : \mathfrak{g} \to \mathfrak{g}$ would be nilpotent for each $X \in \mathfrak{g}$ which by Engel's Theorem would imply that $\mathrm{ad}(\mathfrak{g}) \cong \mathfrak{g}$ is nilpotent, see Corollary 2.2.

Given a nonzero semisimple element $X \in \mathfrak{g}$, we can decompose $\mathfrak{g} := \oplus_\lambda \mathfrak{g}_\lambda(X)$ into eigenspaces for $\mathrm{ad}(X)$. By the Jacobi identity, $[\mathfrak{g}_\lambda(X), \mathfrak{g}_\mu(X)] \subset \mathfrak{g}_{\lambda+\mu}(X)$. Moreover, by part (3) of Theorem 3.1, for any finite dimensional complex representation $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ the map $\rho(X) : V \to V$ is diagonalizable. Let $V = \oplus V_\mu(X)$ be the resulting decomposition into eigenspaces for the linear map $\rho(X)$. Since

$$X \cdot Y \cdot v = [X, Y] \cdot v + Y \cdot X \cdot v$$

we immediately conclude that for $Y \in \mathfrak{g}_\lambda(X)$ and $v \in V_\mu(X)$ we have $Y \cdot v \in V_{\lambda+\mu}(X)$, so we immediately get information on the possible eigenvalues.

Of course, we may try to refine this decomposition, by passing to a family of elements $X \in \mathfrak{g}$ which is simultaneously diagonalizable. First we have to recall a result from linear algebra:

Lemma. *Let V be a finite dimensional vector space and $A \subset L(V,V)$ be a family of linear maps such that each $f \in A$ is diagonalizable. Then there is a basis of V with respect to which all the elements of A are represented by diagonal matrices if and only if $f \circ g = g \circ f$ for all $f, g \in A$.*

Proof. Since any two diagonal matrices commute, one direction is obvious. On the other hand, since any linear combination of diagonal matrices is diagonal, we may pass to a subset of A which is a basis for the linear subspace of $L(V,V)$ spanned by A to conclude that simultaneous diagonalizability of this finite set implies simultaneous diagonalizability of all elements of this subspace. Hence it suffices to show that finitely many commuting diagonalizable linear maps $f_1, \dots, f_n$ are simultaneously diagonalizable, and we do this by induction on n.

For $n = 1$, there is nothing to prove, so assume that $n > 1$ and we have found a basis of V with respect to which $f_1, \dots, f_{n-1}$ are represented by diagonal matrices. Consider the joint eigenspace $V_{(\lambda_1,\dots,\lambda_{n-1})} = \{v \in V : f_i(v) = \lambda_i v \quad \forall i = 1, \dots, n-1\}$. By assumption, V is the direct sum of these joint eigenspaces. Now $f_n \circ f_i = f_i \circ f_n$ for all $i = 1, \dots, n-1$ and thus for $v \in V_{(\lambda_1,\dots,\lambda_{n-1})}$ we get $f_i(f_n(v)) = f_n(f_i(v)) = \lambda_i f_n(V)$, and thus $f_n(v) \in V_{(\lambda_1,\dots,\lambda_{n-1})}$. By part (1) of Lemma 3.1 the restriction of f_n to $V_{(\lambda_1,\dots,\lambda_{n-1})}$ is diagonalizable, so we find a basis of that space consisting of elements which are eigenvectors for all the maps $f_1, \dots, f_n$. Taking the union of such bases as a basis of V, we obtain the result. □

Now for semisimple elements X_i of $\mathfrak{g}$, the maps $\operatorname{ad}(X_i)$ are simultaneously diagonalizable if and only if $0 = [\operatorname{ad}(X_i), \operatorname{ad}(X_j)] = \operatorname{ad}([X_i, X_j])$ for all i, j. In view of the injectivity of ad, we are led to looking at commutative subalgebras of $\mathfrak{g}$, which consist of semisimple elements. For X, Y in such a subalgebra $[X, Y] = 0$ implies $[\rho(X), \rho(Y)] = 0$ for any representation ρ, so we see that these maps are again simultaneously diagonalizable, so we obtain a decomposition of any representation space into eigenspaces.

Of course, we obtain the finest possible decomposition if we use a maximal commutative subalgebra which consists entirely of semisimple elements. Such a subalgebra is called a *Cartan subalgebra* of $\mathfrak{g}$. It should be noted here that there is a more general notion of Cartan subalgebras in arbitrary Lie algebras (defined as maximal nilpotent subalgebras which coincide with their own normalizer), but we will restrict to the semisimple case here.

To see that Cartan subalgebras in semisimple Lie algebra do exist, one considers the *centralizer* $\mathfrak{c}(H) := \{X \in \mathfrak{g} : [X, H] = 0\}$ of a semisimple element H. By construction, this is a subalgebra of $\mathfrak{g}$. The minimal dimension of the subspace $\mathfrak{c}(H)$ when H ranges through all semisimple elements of $\mathfrak{g}$ is called the *rank* of the semisimple Lie algebra $\mathfrak{g}$, and elements H for which this minimal dimension is attained are called *regular*.

In the statement of uniqueness of Cartan subalgebras below, the notion of an inner automorphism of $\mathfrak{g}$ is used. We have already seen in 2.9 that for any Lie algebra $\mathfrak{g}$ the automorphism group $\operatorname{Aut}(\mathfrak{g})$ is a closed subgroup of $GL(\mathfrak{g})$ and thus a matrix group in the sense of 1.4. We also noted that the Lie algebra of this subgroup is $\mathfrak{der}(\mathfrak{g})$ which is isomorphic to $\mathfrak{g}$ if $\mathfrak{g}$ is semisimple. One defines the group $\operatorname{Int}(\mathfrak{g})$ of inner automorphisms as the connected component of the identity of $\operatorname{Aut}(\mathfrak{g})$. For any Lie group G with Lie algebra $\mathfrak{g}$ the adjoint action Ad (see 1.5) defines a smooth homomorphism $G \to \mathrm{Aut}(\mathfrak{g})$, which has values in $\mathrm{Int}(\mathfrak{g})$ if G is connected. It turns out that this homomorphism is always surjective onto $\mathrm{Int}(\mathfrak{g})$. Hence given any G an inner automorphism of $\mathfrak{g}$ can always be realized in the form $\mathrm{Ad}(g)$ for some $g \in G$.

Finally, we have to recall some elementary facts on polynomial mappings for the proof: Let V be a finite dimensional complex vector space. A map $V \to \mathbb{C}$ is called a polynomial if it can be written as a sum of functions of the form $v \mapsto \varphi_1(v) \cdots \varphi_k(v)$, where each $\varphi_i : V \to \mathbb{C}$ is a linear functional. For $V = \mathbb{C}^n$ this is equivalent to the usual definition as maps of the form $(z_1, \dots, z_n) \mapsto \sum a_{i_1 \dots i_n} z_1^{i_1} \cdots z_n^{i_n}$. Now let $p : V \to \mathbb{C}$ be a nonzero polynomial and consider the subset $\{v \in V : p(v) \neq 0\}$. Of course, this is open and non–empty (since $p \neq 0$) but even more, it is dense, i.e. its closure is all of V. Indeed, taking a point $v \in V$ such that $p(v) = 0$ we can find an affine complex line through v such that the restriction of p to this line is not identically zero. But then this is a polynomial function on $\mathbb{C}$ and thus has only finitely many zeros. Hence we can find points in which p is nonzero arbitrarily close to v. Now it is easy to see that the intersection of two open dense subsets U_1, U_2 of $V \cong \mathbb{C}^n$ is open and dense. Indeed, if $z \in \mathbb{C}^n$ is arbitrary and $\epsilon > 0$, then since U_1 is dense we find a point $a \in U_1$ whose distance to z is less than $\epsilon/2$. Since U_1 is open, it contains some ball of radius δ around a and we may assume that $\delta \leq \epsilon/2$. Since U_2 is dense, this ball contains an element of U_2, which thus lies in $U_1 \cap U_2$ and has distance $< \epsilon$ to z. As an intersection of two open sets, $U_1 \cap U_2$ is of course open. This immediately generalizes to finitely many sets. (In fact, by Baire's theorem, even the intersection of countable many dense open subsets of the complete metric space $\mathbb{C}^n$ is non–empty.) In particular, we also see that if W is another finite dimensional complex vector space and $p : V \to W$ is a nonzero polynomial (which means that for each linear functional $\varphi : W \to \mathbb{C}$ the map $\varphi \circ p : V \to \mathbb{C}$ is a polynomial) the set $\{v \in V : p(v) \neq 0\}$ is open and dense in V.

Armed with these observations we can now prove:

Theorem. *Let $\mathfrak{g}$ be a complex semisimple Lie algebra.*
(1) If $H \in \mathfrak{g}$ is a regular semisimple element, then $\mathfrak{c}(H) \leq \mathfrak{g}$ is a Cartan subalgebra.
(2) Any two Cartan subalgebras in $\mathfrak{g}$ are conjugate by an inner automorphism of $\mathfrak{g}$.

Proof. (1) Assume that H is a regular semisimple element of $\mathfrak{g}$, and let $\mathfrak{g} = \oplus \mathfrak{g}_\lambda(H)$ be the decomposition of $\mathfrak{g}$ according to eigenvalues of $\mathrm{ad}(H)$. In particular, $\mathfrak{c}(H) = \mathfrak{g}_0(H)$. We first claim that $\mathfrak{c}(H)$ is a nilpotent Lie subalgebra of $\mathfrak{g}$. If $\mathfrak{c}(H)$ would not be nilpotent, then by Engel's theorem in 2.2, there is an element $Z \in \mathfrak{c}(H)$ such that the restriction of $\mathrm{ad}(Z)$ to $\mathfrak{c}(H)$ is not nilpotent, i.e. $(\mathrm{ad}(Z)|_{\mathfrak{c}(H)})^{\dim(\mathfrak{c}(H))} \neq 0$. Of course, mapping Z to $(\mathrm{ad}(Z)|_{\mathfrak{c}(H)})^{\dim(\mathfrak{c}(H))}$ is a polynomial mapping, so if there is one such Z then the subset of all those Z is open and dense in $\mathfrak{c}(H)$.

On the other hand, we have observed above, that for any $Z \in \mathfrak{g}_0(H)$, the adjoint action $\mathrm{ad}(Z)$ preserves the decomposition $\mathfrak{g} = \oplus \mathfrak{g}_\lambda(H)$, and we can look at those elements for which the restriction of $\mathrm{ad}(Z)$ to $\oplus_{\lambda \neq 0} \mathfrak{g}_\lambda(H)$ is invertible. Again this set is the complement of the zeros of a polynomial mapping, it is non–empty since it contains H, and thus it is open and dense. Since two dense open subsets of $\mathfrak{c}(H)$ must have a nontrivial intersection, we find an element $Z \in \mathfrak{c}(H)$ whose generalized eigenspace corresponding to the eigenvalue zero is strictly contained in $\mathfrak{c}(H)$. But this generalized eigenspace coincides with $\mathfrak{c}(Z_S)$, the semisimple part of Z, which contradicts regularity of H. Hence $\mathfrak{c}(H)$ is indeed nilpotent.

The second main ingredient is to observe the behavior of the Killing form with respect to the decomposition $\mathfrak{g} = \oplus \mathfrak{g}_\lambda(H)$. Indeed for $X \in \mathfrak{g}_\lambda(H)$ and $Y \in \mathfrak{g}_\mu(H)$ we see that $\mathrm{ad}(X)\circ\mathrm{ad}(Y)$ maps $\mathfrak{g}_\nu(H)$ to $\mathfrak{g}_{\lambda+\mu+\nu}(H)$, and thus is tracefree unless $\lambda+\mu=0$. In particular, $\mathfrak{c}(H)=\mathfrak{g}_0(H)$ is perpendicular to $\oplus_{\lambda\neq 0}\mathfrak{g}_\lambda(H)$, and thus the restriction of the Killing form B to $\mathfrak{c}(H)$ is non–degenerate.

Now $\mathfrak{c}(H)$ is nilpotent and thus solvable, so the image $\mathrm{ad}(\mathfrak{c}(H))\subset\mathfrak{gl}(\mathfrak{g})$ is solvable, too. By Lie's theorem from 2.2 there is a basis of $\mathfrak{g}$ with respect to which all $\mathrm{ad}(X)$ for $X\in\mathfrak{c}(H)$ are represented by upper triangular matrices. But this implies that for $X,Y,Z\in\mathfrak{c}(H)$ the map $\mathrm{ad}([X,Y])=[\mathrm{ad}(X),\mathrm{ad}(Y)]$ is represented by a matrix that is strictly upper triangular, and thus $\mathrm{ad}([X,Y])\circ\mathrm{ad}(Z)$ is tracefree. But this implies that $B([X,Y],Z)=0$ for all $Z\in\mathfrak{c}(H)$ and thus $[X,Y]=0$ for all $X,Y\in\mathfrak{c}(H)$ by non–degeneracy of the restriction of the Killing form. Hence $\mathfrak{c}(H)$ is an Abelian Lie subalgebra of $\mathfrak{g}$. Moreover, $\mathfrak{c}(H)$ is by definition a maximal Abelian Lie subalgebra, since any element that commutes with H already lies in $\mathfrak{c}(H)$. Hence we only have to verify that all elements of $\mathfrak{c}(H)$ are semisimple.

An arbitrary element $Y\in\mathfrak{c}(H)$ by definition commutes with H and since H is semisimple this implies that $\mathrm{ad}(Y)$ preserves the decomposition $\mathfrak{g}=\oplus\mathfrak{g}_\lambda(H)$ into $\mathrm{ad}(H)$–eigenspaces. But then any polynomial in $\mathrm{ad}(Y)$ preserves this decomposition and hence commutes with $\mathrm{ad}(H)$. In particular, the nilpotent part and semisimple part in the Jordan decomposition of $\mathrm{ad}(Y)$ have this property, which implies that $Y_N\in\mathfrak{c}(H)$. For any $X\in\mathfrak{c}(H)$, the maps $\mathrm{ad}(X)$ and $\mathrm{ad}(Y_N)$ commute, so since $\mathrm{ad}(Y_N)$ is nilpotent, also $\mathrm{ad}(X)\circ\mathrm{ad}(Y_N)$ is nilpotent and thus in particular tracefree. Hence $B(Y_N,X)=0$ for all $X\in\mathfrak{c}(H)$, which implies $Y_N=0$ by non–degeneracy of the restriction of the Killing form, and hence Y is semisimple.
(2) This is proved using some elementary algebraic geometry, see [**Fulton-Harris**, Appendix D] or [**Knapp**, section II.3] □

Choosing a Cartan subalgebra $\mathfrak{h}\subset\mathfrak{g}$ we have noted already above that for any finite dimensional complex representation $\rho:\mathfrak{g}\to\mathfrak{gl}(V)$ the operators $\rho(H):V\to V$ are simultaneously diagonalizable. Moreover, on a simultaneous eigenspace, the eigenvalue depends linearly on H, so it is described by a linear functional $\lambda:\mathfrak{h}\to\mathbb{C}$. The functionals corresponding to nontrivial eigenspaces are called the *weights* of the representation V, i.e. $\lambda:\mathfrak{h}\to\mathbb{C}$ is a weight of V if and only if there is a nonzero vector $v\in V$ such that $\rho(H)(v)=\lambda(H)v$ for all $H\in\mathfrak{h}$. If λ is a weight then the *weight space* V_λ corresponding to λ is defined by $V_\lambda:=\{v\in V:\rho(H)(v)=\lambda(H)v\quad\forall H\in\mathfrak{h}\}$. The set of weights of V is a finite subset of $\mathfrak{h}^*$, that will be denoted by $\mathrm{wt}(V)$.

In particular, we may consider the adjoint representation. The nonzero weights of the adjoint representation are called the *roots* of the Lie algebra $\mathfrak{g}$ with respect to $\mathfrak{h}$, and the weight space $\mathfrak{g}_\alpha$ corresponding to a root α is called a *root space*. We will denote the set of all roots of $\mathfrak{g}$ by Δ. The weight space corresponding to the weight zero is exactly the Cartan subalgebra $\mathfrak{h}$, and we obtain the *root decomposition* $\mathfrak{g}=\mathfrak{h}\oplus\bigoplus_{\alpha\in\Delta}\mathfrak{g}_\alpha$. For $X\in\mathfrak{g}_\alpha$, $Y\in\mathfrak{g}_\beta$ and $H\in\mathfrak{h}$ we have

$$[H,[X,Y]]=[[H,X],Y]+[X,[H,Y]]=(\alpha(H)+\beta(H))[X,Y].$$

Thus $[\mathfrak{g}_\alpha,\mathfrak{g}_\beta]\subset\mathfrak{g}_{\alpha+\beta}$ for $\alpha,\beta\in\Delta$, where we agree that $\mathfrak{g}_\gamma=\{0\}$ if γ is not a root.

More generally, the definition of an action implies that if $V=\oplus_{\lambda\in\mathrm{wt}(V)}V_\lambda$ is the weight decomposition of a representation then for $v\in V_\lambda$, $X\in\mathfrak{g}_\alpha$, and $H\in\mathfrak{h}$ we have

$$H\cdot X\cdot v=[H,X]\cdot v+X\cdot H\cdot v=(\alpha(H)+\lambda(H))X\cdot v,$$

and thus $X\cdot v\in V_{\lambda+\alpha}$. Hence we immediately get information on the possible eigenvalues.

The root system of a complex semisimple Lie algebra

3.3. Example: $\mathfrak{sl}(n,\mathbb{C})$. We know from example (1) of 2.3 that $\mathfrak{g} := \mathfrak{sl}(n,\mathbb{C})$ is a simple Lie algebra for $n \geq 2$, and we have essentially already constructed the root decomposition there. Let us redo this in the language introduced above. Let $\mathfrak{h} \subset \mathfrak{g}$ be the subspace of all tracefree diagonal matrices. Of course, this is a commutative subalgebra of $\mathfrak{g}$. Moreover, for $i \neq j$, consider the elementary matrix E_{ij} which by definition has all entries equal to zero, except from a 1 in the jth column of the ith row. Let $H \in \mathfrak{h}$ be a matrix with diagonal entries $x_{11}, \dots, x_{nn}$ (so we know that $x_{11} + \cdots + x_{nn} = \mathrm{tr}(H) = 0$). Clearly, $HE_{ij} = x_{ii}E_{ij}$ and $E_{ij}H = x_{jj}E_{ij}$, and thus $[H, E_{ij}] = (x_{ii} - x_{jj})E_{ij}$. In particular, we can find some $H \in \mathfrak{h}$ such that $[H, E_{ij}] \neq 0$, for example $H = E_{ii} - E_{jj}$. Moreover, denoting by $e_i : \mathfrak{h} \to \mathbb{C}$ the linear functional which extracts the i–th diagonal entry of a diagonal matrix, we see that E_{ij} is a joint eigenvector for the maps $\mathrm{ad}(H)$ with $H \in \mathfrak{h}$ with eigenvalue $e_i - e_j \in \mathfrak{h}^*$.

This immediately shows that $\mathfrak{h}$ is a maximal commutative subalgebra of $\mathfrak{g}$. Indeed, any element $X \in \mathfrak{g}$ can be written as the sum of some element $H' \in \mathfrak{h}$ and the matrices E_{ij} with $i \neq j$. Vanishing of $[H, X]$ for all $H \in \mathfrak{h}$ then immediately implies that all the coefficients of the E_{ij} have to vanish. On the other hand, for any element $H \in \mathfrak{h}$ the map $\mathrm{ad}(\mathfrak{h}) : \mathfrak{g} \to \mathfrak{g}$ is diagonalizable, so $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$. Moreover, if $H_0 \in \mathfrak{h}$ has the property that all its diagonal elements are different, then even vanishing of $[H_0, X]$ implies $X \in \mathfrak{h}$, and thus $\mathfrak{h} = \mathfrak{c}(H_0)$. In particular, the rank of $\mathfrak{sl}(n,\mathbb{C})$ is $n - 1$.

It also follows immediately that the set of roots in this case is given by $\Delta = \{e_i - e_j : i \neq j\}$. Each of the root spaces $\mathfrak{g}_{e_i - e_j}$ is one dimensional and spanned by E_{ij}. Moreover, for any root $\alpha \in \Delta$ we also have $-\alpha \in \Delta$, and no other complex multiple of α is a root. Suppose that $H \in \mathfrak{h}$ has the property that $\alpha(H) = 0$ for all $\alpha \in \Delta$. Then of course, all the diagonal entries of H must be equal, but then $\mathrm{tr}(H) = 0$ implies $H = 0$. This exactly means that the roots span the dual space $\mathfrak{h}^*$. We also have a distinguished real subspace $\mathfrak{h}_0 \subset \mathfrak{h}$ consisting of all real tracefree diagonal matrices and visibly $\mathfrak{h}$ is the complexification of $\mathfrak{h}_0$. Suppose that $H \in \mathfrak{h}$ has the property that $\alpha(H) \in \mathbb{R}$ for all $\alpha \in \Delta$. Then any two diagonal entries of H differ by a real number, so H must be the sum of a purely imaginary multiple of the identity and a real diagonal matrix. Vanishing of the imaginary part of the trace of H then implies that all diagonal entries of H are real, so we can characterize $\mathfrak{h}_0$ as the subspace of $\mathfrak{h}$ on which all roots take real values.

Our next task is to show that all these properties are true for arbitrary complex semisimple Lie algebras. The basis for this will be to study the representation theory of $\mathfrak{sl}(2,\mathbb{C})$.

3.4. Digression on $\mathfrak{sl}(2,\mathbb{C})$. In the special case $n = 2$ the root decomposition from 3.3 above becomes very simple. The Cartan subalgebra $\mathfrak{h}$ has dimension one, and the obvious choice for a basis is $H := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in \mathfrak{h}_0$. For the two root spaces we take $E := E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, and $F := E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ as basis elements. This precisely gives the basis from example 2.6 and we have already seen that the commutation relations between these basis elements are $[E, F] = H$, $[H, E] = 2E$ and $[H, F] = 2F$. Since we have chosen the basis $\{H\}$ for $\mathfrak{h}$ we may view linear functionals on $\mathfrak{h}$ simply as complex numbers. In this picture $\Delta = \{2, -2\}$ and the root decomposition is $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{g}_{-2} \oplus \mathfrak{g}_2$, where $\mathfrak{g}_2 = \mathbb{C} \cdot E$ and $\mathfrak{g}_{-2} = \mathbb{C} \cdot F$.

Next, let $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ be a finite dimensional irreducible representation of $\mathfrak{g}$ with weight decomposition $V = \oplus V_\lambda$, where $\lambda \in \mathbb{C}$ and $V_\lambda = \{v \in V : \rho(H)(v) = \lambda v\}$. Let λ_0 be the weight of V which has maximal real part, and let $v_0 \in V_{\lambda_0}$ be a nonzero element. Then $\rho(E)(v_0) = 0$, since if this were nonzero it would be a weight vector of weight λ_0+2. Then consider the subspace of V which is spanned by $\{v_n := \rho(F)^n(v_0) : n \in \mathbb{N}\}$. Of course, if v_n is nonzero, then it is an eigenvector for $\rho(H)$ with eigenvalue $\lambda_0 - 2n$, so the nonzero elements in this set are linearly independent. Hence there must be a minimal index n such that $v_{n+1} = 0$. By construction, $\rho(F)(v_j) = v_{j+1}$, so our subspace is invariant under $\rho(F)$. On the other hand, we can compute the action of $\rho(E)$ on these elements:

$$\rho(E)(v_1) = \rho(E)(\rho(F)(v_0)) = [\rho(E), \rho(F)](v_0) + \rho(F)(\rho(E)(v_0)) = \lambda_0 v_0 + 0.$$

Similarly,

$$\rho(E)(v_2) = \rho(E)(\rho(F)(v_1)) = [\rho(E), \rho(F)](v_1) + \rho(F)(\rho(E)(v_1)) = (\lambda_0 - 2 + \lambda_0)v_1.$$

Inductively, this immediately implies that $\rho(E)(v_j) = a_j v_{j-1}$, where $a_j = a_{j-1} + \lambda_0 - 2j$ and $a_1 = \lambda_0$, and we obtain $a_j = j\lambda_0 - j(j-1) = j(\lambda_0 - j + 1)$. Thus we see that our subspace is invariant, so it has to coincide with V by irreducibility, which implies that $\{v_0, \dots, v_n\}$ is a basis for V. Now it remains to find the relation between λ_0 and n, which however also is rather easy. On one hand, $\rho(H) = [\rho(E), \rho(F)]$, so we know that $\rho(H)$ must be tracefree. On the other hand, since $\rho(H)$ is diagonal with respect to the basis above, we see that the trace of $\rho(H)$ is given by $\sum_{j=0}^n (\lambda_0 - 2j) = (n+1)(\lambda_0 - n)$, which shows that $\lambda_0 = n$, so in particular, all weights showing up in finite dimensional representations are integers.

In particular, these observations imply that an irreducible representation of $\mathfrak{sl}(2, \mathbb{C})$ is determined up to isomorphism by its dimension. Indeed, if we have two irreducible representations V and V' of dimension $n+1$, then we consider the bases $\{v_0, \dots, v_n\}$ and $\{v'_0, \dots, v'_n\}$ as constructed above, and then the mapping defined by $v_j \mapsto v'_j$ is an isomorphism of representations, since the actions of the three generators on each of the basis elements are determined by the dimension.

To show that for each $n \in \mathbb{N}$ there exists an irreducible representation of $\mathfrak{sl}(2, \mathbb{C})$ one may consider the standard basis $\{e_0, \dots, e_n\}$ of $\mathbb{C}^{n+1}$ and define $\rho : \mathfrak{sl}(2, \mathbb{C}) \to \mathfrak{gl}(n+1, \mathbb{C})$ by $\rho(F)(e_j) = e_{j+1}$, $\rho(F)(e_n) = 0$, $\rho(H)(e_j) = (n-2j)e_j$, $\rho(E)(e_0) = 0$ and $\rho(E)(e_j) = j(n-j+1)e_{j-1}$, and verify directly that this indeed defines a representation. A more conceptual way to obtain these representations will be discussed in 4.7. Thus we obtain:

PROPOSITION. *For any $n \in \mathbb{N}$ there is a unique (up to isomorphism) irreducible representation of $\mathfrak{sl}(2, \mathbb{C})$ of dimension $n+1$. The weights of this representation are $\{n, n-2, \dots, -n+2, -n\}$ and all weight spaces are one–dimensional. More precisely, there is a basis $\{v_0, \dots, v_n\}$ of V such that $H \cdot v_j = (n-2j)v_j$, $F \cdot v_j = v_{j+1}$ for $j < n$ and $F \cdot v_n = 0$ and $E \cdot v_j = j(n-j+1)v_{j-1}$.*

3.5. The root system of a complex semisimple Lie algebra. Let us return to a general complex semisimple Lie algebra $\mathfrak{g}$ with a Cartan subalgebra $\mathfrak{h}$ and the corresponding root decomposition $\mathfrak{g} = \mathfrak{h} \oplus_{\alpha \in \Delta} \mathfrak{g}_\alpha$. We already know that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$. In particular, this implies that for $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_\beta$ the map $\operatorname{ad}(X) \circ \operatorname{ad}(Y)$ maps $\mathfrak{g}_\gamma$ to $\mathfrak{g}_{\alpha+\beta+\gamma}$, so this is tracefree unless $\beta = -\alpha$. Hence $B(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ unless $\beta = -\alpha$. Non–degeneracy of B then implies that for $\alpha \in \Delta$ also $-\alpha \in \Delta$ and B induces a non–degenerate pairing between $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$. Moreover, the restriction of the Killing form to $\mathfrak{h}$ is non–degenerate.

Secondly, assume that $H \in \mathfrak{h}$ is such that $\alpha(H) = 0$ for all $\alpha \in \Delta$. Then H commutes with any element of any root space, and since $H \in \mathfrak{h}$ it also commutes with any element of $\mathfrak{h}$, so $H \in \mathfrak{z}(\mathfrak{g})$. Since $\mathfrak{g}$ is semisimple, this implies $H = 0$, which in turn implies that the roots generate the dual space $\mathfrak{h}^*$.

Now consider an element $X \in \mathfrak{g}_\alpha$. From above we know that there is a $Y \in \mathfrak{g}_{-\alpha}$ such that $B(X, Y) \neq 0$. On the other hand, we know that there is an element $H \in \mathfrak{h}$ such that $\alpha(H) \neq 0$. But then $B(H, [X, Y]) = B([H, X], Y) = \alpha(H)B(X, Y) \neq 0$, so $[X, Y] \neq 0$. Now since the restriction of B to $\mathfrak{h}$ is non–degenerate, there is a unique element $H_\alpha \in \mathfrak{h}$ such that $B(H, H_\alpha) = \alpha(H)$ for all $H \in \mathfrak{h}$. Since $[X, Y] \in \mathfrak{h}$, the equation $B(H, [X, Y]) = \alpha(H)B(X, Y)$ for all $H \in \mathfrak{h}$ then implies that $[X, Y] = B(X, Y)H_\alpha$.

Next, we want to show that $\alpha(H_\alpha) = B(H_\alpha, H_\alpha) \neq 0$. To do this, we introduce an idea which will be used frequently in the sequel. Let $\beta \in \Delta$ be a root and consider the subspace $\mathfrak{k} := \oplus_{n\in\mathbb{Z}}\mathfrak{g}_{\beta+n\alpha}$, which is called the *$\alpha$–string through β*. As above, we may choose $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$ such that $[X, Y] = H_\alpha$. Then $\mathfrak{k}$ is invariant under $\mathrm{ad}(X)$ and $\mathrm{ad}(Y)$ and $\mathrm{ad}(H_\alpha) = [\mathrm{ad}(X), \mathrm{ad}(Y)]$ on $\mathfrak{k}$. In particular, $\mathrm{ad}(H_\alpha)$ is tracefree on $\mathfrak{k}$, which implies that $0 = \sum_{n\in\mathbb{Z}}(\beta(H_\alpha) + n\alpha(H_\alpha)) \dim(g_{\beta+n\alpha})$. In particular, $\alpha(H_\alpha) = 0$ implies $\beta(H_\alpha) = 0$ for all $\beta \in \Delta$ which leads to the contradiction $H_\alpha = 0$. Moreover, this equation shows that for any $\beta \in \Delta$ the number $\beta(H_\alpha)$ is a rational multiple of $\alpha(H_\alpha)$.

This is the place where the representation theory of $\mathfrak{sl}(2, \mathbb{C})$ comes into the game. Namely, since $B(H_\alpha, H_\alpha) = \alpha(H_\alpha) \neq 0$, we may choose $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$ such that $\alpha([X, Y]) = 2$. But this implies that the elements X, Y and $H := [X, Y]$ we have $[H, X] = \alpha(H)X = 2X$ and $[H, Y] = -\alpha(H)Y = -2Y$. Hence these three elements span a Lie subalgebra $\mathfrak{s}_\alpha \leq \mathfrak{g}$, which is obviously isomorphic to $\mathfrak{sl}(2, \mathbb{C})$. Now consider the (finite–dimensional) subspace $\mathfrak{h} \oplus \bigoplus_{z\in\mathbb{C}\setminus\{0\}} \mathfrak{g}_{z\alpha}$, which obviously is invariant under the action of $\mathfrak{s}_\alpha$. From complete reducibility and Proposition 3.4 we conclude that $\mathfrak{g}_{z\alpha}$ is trivial, unless z is half integer. Moreover, $\mathfrak{s}_\alpha$ acts trivially on the codimension one subspace $\ker(\alpha) \subset \mathfrak{h}$, and irreducible on $\mathfrak{s}_\alpha$, which adds a one–dimensional weight space of weight zero spanned by H. Since we have used up all the zero weight space, we see that apart from these two irreducible components no other irreducible components of even highest weight can show up. In particular, 2α must not be a root. But this in turn implies that $\frac{1}{2}\alpha$ cannot be a root, which shows that the weight space of weight one is trivial. But this implies that there are no other irreducible components at all, i.e. our space just has the form $\ker(\alpha) \oplus \mathfrak{s}_\alpha$, which in particular implies that $\mathfrak{g}_\alpha$ is one–dimensional.

Let us formalize things a little more. Since the restriction of B to $\mathfrak{h}$ is non–degenerate, for any linear functional $\varphi \in \mathfrak{h}^*$ there is a unique element $H_\varphi \in \mathfrak{h}$ such that $B(H, H_\varphi) = \varphi(H)$ for all $H \in \mathfrak{h}$. Using this, we can define a non–degenerate complex bilinear form $\langle\ ,\ \rangle$ on $\mathfrak{h}^*$ by putting $\langle\varphi, \psi\rangle := B(H_\varphi, H_\psi) = \varphi(H_\psi) = \psi(H_\varphi)$. By construction, the distinguished element $H \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ from above is characterized by $\alpha(H) = 2$, which implies that $H = \frac{2}{\langle\alpha,\alpha\rangle}H_\alpha$.

The next step is then to consider two roots $\alpha, \beta \in \Delta$ which are not proportional to each other, and consider the α–string through β, i.e. the subspace $\oplus_{n\in\mathbb{Z}}\mathfrak{g}_{\beta+n\alpha}$. By construction, this subspace of $\mathfrak{g}$ is invariant under the adjoint action of $\mathfrak{s}_\alpha$, so we may again apply the representation theory of $\mathfrak{sl}(2, \mathbb{C})$. By construction, the eigenvalue of the action of the distinguished element $\frac{2}{\langle\alpha,\alpha\rangle}H_\alpha$ on $\mathfrak{g}_{\beta+n\alpha}$ is given by $\frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle} + 2n$, so these eigenvalues are all different. From above, we know that each of the spaces $\mathfrak{g}_{\beta+n\alpha}$ is of dimension at most one. On the other hand, by complete reducibility we know the we get a decomposition into a direct sum of irreducible representations. Using the description of the irreducible representations of $\mathfrak{sl}(2,\mathbb{C})$ in Proposition 3.4 above and our information on the eigenvalues, we conclude that any irreducible component must have the form of an unbroken string $\oplus_{n=n_0}^{n_1}\mathfrak{g}_{\beta+n\alpha}$ of 1–dimensional spaces. On the other hand, we know that the eigenvalues of $\frac{2}{\langle\alpha,\alpha\rangle}H_\alpha$ must be symmetric around zero, which implies that there is just one irreducible component. Thus, the α–string through β must have the form $\oplus_{n=-p}^{q}\mathfrak{g}_{\beta+n\alpha}$ with $p,q\geq 0$ and the maximal eigenvalue $\frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle}+2q$ must equal $p+q$, which implies $p-q=\frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle}$. In particular $\frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle}\in\mathbb{Z}$ for all $\alpha,\beta\in\Delta$. Moreover, we see that for $\alpha,\beta\in\Delta$ such that $\alpha+\beta\in\Delta$, we have $[\mathfrak{g}_\alpha,\mathfrak{g}_\beta]=\mathfrak{g}_{\alpha+\beta}$.

This completes our information on the root decomposition of a complex semisimple Lie algebra, which we collect in the following

Proposition. *Let $\mathfrak{g}$ be a complex semisimple Lie algebra, $\mathfrak{h}\leq\mathfrak{g}$ a Cartan subalgebra $\Delta\subset\mathfrak{h}^*$ the corresponding set of roots and $\langle\ ,\ \rangle$ the complex bilinear form on $\mathfrak{h}^*$ induced by the Killing form. Then we have:*
(1) For any $\alpha\in\Delta$ also $-\alpha\in\Delta$ and these are the only nonzero complex multiples of α which are roots.
(2) For any $\alpha\in\Delta$ the root space $\mathfrak{g}_\alpha$ is one–dimensional, and the subspace $\mathfrak{s}_\alpha$ of $\mathfrak{g}$ spanned by $\mathfrak{g}_\alpha$, $\mathfrak{g}_{-\alpha}$, and $[\mathfrak{g}_\alpha,\mathfrak{g}_{-\alpha}]\subset\mathfrak{h}$ is a Lie subalgebra isomorphic to $\mathfrak{sl}(2,\mathbb{C})$.
(3) For $\alpha,\beta\in\Delta$ with $\beta\neq-\alpha$ we have $[\mathfrak{g}_\alpha,\mathfrak{g}_\beta]=\mathfrak{g}_{\alpha+\beta}$ if $\alpha+\beta\in\Delta$ and $[\mathfrak{g}_\alpha,\mathfrak{g}_\beta]=\{0\}$ otherwise.
(4) For $\alpha,\beta\in\Delta$ with $\beta\neq\pm\alpha$ and $z\in\mathbb{C}$ a functional of the form $\beta+z\alpha$ can only be a root if $z\in\mathbb{Z}$. The roots of this form are an unbroken string $\beta-p\alpha,\beta-(p-1)\alpha,\dots,\beta+q\alpha$, where $p,q\geq 0$ and $p-q=\frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle}$.

To proceed, we need a bit more information on the system Δ of roots. Let $V\subset\mathfrak{h}^*$ be the real span of the set Δ and let $\mathfrak{h}_0\subset\mathfrak{h}$ be the real subspace spanned by the elements H_α for $\alpha\in\Delta$. We claim that $\mathfrak{h}$ is the complexification of $\mathfrak{h}_0$, the real dual of $\mathfrak{h}_0$ is exactly V, and the Killing form B (respectively the inner product $\langle\ ,\ \rangle$) is positive definite on $\mathfrak{h}_0$ (respectively on V).

By definition, $\operatorname{ad}(H_\alpha)$ acts trivially on $\mathfrak{h}$ and by multiplication by $\beta(H_\alpha)=\langle\alpha,\beta\rangle$ on $\mathfrak{g}_\beta$. Therefore, the trace of $\operatorname{ad}(H_\alpha)\circ\operatorname{ad}(H_\alpha)$ is given by $\sum_{\beta\in\Delta}\beta(H_\alpha)^2$. On the other hand, this trace by definition equals $B(H_\alpha,H_\alpha)=\langle\alpha,\alpha\rangle$. Thus, we get

$$(*)\qquad \langle\alpha,\alpha\rangle=2\langle\alpha,\alpha\rangle^2+\sum_{\beta\in\Delta;\beta\neq\pm\alpha}\langle\beta,\alpha\rangle^2.$$

Dividing this equation by $\langle\alpha,\alpha\rangle^2$, we see from above that each summand in the right hand side becomes a rational number which implies that $\langle\alpha,\alpha\rangle=\alpha(H_\alpha)$ is a rational number. From above we know that $\beta(H_\alpha)$ is a rational multiple of $\alpha(H_\alpha)$, so this is a rational number, too. Choosing a complex basis of $\mathfrak{h}^*$ consisting of roots and considering the real span of this basis, we see that the subspace of $\mathfrak{h}$ on which all these forms are real contains all H_α. This implies that the real dimension of the real span of the H_α equals at most the complex dimension of $\mathfrak{h}$. On the other hand, the H_β for the elements β in the above basis are linearly independent, so the dimension of the real span of all H_α equals the complex dimension of $\mathfrak{h}$. From this we conclude that the real subspace of $\mathfrak{h}^*$ constructed above must coincide with V. Finally, since we know that $\langle\beta,\alpha\rangle$ is rational, equation $(*)$ shows that $\langle\alpha,\alpha\rangle>0$ for all α.

This brings us the final idea in this context, which is to consider the reflections of the real Euclidean vector space $\mathfrak{h}_0^*$ in the hyperplanes orthogonal to roots. For $\alpha\in\Delta$ the *root reflection* $s_\alpha:\mathfrak{h}_0^*\to\mathfrak{h}_0^*$ on the hyperplane $\alpha^\perp$ is given by $s_\alpha(\varphi)=\varphi-\frac{2\langle\varphi,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha$.

As a reflection on a hyperplane this is an orthogonal mapping of determinant -1, and $s_\alpha(\alpha) = -\alpha \in \Delta$. For a root $\beta \neq \pm\alpha$ we have $s_\alpha(\beta) = \beta - (p-q)\alpha$, where the α–string through β has the form $\mathfrak{g}_{\beta-p\alpha} \oplus \dots \oplus \mathfrak{g}_{\beta+q\alpha}$. Since $-p \leq q-p \leq q$, we conclude that $s_\alpha(\beta) \in \Delta$.

Collecting our information about the set of roots we get:

Theorem. *Let $\mathfrak{g}$ be a semisimple Lie algebra of rank r, let $\mathfrak{h} \leq \mathfrak{g}$ be a Cartan subalgebra, $\Delta \subset \mathfrak{h}^*$ the corresponding set of roots, $\mathfrak{h}_0^* \subset \mathfrak{h}^*$ the real span of Δ and let $\langle\ ,\ \rangle$ be the complex bilinear form on $\mathfrak{h}^*$ induced by the Killing form. Then*
(1) $\mathfrak{h}_0^*$ has real dimension r and the restriction of $\langle\ ,\ \rangle$ to $\mathfrak{h}_0^*$ is positive definite.
(2) The subset $\Delta \subset \mathfrak{h}_0^*$ is an abstract root system in the Euclidean vector space $(\mathfrak{h}_0^*, \langle\ ,\ \rangle)$, i.e. a finite set of nonzero elements which span the whole space, such that any reflection s_α for $\alpha \in \Delta$ maps Δ to itself and such that for $\alpha, \beta \in \Delta$ we have $\frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle} \in \mathbb{Z}$.
(3) The abstract root system $\Delta \subset \mathfrak{h}_0^*$ is reduced root system!reduced, i.e. $\alpha \in \Delta$ implies $2\alpha \notin \Delta$.

There is a final point to make here, which concerns simplicity of $\mathfrak{g}$. Note that there is an obvious notion of direct sum of abstract root systems, by taking the orthogonal direct sum of the surrounding Euclidean spaces and the union of the abstract root systems. An abstract root system is called *irreducible* if and only if it does not decompose as a direct sum in a nontrivial way.

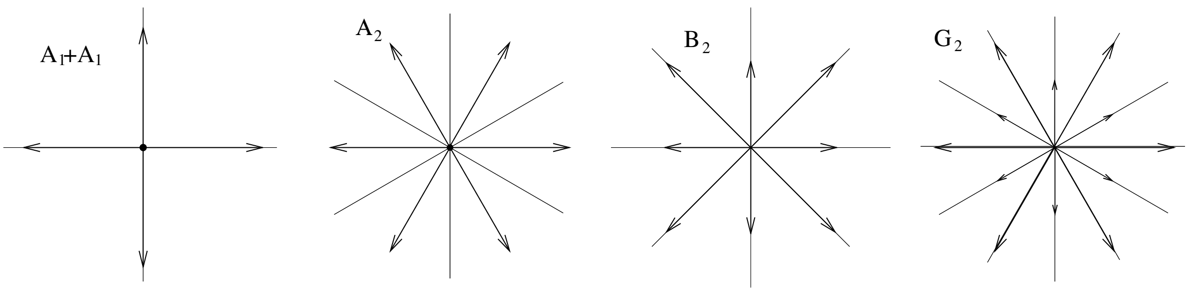
Assume that $\mathfrak{g}$ is a complex semisimple Lie algebra with a Cartan subalgebra $\mathfrak{h}$ such that the associated root system decomposes into a direct sum. Then $\mathfrak{h}_0^*$ is an orthogonal direct sum and Δ is the disjoint union $\Delta = \Delta_1 \sqcup \Delta_2$. Define $\mathfrak{h}_1 \subset \mathfrak{h}$ to be the subspace on which all $\alpha \in \Delta_2$ vanish and likewise for $\mathfrak{h}_2$. Taking a basis for $\mathfrak{h}^*$ consisting of elements of Δ we immediately see that $\mathfrak{h} = \mathfrak{h}_1 \oplus \mathfrak{h}_2$. Now define $\mathfrak{g}_i := \mathfrak{h}_i \oplus \oplus_{\alpha\in\Delta_i}\mathfrak{g}_\alpha$ for $i = 1, 2$. Clearly we get $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ as a vector space, and by construction each $\mathfrak{g}_i$ is a Lie subalgebra of $\mathfrak{g}$. Taking $H \in \mathfrak{h}_1$, we see that $[H, \mathfrak{h}_2] = 0$ and since any $\beta \in \Delta_2$ vanishes on H we conclude $[H, \mathfrak{g}_2] = 0$. Similarly for $\alpha \in \Delta_1$ and $X \in \mathfrak{g}_\alpha$ we have $[\mathfrak{h}_2, X] = 0$. For $\beta \in \Delta_2$, if $\alpha + \beta \in \Delta$ then this would have to be either in Δ_1 or in Δ_2 so it would have to vanish identically either on $\mathfrak{h}_1$ or on $\mathfrak{h}_2$ which would imply $\alpha = 0$ respectively $\beta = 0$. Hence we conclude $\alpha + \beta \notin \Delta$ which together with the above implies that $[\mathfrak{g}_1, \mathfrak{g}_2] = 0$, so each $\mathfrak{g}_i$ is an ideal in $\mathfrak{g}$, so $\mathfrak{g}$ is not simple.

Conversely, if $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ is a direct sum of ideals, we can choose Cartan subalgebras $\mathfrak{h}_i \subset \mathfrak{g}_i$. Then $\mathfrak{h} := \mathfrak{h}_1 \oplus \mathfrak{h}_2$ is a commutative subalgebra of $\mathfrak{g}$ which consists of semisimple elements, and one immediately verifies that $\mathfrak{h}$ is a maximal commutative subalgebra and thus a Cartan subalgebra. Moreover, we know from 2.8 that the decomposition $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ is orthogonal with respect to the Killing form. Denoting by Δ_i the set of roots for $(\mathfrak{g}_i, \mathfrak{h}_i)$ for $i = 1, 2$ we immediately conclude that the root system Δ of $(\mathfrak{g}, \mathfrak{h})$ is the direct sum of Δ_1 and Δ_2. Hence we conclude that a complex semisimple Lie algebra $\mathfrak{g}$ is actually simple if and only if its root system is irreducible.

3.6. Example: Two dimensional abstract root systems. The further steps towards and efficient description and the classification of complex semisimple Lie algebras can be entirely done in the general setting of abstract root systems. Notice that there is a simple notion of isomorphism of abstract root system, namely $(V, \Delta) \cong (V', \Delta')$ if and only if there is a linear map $\varphi : V \to V'$ such that $\varphi(\Delta) = \Delta'$ and such that $\frac{2\langle\varphi(\beta),\varphi(\alpha)\rangle}{\langle\varphi(\alpha),\varphi(\alpha)\rangle} = \frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle}$ for all $\alpha, \beta \in \Delta$, where the first inner product is in V' and the second is in V. As an example, we consider here reduced abstract root systems in $\mathbb{R}$ and $\mathbb{R}^2$.

For the case of $\Delta \subset \mathbb{R}$, we must have a nonzero element $\alpha \in \Delta$, since Δ spans $\mathbb{R}$. Of course, $s_\alpha(\alpha) = -\alpha$, so this must lie in Δ, too. Any other element $\beta \in \Delta$ then can be written as $r\alpha$ for some $r \in \mathbb{R} \setminus \{0\}$. By definition, $\frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle} = 2r \in \mathbb{Z}$ so r must be half integer. Exchanging the roles of α and β we see that also $\frac{2r}{r^2} = \frac{2}{r}$ must be an integer, which only leaves the possibilities $r = \pm 1$ and $r = \pm 2$. Since we want to get a reduced root system, the case $r = \pm 2$ is impossible, since neither 2α nor $2(-\alpha)$ may be contained in Δ. Hence we have $\Delta = \{\alpha, -\alpha\}$. Since rescaling of α defines an isomorphism, we see that up to isomorphism the root system $\Delta = \{-2, 2\}$ of $\mathfrak{sl}(2,\mathbb{C})$ is the only one–dimensional abstract root system. This root system is commonly called A_1 and it can be pictured as .

We already know one abstract root system in $\mathbb{R}^2$, namely the root system $\Delta = \{e_1 - e_2, e_1 - e_3, e_2 - e_3, e_2 - e_1, e_3 - e_1, e_3 - e_2\}$ of $\mathfrak{sl}(3,\mathbb{C})$, which is commonly called A_2. Putting $\alpha_1 = e_1 - e_2$ and $\alpha_2 = e_2 - e_3$ we clearly get two linearly independent elements of Δ, and we immediately see that $\Delta = \{\alpha_1, \alpha_2, \alpha_1 + \alpha_2, -\alpha_1, -\alpha_2, -\alpha_1 - \alpha_2\}$. To picture this system, we have to determine the inner product on $\mathfrak{h}_0^*$. Since a rescaling of all elements clearly leads to an isomorphism of abstract root systems, it suffices to do this up to some factor. Essentially, we therefore have to determine the angle between α_1 and α_2. We will do this for $\mathfrak{sl}(n,\mathbb{C})$ by observing that for two diagonal matrices H and H' with diagonal entries x_i and y_i, we get $\operatorname{tr}(HH') = \sum x_i y_i$. From 2.6 we know that this trace form is invariant on $\mathfrak{sl}(n,\mathbb{C})$ and hence part (3) of corollary 2.8 implies that it is a nonzero multiple of the Killing form. Thus we conclude that the element of $\mathfrak{h}$ which represents the functional e_i has the ith diagonal entry equal to $\frac{n-1}{n}$ and all other diagonal entries equal to $\frac{-1}{n}$, and we can write this as the ith unit vector minus $\frac{1}{n}\,\mathrm{id}$. Since a functional of the form $e_j - e_k$ obviously vanishes on multiples of the identity, we conclude that $\langle e_i, e_j - e_k\rangle = \delta_{ij} - \delta_{ik}$. For $\mathfrak{sl}(3,\mathbb{C})$, this immediately gives $\langle\alpha_i, \alpha_i\rangle = 2$ for $i = 1, 2$ and $\langle\alpha_1, \alpha_2\rangle = -1$. In particular, the angle between α_1 and α_2 is $\frac{2\pi}{3}$, and we obtain the picture for A2 below.



The first root system $A_1 + A_1$ in this picture is just the direct sum of two copies of A_1. The hyperplanes orthogonal to the elements of the root systems are indicated, and it is obvious that in each case the root system is invariant under the reflections in these hyperplanes. The condition that $\frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle} \in \mathbb{Z}$ can be checked directly for B_2 and G_2. We shall see later on, that these are all abstract root systems in $\mathbb{R}^2$.

3.7. Positive and simple roots, Dynkin diagram. As remarked above, the next steps in the analysis of the structure of complex semisimple Lie algebras can be done entirely in the setting of abstract root systems. To do this, one first observes that for any abstract root system $\Delta \subset V$ and $\alpha \in \Delta$, one has $s_\alpha(\alpha) = -\alpha$ and hence $-\alpha \in \Delta$. Moreover, as in the example of A_1 above, one concludes that in a reduced abstract root system these are the only multiples of α which lie in Δ. Next, we claim that for α, β in a reduced root system Δ, the only possible values for the integer $\frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle}$

are 0, ± 1, ± 2 and ± 3. Recall that the Cauchy–Schwarz inequality states that for any $v, w \in V$ we have $|\langle v, w\rangle|^2 \leq \langle v, v\rangle\langle w, w\rangle$ and equality is only possible if v and w are proportional. We may assume that α, β are not proportional (since otherwise $\beta = \pm\alpha$, so we obtain ± 2), and hence obtain $\left|\frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle}\frac{2\langle\beta,\alpha\rangle}{\langle\beta,\beta\rangle}\right| < 4$, and since both factors are integers, they both must have absolute value ≤ 3. Moreover, if the two numbers have different absolute value, then the smaller absolute value must actually be at most one, so we see that if $|\alpha| \leq |\beta|$, then $\frac{2\langle\beta,\alpha\rangle}{\langle\beta,\beta\rangle}$ must equal 0 or ± 1. Finally, one can show that part (4) of proposition 3.5 continues to hold for abstract root systems, i.e. given $\alpha, \beta \in \Delta$ then the set of all elements of Δ of the form $\beta + n\alpha$ for $n \in \mathbb{Z}$ is an unbroken string of the form $\beta - p\alpha, \beta - (p-1)\alpha, \dots, \beta + q\alpha$, where $p, q \geq 0$ and $p - q = \frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle}$. In particular, replacing β by the starting point $\beta - p\alpha$ of the string, we see that any such string consists of at most four roots.

Since for any $\alpha \in \Delta$ also $-\alpha \in \Delta$, it is a natural idea is to split the set of roots into positive and negative roots. It is better to use a slightly stronger concept, namely to introduce a notion of positivity in V, i.e. selecting a subset $V^+ \subset V \setminus \{0\}$ such that V is the disjoint union of V^+, $\{0\}$ and $-V^+$, and such that V^+ is stable under addition and multiplication by positive scalars. Once having chosen such a decomposition, we get a total ordering on V, defined by $v \leq w$ if an only if $w - v \in V^+ \cup \{0\}$. A simple way to get such an ordering is to choose a basis $\varphi_1, \dots, \varphi_n$ for the dual space V^* and define $v \in V^+$ if and only if there is an index j such that $\varphi_i(v) = 0$ for $i < j$ and $\varphi_j(v) > 0$. For example, taking the standard bases in $\mathbb{R}^2$ and in $\mathbb{R}^{2*}$, the positive subset of $\mathbb{R}^2$ is the open right half plane together with the positive y–axis. Otherwise put $(x, y) > 0$ if either $x > 0$ or $x = 0$ and $y > 0$.

Having chosen V^+, we define the set Δ^+ of *positive roots* by $\Delta^+ = \Delta \cap V^+$, so Δ is the disjoint union of Δ^+ and $-\Delta^+$. Moreover, sums of positive roots are positive. We will write $\alpha > 0$ to indicate that the root α is positive. Further, we define the subset $\Delta^0 \subset \Delta^+$ of *simple roots* as the set of those positive roots, which cannot be written as the sum of two positive roots.

Lemma. *Let Δ be an abstract root system in the Euclidean space $(V, \langle\ ,\ \rangle)$. Let Δ^+ and Δ^0 be the set of positive respectively simple roots associated to the choice of $V^+ \subset V$. Then we have:*
(1) For $\alpha, \beta \in \Delta^0$, $\alpha - \beta$ is not a root and $\langle\alpha, \beta\rangle \leq 0$.
(2) Δ^0 is a basis for V and the coefficients in the expansion of any $\alpha \in \Delta^+$ with respect to this basis are non–negative integers.

Proof. (1) If $\alpha - \beta \in \Delta$, then it either lies in Δ^+ or in $-\Delta^+$. If $\alpha - \beta$ is positive, then $\alpha = \beta + (\alpha - \beta)$ contradicts $\alpha \in \Delta^0$, in the other case $\beta = \alpha + (\beta - \alpha)$ contradicts $\beta \in \Delta^0$. Consequently, the α–string through β must have the form $\beta, \beta + \alpha, \dots, \beta + q\alpha$ for some $q \geq 0$ and $-q = \frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle} \leq 0$.
(2) The choice of V^+ gives us a total ordering on Δ, and let us enumerate the positive roots as $\beta_1, \dots, \beta_k$ such that $\beta_1 < \dots < \beta_k$. We show by induction on k, that any β_i can be written as a linear combination of simple roots with non–negative integral coefficients. Observe that $\alpha = \alpha_1 + \alpha_2$ for positive roots α, α_1 and α_2 implies that $\alpha_1, \alpha_2 < \alpha$ since $\alpha - \alpha_1 = \alpha_2 \in V^+$ and likewise for α_2. In particular, the smallest root β_1 itself has to be simple. Assume inductively that $i > 1$ and β_j has been written as appropriate linear combinations for all $j < i$. Then either β_i is simple and we are done, or there are positive roots α_1, α_2 such that $\beta_i = \alpha_1 + \alpha_2$. But then $\alpha_1, \alpha_2 < \beta_i$, so they both must equal some β_j for $j < i$, and the claim follows by induction.

Of course, this implies that any element of Δ can be written as an integral linear combination of simple roots with all coefficients having the same sign. By definition, the set Δ spans V, so we see that also Δ^0 spans V, and we only have to show that the set is linearly independent. Suppose that we write 0 as a linear combination $\sum x_i\alpha_i$ of simple roots. Renumber the $\alpha's$ in such a way that $x_1,\dots,x_i \geq 0$ and $x_{i+1},\dots,x_n \leq 0$, and we put $\beta := x_1\alpha_1 + \dots + x_i\alpha_i = -x_{i+1}\alpha_{i+1} - \dots - x_n\alpha_n$. Then we compute

$$\langle \beta, \beta \rangle = \langle \sum_{j=1}^{i} x_j\alpha_j, -\sum_{k=i+1}^{n} x_k\alpha_k \rangle = \sum_{j,k} -x_jx_k\langle \alpha_j, \alpha_k \rangle$$

Of course, the left hand side is ≥ 0. On the other hand, $x_j \geq 0$ and $x_k \leq 0$, so $-x_jx_k \geq 0$ and by part (1) $\langle \alpha_j, \alpha_k \rangle \leq 0$ since $j \neq k$. Hence the right hand side is ≤ 0, which implies that $\beta = x_1\alpha_1 + \dots + x_i\alpha_i = 0$ and also $-x_{i+1}\alpha_{i+1} - \dots - x_n\alpha_n = 0$ and in each expression all coefficients are ≥ 0. But if at least one of them would be really positive then the result would lie in V^+ and thus be nonzero. □

The next step is to associate the *Cartan matrix* to $\Delta^0 = \{\alpha_1,\dots,\alpha_n\}$. This is simply the $n \times n$–matrix $A = (a_{ij})$ with $a_{ij} := \frac{2\langle \alpha_i,\alpha_j \rangle}{\langle \alpha_i,\alpha_i \rangle}$. Clearly, this matrix depends on the numbering of the simple roots, so it is as best unique up to conjugation with permutation matrices. We already know that all a_{ij} are integers, the diagonal entries are 2, while the off–diagonal ones are ≤ 0 by part (1) of the lemma. Obviously, if $a_{ij} = 0$ then also $a_{ji} = 0$. Finally, let D be the diagonal matrix whose ith diagonal entry equals $|\alpha_i| > 0$. Then one immediately checks that the component s_{ij} of $S := DAD^{-1}$ is given by $\frac{2\langle \alpha_i,\alpha_j \rangle}{|\alpha_i||\alpha_j|}$. In particular, this matrix is symmetric. Moreover $s_{ij} = 2\langle v_i, v_j \rangle$, where for each i we put $v_i := \frac{\alpha_i}{|\alpha_i|}$. But then for any $x = (x_1,\dots,x_n) \in \mathbb{R}^n$ the expression $xSx = \sum_{i,j} x_is_{ij}x_j$ simply equals $2\sum_{ij} x_ix_j\langle v_i, v_j \rangle = 2\langle \sum x_iv_i, \sum x_jv_j \rangle$ so this is ≥ 0 and it only equals 0 if $x = 0$. Hence the symmetric matrix S is positive definite.

The final step associates to the Cartan matrix A a certain graph called the *Dynkin diagram* of the root system. One simply takes one vertex for each of the simple roots α_i, and connects the vertices corresponding to α_i and α_j by $n_{ij} := a_{ij}a_{ji}$ lines. Note that n_{ij} is always a non–negative integer and $n_{ij} = 0$ if and only if α_i and α_j are orthogonal. We have seen above that $|\alpha_i| \leq |\alpha_j|$ implies that a_{ji} equals 0 or -1, and hence in this case $a_{ij} = -n_{ij}$. If $n_{ij} = 1$, this implies that $a_{ij} = a_{ji} = -1$ and hence $|\alpha_i| = |\alpha_j|$. If $n_{ij} > 1$ (and hence equals 2 or 3), then we conclude that $|\alpha_j| = \sqrt{n_{ij}}|\alpha_i|$, and we indicate the relation of the lengths by adding an arrow which points from the longer root to the shorter root. Having done this, we conclude that we can read the relations of the lengths of the various simple roots in each connected component of the Dynkin diagram. Having these, we can of course recover the Cartan matrix from the Dynkin diagram.

Let us finally discuss irreducibility in this picture. Consider a direct sum $\Delta = \Delta_1 \sqcup \Delta_2 \subset V_1 \oplus V_2$ of two abstract root systems. Having chosen V^+, we use $V_i^+ := V^+ \cap V_i$ for $i = 1,2$ to define positivity on the two subspaces. Of course, this leads to the positive roots $\Delta^+ = \Delta_1^+ \sqcup \Delta_2^+$, and since the sum of an element of Δ_1 and an element of Δ_2 never is a root, we see that $\Delta^0 = \Delta_1^0 \sqcup \Delta_2^0$. By construction the elements of these two sets are mutually orthogonal, so numbering roots appropriately the resulting Cartan matrix of Δ^0 is block diagonal with two blocks which equal the Cartan matrices of Δ_1^0 and Δ_2^0. Clearly, this implies that the Dynkin diagram of Δ^0 is the disjoint union of the Dynkin diagrams of Δ_1^0 and Δ_0^2.

Conversely, a decomposition of the Dynkin diagram associated to $\Delta^0 \subset \Delta \subset V$ into two connected components leads to a block diagonal form of the Cartan matrix, possibly after renumbering of the roots (i.e. conjugation by a permutation matrix). Hence the simple system of the original root system decomposes into two subsets Δ_1^0 and Δ_2^0 of mutually orthogonal roots. The linear spans of these subsets decompose the vector space V into an orthogonal direct sum of two subspaces, we define Δ_i to be the set of those roots which can be written as linear combinations of elements of Δ_i^0 for $i = 1, 2$. Of course, these are disjoint subsets of Δ and we claim that the union equals Δ. If not, then there is a smallest root $\alpha \in \Delta^+$ which is not contained in the union, which means that expanding this into a linear combination of simple roots, one needs both roots from Δ_1^0 and from Δ_2^0. Writing $\alpha = \sum a_i\alpha_i$ for the expansion, all a_i are nonnegative integers, and we get $0 < |\alpha|^2 = \sum a_i\langle\alpha, \alpha_i\rangle$ so there must be a simple root α_i such that the inner product is positive. Exchanging the two subsets if necessary, we assume that $\alpha_i \in \Delta_2^0$. From the description of the α_i–string through α we conclude that $\langle\alpha, \alpha_i\rangle > 0$ implies that $\alpha - \alpha_i$ is a root, and since this is smaller than α, the only possibility is that $a_i = 1$ and $\alpha - \alpha_i \in \Delta_1$, i.e. α_i was the only element of Δ_2^0 which shows up in the expansion of α. But then $\langle\alpha - \alpha_i, \alpha_i\rangle = 0$, and thus the α_i–string through $\alpha - \alpha_i$ is symmetric. But since $\alpha - \alpha_i - \alpha_i$ certainly cannot be a root, $\alpha = \alpha - \alpha_i + \alpha_i$ cannot be a root either, which is a contradiction. Hence we see that an abstract root system is irreducible if and only if its Dynkin diagram is connected.

3.8. The classical examples. We next discuss the root systems, Cartan matrices and Dynkin diagrams of the so called classical simple Lie algebras. As we shall later see, these cover all but 5 irreducible abstract root systems.

(1) We have already described the root system of $\mathfrak{sl}(n, \mathbb{C})$ in 3.3: Taking the subspace of tracefree diagonal matrices as a Cartan subalgebra $\mathfrak{h}$, we get $\Delta = \{e_i - e_j : i \neq j\}$, where $e_i : \mathfrak{h} \to \mathbb{C}$ extracts the ith entry of a diagonal matrix. In particular, $\mathfrak{h}_0$ is the subspace of all real diagonal matrices. To define positivity, we use the basis $\{E_{11} - E_{nn}, \dots, E_{n-1,n-1} - E_{nn}\}$ of $\mathfrak{h}_0$, where as before E_{ij} denotes an elementary matrix. This immediately leads to $\Delta^+ = \{e_i - e_j : i < j\}$. The corresponding set of simple roots is then $\{\alpha_1, \dots, \alpha_{n-1}\}$ with $\alpha_i = e_i - e_{i+1}$. In terms of this set of simple roots, the positive roots are exactly all combinations of the form $\alpha_i + \alpha_{i+1} + \dots + \alpha_{i+j}$ for $i \in \{1, \dots, n-1\}$ and $0 \leq j \leq n - 1 - i$. We also have observed in 3.6 that the inner product on $\mathfrak{h}_0^*$ induced by the Killing form is given (up to a factor) by $\langle e_i, \alpha_{i-1}\rangle = -1$, $\langle e_i, \alpha_i\rangle = 1$ and $\langle e_i, \alpha_j\rangle = 0$ for $j \neq i - 1, i$. Consequently, we get $\langle\alpha_i, \alpha_i\rangle = 2$ for all $i = 1, \dots, n-1$, so all simple roots have the same length, $\langle\alpha_i, \alpha_{i+1}\rangle = -1$ and $\langle\alpha_i, \alpha_j\rangle = 0$ for $|j - i| > 1$. Thus the Cartan matrix of $\mathfrak{sl}(n, \mathbb{C})$ has each entry on the main diagonal equal to 2, in the two diagonals above and below the main diagonal all entries equal -1, while all other entries are zero. The Dynkin diagram has $n-1$ vertices, and each of these vertices is connected to its (one or two) neighbors by one edge, i.e. we get $\circ\!\!-\!\!\!-\!\!\circ\!\!-\cdots-\!\!\circ\!\!-\!\!\!-\!\!\circ$. This Cartan matrix and the Dynkin diagram are usually called A_{n-1}.

(2) The second series of classical examples of a complex simple Lie algebras is provided by the algebras $\mathfrak{so}(n, \mathbb{C})$ of matrices which are skew symmetric with respect to a non–degenerate complex bilinear form. Taking the standard from $(z, w) \mapsto \sum z_i w_i$ one simply obtains the Lie algebra of skew symmetric complex $n \times n$–matrices. This subalgebra is reductive by 2.10 and since the standard representation $\mathbb{C}^n$ is irreducible even for the subalgebra $\mathfrak{so}(n, \mathbb{R})$, we conclude that $\mathfrak{so}(n, \mathbb{C})$ is semisimple.

However, for structural purposes it is better to take a different non–degenerate bilinear form, which by linear algebra is equivalent to the standard bilinear form. Moreover, one has to distinguish between even and odd dimensions.

For even dimensions, the good choice is to use the form

$$b(z,w) := \sum_{i=1}^{n}(z_i w_{n+i} + z_{n+i}w_i)$$

on $\mathbb{C}^{2n}$. One easily verifies directly that a $2n\times 2n$–matrix is skew symmetric with respect to this form if and only if it has the block form $\begin{pmatrix} A & B \\ C & -A^T\end{pmatrix}$ with an arbitrary $n\times n$–matrix A and skew symmetric $n\times n$–matrices B and C. Consider the subspace $\mathfrak{h}$ of all diagonal matrices contained in $\mathfrak{g}$, so these have diagonal entries of the form $a_1,\dots,a_n,-a_1,\dots,-a_n$, and we denote by e_i (for $i=1,\dots,n$) the functional which takes out the ith diagonal element.

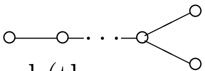
For block diagonal matrices (i.e. matrices with trivial B and C–part), the adjoint action preserves this block decomposition, and explicitly we have

$$\left[\begin{pmatrix} A & 0 \\ 0 & -A^t\end{pmatrix},\begin{pmatrix} A' & B' \\ C' & -(A')^t\end{pmatrix}\right] = \begin{pmatrix} [A,A'] & AB'+B'A^t \\ -A^tC'-C'A & -[A,A']^t\end{pmatrix}.$$

for $i\neq j$, the elementary matrix E_{ij} put in the A–block (i.e. the matrix $E_{ij}-E_{n+i,n+j}$) is an eigenvector corresponding to the eigenvalue e_i-e_j. For the other blocks, a basis is given by the elements $E_{ij}-E_{ji}$ (put into the appropriate block) for $i<j$. From the description above it follows immediately that this element in the B–block is an eigenvector for the eigenvalue e_i+e_j, while in the C–block it produces the eigenvalue $-e_i-e_j$. In particular, we conclude that $\mathfrak{h}\subset\mathfrak{g}$ is a Cartan subalgebra, so $\mathfrak{so}(2n,\mathbb{C})$ has rank n, and the set of roots is $\Delta=\{\pm e_i\pm e_j : 1\leq i<j\leq n\}$. As in (1), $\mathfrak{h}_0$ is the space of real diagonal matrices and defining positivity using the basis $\{E_{11}-E_{n+1,n+1},\dots,E_{nn}-E_{2n,2n}\}$, we get $\Delta^+ := \{e_i\pm e_j : i<j\}$. As for the simple roots, the elements $\alpha_1,\dots,\alpha_{n-1}$, defined by $\alpha_j=e_j-e_{j+1}$ clearly are simple. A moment of thought shows that among the roots e_i+e_j only $\alpha_n=e_{n-1}+e_n$ can be simple, so we have found the set $\Delta^0=\{\alpha_1,\dots,\alpha_n\}$ of simple roots.

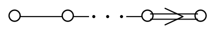
To express all positive roots as linear combinations of simple roots, one first may use all expressions of the form $\alpha_i+\alpha_{i+1}+\dots+\alpha_j$ for $i<j<n$, which give the root e_i-e_{j+1}. Next, the roots e_i+e_n for $i<n-1$ are obtained as $\alpha_i+\dots+\alpha_{n-2}+\alpha_n$. The remaining positive roots e_i+e_j for $i<j<n$ are obtained as $\alpha_i+\dots+\alpha_{j-1}+2\alpha_j+\dots+2\alpha_{n-2}+\alpha_{n-1}+\alpha_n$.

Next, we have to compute the Killing form, and as before we may equivalently use the trace form of the standard representation. This visibly is twice the standard inner product, so to determine the Cartan matrix and the Dynkin diagram we may assume that the e_i are orthonormal. In particular $\langle\alpha_j,\alpha_j\rangle=2$ for all j, so all simple roots have the same length. As for the other inner products, we see that for $j=1,\dots,n-2$ we have $\langle\alpha_j,\alpha_{j+1}\rangle=-1$, but $\langle\alpha_{n-1},\alpha_n\rangle=0$ and $\langle\alpha_{n-2},\alpha_n\rangle=-1$, while all other inner products are zero. For the Cartan matrix, this means that the upper part looks like the Cartan matrix A_n, while the bottom right 3×3–block has the form $\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 0 \\ -1 & 0 & 2\end{pmatrix}$,

and for the Dynkin diagram we get , which is usually called D_n. For $n\geq 3$, this Dynkin diagram is connected (the case $n=2$ will be discussed immediately) which implies that $\mathfrak{so}(2n,\mathbb{C})$ is simple for $n\geq 3$.

For $n = 1$, the Lie algebra $\mathfrak{so}(2,\mathbb{C})$ is isomorphic to $\mathbb{C}$ and thus not semisimple. If $n = 2$, then the roots are just $\pm e_1 \pm e_2$, so this is simply an orthogonal direct sum of two copies of A_1. (In terms of the Dunking diagram only the upper and the lower root, which are not connected, are there.) Indeed, this reflects an isomorphism $\mathfrak{so}(4,\mathbb{C}) \cong \mathfrak{sl}(2,\mathbb{C}) \times \mathfrak{sl}(2,\mathbb{C})$. Similarly, for $n = 3$ we obtain the same Dynkin diagram as for $\mathfrak{sl}(4,\mathbb{C})$, and again this reflects the fact that $\mathfrak{so}(6,\mathbb{C}) \cong \mathfrak{sl}(4,\mathbb{C})$. Describing these isomorphisms explicitly is interesting, but would lead us too far away from the main subject here.

(3) For the odd orthogonal algebras $\mathfrak{so}(2n+1,\mathbb{C})$ the situation is very similar to the even orthogonal case. The good choice for the bilinear form on $\mathbb{C}^{2n+1}$ is $b(z,w) = \sum_{i=1}^n (z_i w_{n+i} + z_{n+i} w_i) + z_{2n+1} w_{2n+1}$. In particular, this implies that we have $\mathfrak{so}(2n,\mathbb{C})$ in the form presented above included as a subalgebra, with only one additional column and line, which are negative transposes of each other. Denoting by $\mathfrak{h}$ the intersection of $\mathfrak{so}(2n+1,\mathbb{C})$ with the space of diagonal matrices we get the same subalgebra as in (2). Elements of $\mathfrak{h}$ have diagonal entries of the form $a_1, \dots, a_n, -a_1, \dots, -a_n, 0$, and we denote by e_j the functional which extracts a_j. From the subalgebra $\mathfrak{so}(2n,\mathbb{C})$ we get eigenspaces for the adjoint action of an element of $\mathfrak{h}$ with eigenvalues $\pm e_i \pm e_j$ with $i < j$, and the additional row and column are spanned by eigenvectors with eigenvalues $\pm e_j$ for $j = 1, \dots, n$. Hence we see that $\mathfrak{h}$ is a Cartan subalgebra, so $\mathfrak{so}(2n+1,\mathbb{C})$ has rank n, and the set of roots is given by $\Delta = \{\pm e_i \pm e_j : 1 \le i < j \le n\} \cup \{\pm e_j : 1 \le j \le n\}$. Defining positivity as in (2), we get $\Delta^+ = \{e_i \pm e_j : i < j\} \cup \{e_j\}$, and the simple subsystem $\Delta^0 = \{\alpha_1, \dots, \alpha_n\}$ is given by $\alpha_j = e_j - e_{j+1}$ for $j = 1, \dots, n-1$ and $\alpha_n = e_n$.

Replacing the Killing form by the trace form of the standard representation as before, we again obtain the standard form, so for determining the Cartan matrix and the Dynkin diagram we may assume that the e_i are orthonormal. In particular, for $j < n$ we get $\langle \alpha_j, \alpha_j \rangle = 2$ while $\langle \alpha_n, \alpha_n \rangle = 1$, so the last simple root is shorter than the others. On the other hand, for all $j < n$ we have $\langle \alpha_j, \alpha_{j+1} \rangle = -1$, and all other inner products are zero. Thus, the Cartan matrix looks like the one for A_n, with the only exception that the bottom right 2×2–block has the form $\begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix}$ instead of $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$. The Dynkin diagram of $\mathfrak{so}(2n+1,\mathbb{C})$ thus is given by  and this diagram is usually called B_n. Connectedness of this diagram shows that $\mathfrak{so}(2n+1,\mathbb{C})$ is simple for $n \ge 1$.

For $n = 1$ we see that $B_1 = A_1$ and this reflects an isomorphism $\mathfrak{so}(3,\mathbb{C}) \cong \mathfrak{sl}(2,\mathbb{C})$. This comes from the adjoint representation of $\mathfrak{sl}(2,\mathbb{C})$ (which has dimension three) for which the Killing form is invariant.

(4) The final classical example of complex simple Lie algebras is provided by the so called symplectic Lie algebras. These are defined as the matrices which are skew symmetric with respect to a non–degenerate skew symmetric bilinear form. Linear algebra shows that such forms only exist on spaces of even dimension, and there they are unique up to isomorphism. We choose our form on $\mathbb{C}^{2n}$ to be given by $\omega(z,w) := \sum_{i=1}^n (z_i w_{n+i} - z_{n+i} w_i)$. One easily verifies that skew symmetry with respect to this form is equivalent to a block form $\begin{pmatrix} A & B \\ C & -A^t \end{pmatrix}$, with an arbitrary $n \times n$–matrix A and symmetric $n \times n$–matrices B and C. One easily verifies directly that the matrices of this form form a Lie subalgebra $\mathfrak{sp}(2n,\mathbb{C}) \subset \mathfrak{sl}(2n,\mathbb{C})$, and visibly this subalgebra is closed under conjugate

transposition. Hence from 2.10 we conclude that $\mathfrak{sp}(2n,\mathbb{C})$ is reductive, and since the standard representation $\mathbb{C}^{2n}$ is easily seen to be irreducible, it is actually semisimple.

As before, we denote by $\mathfrak{h}$ the space of diagonal matrices contained in $\mathfrak{sp}(2n,\mathbb{C})$, so these have diagonal entries $(a_1,\dots,a_n,-a_1,\dots,-a_n)$, and we denote by e_j the linear functional that extracts a_j. The formula for the adjoint action of such an element from (2) remains valid. Then off-diagonal elementary matrices in the A–block are eigenvectors with eigenvalues $e_i - e_j$ for $i \neq j$, elementary matrices of the form $E_{ij} + E_{ji}$ (with $i \neq j$) are eigenvectors with eigenvalue $e_i + e_j$ in the B–block and with eigenvalue $-e_i - e_j$ in the C–block. Finally, the elementary matrices E_{ii} are eigenvectors with eigenvalue $2e_i$ in the B–block and with eigenvalue $-2e_i$ in the C–block.

Thus, we see that $\mathfrak{h}$ is a Cartan subalgebra, so $\mathfrak{sp}(2n,\mathbb{C})$ has rank n. Moreover, the set of roots is given by $\Delta = \{\pm e_i \pm e_j : 1 \leq i < j \leq n\} \cup \{\pm 2e_j : 1 \leq j \leq n\}$. Defining positivity as in (2), we get $\Delta^+ = \{e_i \pm e_j : i < j\} \cup \{2e_j\}$. A moment of thought shows that the set $\Delta^0 = \{\alpha_1,\dots,\alpha_n\}$ of simple roots is given by $\alpha_j = e_j - e_{j+1}$ for $j < n$ and $\alpha_n = 2e_n$.

To get the Cartan matrix and Dynkin diagram we may proceed as before and assume that the e_i are orthonormal. This implies that for $j < n$ we have $\langle \alpha_j, \alpha_j \rangle = 2$, while $\langle \alpha_n, \alpha_n \rangle = 4$, so this time there is one root longer than the others. The remaining nontrivial inner products are $\langle \alpha_j, \alpha_{j+1} \rangle = -1$ for $j < n-1$ and $\langle \alpha_{n-1}, \alpha_n \rangle = -2$. Hence, the Cartan matrix is just the transpose of the Cartan matrix B_n, and the Dynkin diagram is $\circ\!\!-\!\!\!-\!\!\circ\!\!-\cdots-\!\!\circ\!\!\Leftarrow\!\!\circ$, which is usually called C_n. Connectedness of the Dynkin diagram shows simplicity of $\mathfrak{sp}(2n,\mathbb{C})$ for all $n \geq 1$.

As for low dimensional special cases, for $n = 1$, $\mathfrak{sp}(2,\mathbb{C})$ by our description coincides with $\mathfrak{sl}(2,\mathbb{C})$. For $n = 2$ the Dynkin diagrams B_2 and C_2 coincide which reflects an isomorphism $\mathfrak{so}(5,\mathbb{C}) \cong \mathfrak{sp}(4,\mathbb{C})$.

3.9. The Weyl group. The Weyl group is the last essential ingredient in the structure theory of complex semisimple Lie algebras. This group can be introduced in the setting of abstract root systems, and we will work in that setting. Our first use of the Weyl group will be to show that the Dynkin diagram associated to an abstract root system is independent of the choice of positivity. Having given a complex semisimple Lie algebra $\mathfrak{g}$ and a Cartan subalgebra $\mathfrak{h}$, the Weyl group of the corresponding abstract root system is called the Weyl group $W = W(\mathfrak{g},\mathfrak{h})$ of $\mathfrak{g}$ (with respect to $\mathfrak{h}$).

If $\Delta \subset V$ is an abstract root system, then for any root $\alpha \in \Delta$, we have the root reflection $s_\alpha : V \to V$, and we know from 3.5 that $s_\alpha(\Delta) \subset \Delta$. Now we define the *Weyl group* $W = W(\Delta)$ of Δ to be the subgroup of the orthogonal group $O(V)$ generated by all the reflections s_α. Then any element $w \in W$ maps the root system Δ to itself. If $w \in W$ fixes all elements of Δ, then w must be the identity map, since Δ spans V. Hence we may also view W as a subgroup of the permutation group of Δ, so W is finite. For later use, we define the *sign* of an element $w \in W$ as the determinant of w, viewed as a linear automorphism of V. Since w is an orthogonal map, the sign really is either 1 or -1, depending on whether w is a product of an even or odd number of reflections.

Originally, we have first defined the positive roots Δ^+ and then obtained the subset Δ^0 of simple roots. Now one can turn around the game and define simple roots directly and then use these to define positive roots. A *simple subsystem* $\Delta^0 = \{\alpha_1,\dots,\alpha_n\} \subset \Delta$ is defined to be a basis of V consisting of roots such that any element of Δ can be written as a linear combination of the α_j with all coefficients of the same sign. Having given such a simple system, one defines the corresponding set of positive roots $\Delta^+ \subset \Delta$

as consisting of those roots for which all coefficients are non–negative. Of course, Δ is the disjoint union of Δ^+ and $-\Delta^+$.

For a given simple system Δ^0, we consider the subgroup $W' \subset W$ generated by the reflections s_{α_j} for $\alpha_j \in \Delta^0$. Now suppose that $\alpha \in \Delta^+$ is not simple, i.e. $\alpha = \sum_j a_j\alpha_j$ and all $a_j \geq 0$. Since α_j is the only positive multiple of α_j that is a root, we conclude that at least two of the coefficients a_i are nonzero. Since $0 < \langle \alpha, \alpha \rangle = \sum_j a_j \langle \alpha, \alpha_j \rangle$, we see that there is a simple root α_j such that $\langle \alpha, \alpha_j \rangle > 0$. Then $s_{\alpha_j}(\alpha) = \alpha - \frac{2\langle \alpha,\alpha_j\rangle}{\langle \alpha_j,\alpha_j\rangle}\alpha_j$ is a root. Expanding this as a linear combination of elements of Δ^0, the coefficient of α_i for $i \neq j$ equals a_i while the coefficient of α_j is obtained by subtracting a positive integer from a_j. In particular, there still must be at least one positive coefficient, so $s_{\alpha_j}(\alpha) \in \Delta^+$. Now we can iterate this, and in each step a positive integer is subtracted from one of the coefficients while all other coefficents remain the same. Thereofore, this process has to terminate, and multiplying up the reflections we obtain an element $w \in W'$ and a simple root $\alpha_k \in \Delta^0$ such that $w(\alpha) = \alpha_k$, and thus $\alpha = w^{-1}(\alpha_k)$. Since $s_{\alpha_k}(\alpha_k) = -\alpha_k$, we see that $-\alpha = (w^{-1} \circ s_{\alpha_k})(\alpha_k)$, and thus $\Delta = W'(\Delta^0)$. Hence we can recover Δ from the simple subsystem Δ^0. Moreover, the above computation shows that the coefficients in the expansion of $\alpha \in \Delta^+$ as a linear combination of elements of Δ^0 are non–negative integers.

On the other hand, if $\alpha = w(\alpha_j)$ then

$$s_\alpha(v) = v - \frac{2\langle \alpha, v\rangle}{\langle \alpha, \alpha\rangle}\alpha = w\left(w^{-1}(v) - \frac{2\langle \alpha_j, w^{-1}(v)\rangle}{\langle \alpha_j, \alpha_j\rangle}\alpha_j\right)$$

and thus $s_{w(\alpha_j)} = w \circ s_{\alpha_j} \circ w^{-1} \in W'$. Hence we conclude that $W' = W$, i.e. the Weyl group is generated by the root reflections corresponding to elements of Δ^0.

Now it is almost obvious that for a simple system $\{\alpha_1, \dots, \alpha_n\}$ and some element $w \in W$, also $\{w(\alpha_1), \dots, w(\alpha_n)\}$ is a simple system. (To obtain a representation of α in terms of the $w(\alpha_j)$ use a representation of $w^{-1}(\alpha) \in \Delta$ in terms of the α_j.) Thus, fixing some simple subsystem Δ^0, we define a mapping from W to the set of all simple subsystems of Δ by mapping w to $w(\Delta^0)$. We claim, that this map is a bijection:

Let $A \subset \Delta$ be any simple subsystem and let $D^+ \subset \Delta$ be the corresponding set of positive roots. Suppose first that $A \subset \Delta^+$. Then of course $D^+ \subset \Delta^+$, so the two sets must actually be equal, and in particular $\Delta^0 \subset D^+$. Hence any element of Δ^0 can be written as a linear combination of elements of A with coefficients in $\mathbb{N}$ and vice versa, and this immediately implies $\Delta^0 = A$. If A is not contained in Δ^+, then take $\alpha \in A \cap -\Delta^+$. Of course, $s_\alpha(\alpha) = -\alpha \in s_\alpha(D^+) \cap \Delta^+$. On the other hand, the above computation shows that if $\beta \in D^+$ and $\beta \neq \alpha$ then $s_\alpha(\beta) \in D^+$. Hence we conclude that $s_\alpha(D^+)$ is obtained from D^+ simply by replacing α by $-\alpha$. Therefore, the intersection of Δ^+ and $s_\alpha(D^+)$, which is the positive system associated to the simple system $s_\alpha(A)$, is strictly larger than $\Delta^+ \cap D^+$. Inductively, this implies that we can find an element $w \in W$ such that $w(D^+) = \Delta^+$, and we have seen that this implies $w(A) = \Delta^0$ and thus $A = w^{-1}(\Delta^0)$. Hence we conclude that the map $w \mapsto w(\Delta^0)$ is surjective onto the set of simple subsystems.

Since each $w \in W$ is an orthogonal mapping and the Cartan matrix and the Dynkin diagram are constructed from the mutual inner products of simple roots, this implies that they are independent of the choice of the simple (or positive) subsystem.

For later use we also verify that the mapping $w \mapsto w(\Delta^0)$ is injective: For $\Delta^0 = \{\alpha_1, \dots, \alpha_n\}$ let us temporarily write s_i for the reflection s_{α_i}. Take $w \in W$ such that $w(\Delta^0) = \Delta^0$ and write $w = s_{i_m} \cdots s_{i_1}$ as a product of simple reflections. By construction, $w(\alpha_{i_1}) \in \Delta^0 \subset \Delta^+$ but $s_{i_1}(\alpha_{i_1}) = -\alpha_{i_1} \in -\Delta^+$. Therefore, we find an index r with $2 \leq$

$r \leq m$ such that putting $w' := s_{i_{r-1}} \cdots s_{i_1}$ we get $w'(\alpha_{i_1}) \in -\Delta^+$ but $(s_{i_r} w')(\alpha_{i_1}) \in \Delta^+$. But since α_{i_r} is the only positive root that becomes negative under s_{i_r} we see that this is only possible if $w'(\alpha_{i_1}) = -\alpha_{i_r}$. But then $s_{i_r} = s_{w'(\alpha_{i_1})} = w' s_{i_1} (w')^{-1}$, and therefore $s_{i_r} w' = w' s_{i_1} = s_{i_{r-1}} \cdots s_{i_2}$. But this implies that we can write w also as a product of $m-2$ simple reflections. If m is even, we come down to 0 reflections and thus $w = \mathrm{id}$. If m is odd we conclude that w is a simple reflection, which is a contradiction since $s_i(\alpha_i) = -\alpha_i \notin \Delta^0$.

To conclude this section, let us briefly describe the Weyl groups for the classical examples from 3.8:

For the root system A_{n-1} of $\mathfrak{sl}(n,\mathbb{C})$, the dual space $\mathfrak{h}_0^*$ is the quotient of the space of all $\sum a_j e_j$ by the line generated by $e_1 + \cdots + e_n$, so we may view it as the space of all $\sum a_j e_j$ such that $\sum a_j = 0$. One immediately verifies that the root reflection $s_{e_i - e_j} : \mathfrak{h}_0^* \to \mathfrak{h}_0^*$ is induced by the map which exchanges e_i and e_j and leaves the e_k for $k \neq i, j$ untouched. This immediately implies that the Weyl group of A_{n-1} is the permutation group $\mathfrak{S}_n$ of n elements.

For the root system B_n of $\mathfrak{so}(2n+1,\mathbb{C})$, the reflections in $e_i - e_j$ again exchanges e_i and e_j, while the reflection in e_j changes the sign of e_j and leaves the other e_k untouched. Thus, we may view the Weyl group W as the subgroup of all permutations σ of the $2n$ elements $\pm e_j$ such that $\sigma(-e_j) = -\sigma(e_j)$ for all $j = 1, \dots, n$. As a set, W is the product of $\mathfrak{S}_n$ and n copies of $\mathbb{Z}_2$, so in particular, it has $2^n n!$ elements. Since from the point of view of the induced reflections, there is no difference between e_j and $2e_j$, we obtain the same Weyl group for the root system C_n.

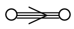
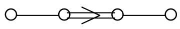
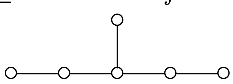
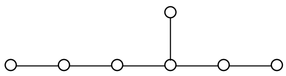
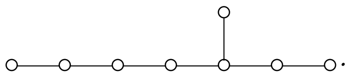
Finally for the even orthogonal root system D_n, the reflections in the roots $e_i - e_j$ again generate permutations of the e_j, while the reflection in $e_i + e_j$ maps e_i to $-e_j$ and e_j to $-e_i$ while all other e_k remain untouched. Consequently, W can be viewed as the subgroup of those permutations σ of the elements $\pm e_j$ which satisfy $\pi(-e_j) = -\pi(e_j)$ and have the property that the number of j such that $\pi(e_j) = -e_k$ for some k is even. In particular, the number of elements in W equals $2^{n-1} n!$.

The classification of root systems and complex simple Lie algebras

We have now developed all the ingredients of the structure theory of complex semisimple Lie algebras, which also form the basis for studying representations. To conclude this chapter we briefly discuss how one can obtain a complete classification of complex simple Lie algebras from these data. In particular, this shows that apart from the classical examples discussed in 3.8 there are (up to isomorphism) only 5 other complex simple Lie algebras.

3.10. The classification of abstract root systems. We have already noted in 3.9 that an abstract root system $\Delta \subset V$ is completely determined by any simple subsystem Δ^0. More precisely, suppose that $\Delta_i \subset V_i$ are abstract root systems for $i = 1, 2$ with simple subsystems Δ_i^0 and we can find a linear isomorphism $\varphi : V_1 \to V_2$ such that $\varphi(\Delta_1^0) = \Delta_2^0$ and $\frac{\langle \varphi(\alpha), \varphi(\beta) \rangle}{\langle \varphi(\alpha), \varphi(\alpha) \rangle} = \frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}$ for all $\alpha, \beta \in \Delta_1^0$. Since Δ_1^0 spans V, the corresponding equation also holds with β replaced by any $v \in V$, which immediately implies that $s_{\varphi(\alpha)} = \varphi \circ s_\alpha \circ \varphi^{-1}$. Since the simple reflections generate the Weyl group, we conclude that $w \mapsto \varphi \circ w \circ \varphi^{-1}$ defines a bijection between the Weyl groups W_1 and W_2. In 3.9 we have observed that $\Delta_1 = \{w(\alpha) : w \in W_1, \alpha \in \Delta_1\}$, and $\varphi(w(\alpha)) = (\varphi \circ w \circ \varphi^{-1})(\varphi(\alpha))$ then immediately implies that $\varphi(\Delta_1) = \Delta_2$. Finally, for $\alpha, \beta \in \Delta_1^0$ and $w, w' \in W_1$ we have $\frac{2\langle w(\alpha), w'(\beta) \rangle}{\langle w(\alpha), w(\alpha) \rangle} = \frac{2\langle \alpha, w^{-1} w'(\beta) \rangle}{\langle \alpha, \alpha \rangle}$. We have observed above, that this equals $\frac{2\langle\varphi(\alpha),\varphi(w^{-1}w'(\beta))\rangle}{\langle\varphi(\alpha),\varphi(\alpha)\rangle}$, which in turn is immediately see to equal $\frac{2\langle\varphi(w(\alpha)),\varphi(w'(\beta))\rangle}{\langle\varphi(w(\alpha)),\varphi(w(\alpha))\rangle}$. Thus we conclude that φ is an isomorphism of root systems as defined in 3.6.

In conclusion, the classification of (irreducible) reduced abstract root systems boils down to the classification of (connected) Dynkin diagrams:

THEOREM. *Let V be a finite dimensional Euclidean vector space and $\Delta \subset V$ a finite dimensional irreducible reduced abstract root system. Then the Dynkin diagram of Δ is A_n for $n \geq 1$, B_n for $n \geq 2$, C_n for $n \geq 3$ or D_n for $n \geq 4$ or one of the following five exceptional diagrams: G_2:* , *F_4:* , *E_6:* ,
E_7: , *E_8:* .

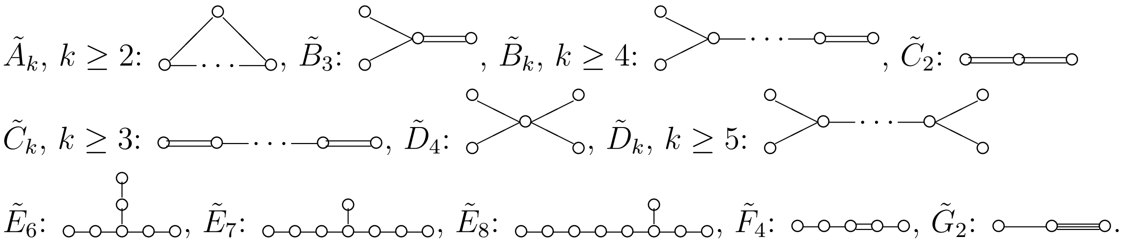
SKETCH OF PROOF. Although this result is very surprising, the proof is completely elementary, but needs a certain amount of explicit verifications. It is only based on the fact that the Dynkin diagram directly leads to an expression of the inner product on V in terms of a certain basis: Consider the Cartan matrix A of Δ. We have observed that there is a diagonal matrix D such that $S := DAD^{-1}$ is the (symmetric) matrix of mutual inner products of the elements of a certain basis of V. From the description in 3.8 it follows immediately that the coefficients s_{ij} of S are given by $s_{ii} = 2$ and $s_{ij} = -\sqrt{n_{ij}}$ for $i \neq j$, where n_{ij} is the number of edges in the Dynkin diagram joining the ith and jth node. In particular, this matrix can be immediately read off from the Dynkin diagram, and the direction of arrows is irrelevant for the matrix S. Recall that we have already verified that $n_{ij} \leq 3$ for all i, j.

It is elementary to verify that the matrix obtained from S by replacing some of the n_{ij} by smaller numbers still describes a positive definite inner product.

It follows from linear algebra that there is an invertible matrix R such that RSR^T is diagonal with positive entries, so in particular $\det(S) > 0$. Moreover, if we take out any block in S of the form $\begin{pmatrix} s_{ii} & \dots & s_{ij} \\ \vdots & & \vdots \\ s_{ji} & \dots & s_{jj} \end{pmatrix}$ then we can apply the same argument replacing V with the space spanned by the simple roots $\alpha_i, \dots, \alpha_j$, so the determinant must be positive, too.

The upshot of this discussion is that if we have given some Diagram and we define a subdiagram as one that is obtained by replacing some multiple edge by a fewer number of edges and then taking a connected subgraph, then the result still has to give rise to the matrix of a positive definite inner product.

Conversely, if we find a diagram such that the associated matrix S has zero determinant, then this cannot occur as a subdiagram of the Dynkin diagram of any abstract root system. There is a conceptual way to produce such a list: Suppose that S is the matrix of a positive definite inner product in a basis $\{v_1, \dots, v_n\}$. Then let v_{n+1} be any nonzero linear combination of the Basis vectors v_i and consider the $(n+1) \times (n+1)$ matrix $\tilde{S}$ associated to the system $\{v_1, \dots, v_{n+1}\}$ in the same way as S is associated to $\{v_1, \dots, v_n\}$. Then of course, the lines of $\tilde{S}$ are linearly dependent, so this has zero determinant. In general, $\tilde{S}$ will not be associated to a diagram, but it turns out that this is always the case if one starts with a root system Δ corresponding to S and takes as the additional element the largest root (in the given ordering). This leads to the following diagrams (where the diagram with index k has $k+1$ vertices):



As an alternative to the approach sketched above, one may of course verify directly that the matrix associated to each of the above diagrams has zero determinant. The fact that none of the above diagrams can be a subdiagram of the Dynkin diagram of an abstract root system actually suffices to push through the classification: First, the Dynkin diagram cannot contain any cycles, since otherwise it would contain a subdiagram of type $\tilde{A}_k$. Next, at any point there may be at most three edges, since otherwise get a subdiagram of type $\tilde{G}_2$, $\tilde{C}_2$, $\tilde{B}_3$ or $\tilde{D}_4$. In particular, if there is a triple edge, G_2 is the only possibility. Let us call a point in which three single edges meet a branch point. If there is a double edge, then there are no branch points, since otherwise some $\tilde{B}_k$ would be a subdiagram. Moreover, there may be at most one double edge since otherwise there is a subdiagram of type $\tilde{C}_k$. Finally, apart from the case of F_4, only one of the two ends of a double edge may be connected to other points, since otherwise we get $\tilde{F}_4$ as a subdiagram. But this shows that in the case of a double edge, only B_n, C_n and F_4 are possible.

Hence we are left with the case that there are no multiple edges. If there is no branch point, then we get a diagram of type A_n. On the other hand, there may be at most one branch point since otherwise we get some $\tilde{D}_k$ as a subdiagram. Thus we are left with the case of three chains meeting in one branch point. Since $\tilde{E}_6$ cannot be a subdiagram, one of these three chains can only consist of a single point. From the remaining two chains at least one has to consist of at most two points, since otherwise $\tilde{E}_7$ is a subdiagram. If there is one chain with two points and one chain with one point, then we only get the possibilities D_5, E_6, E_7 and E_8, since a longer third chain would give a subdiagram of type $\tilde{E}_8$. Finally, if two of the three chains meeting at the branch point consists only of a single point, we get a diagram of type D_n. □

Remark. To obtain the diagrams of type $\tilde{E}$ and $\tilde{F}$ as described in the proof above, one has to know the existence of root systems of type F_4, E_6, E_7, and E_8. These root systems can be constructed directly, but their existence will also follow from the discussion below. There always is the possibility to verify directly that the associated matrices have zero determinant.

3.11. The classification of complex simple Lie algebras. We have seen how to pass from a Lie algebra to a root system and further to a Dynkin diagram. We also know from 3.2 and 3.9 that this Dynkin diagram does not depend on the choices of a Cartan subalgebra and a set of positive roots. In particular, this implies that isomorphic Lie algebras lead to the same Dynkin diagram. Hence the remaining questions are on one hand, whether there exist complex simple Lie algebras corresponding to the exceptional Dynkin diagrams of type E, F and G from Theorem 3.10, and on the other hand, whether two Lie algebras having the same Dynkin diagram must already be isomorphic. These question can be answered (positively) at the same time by giving a universal construction for a simple Lie algebra with a given Dynkin diagram, using the so called Serre relations.

Let us start from a complex simple Lie algebra $\mathfrak{g}$ with a chosen Cartan subalgebra $\mathfrak{h}$, the corresponding set Δ of roots and a chosen simple subsystem $\Delta^0 = \{\alpha_1, \dots, \alpha_n\}$. For any $j = 1, \dots, n$ choose elements E_j and F_j in the root spaces $\mathfrak{g}_{\alpha_j}$ respectively $\mathfrak{g}_{-\alpha_j}$ such that $B(E_j, F_j) = \frac{2}{\langle \alpha_j, \alpha_j \rangle}$. Recall from 3.5 that this exactly means that $H_j := [E_j, F_j]$ satisfies $\alpha_j(H_j) = 2$, so $\{E_j, F_j, H_j\}$ is a standard basis for the subalgebra $\mathfrak{s}_{\alpha_j} \cong \mathfrak{sl}(2, \mathbb{C})$. Moreover, the elements H_j for $j = 1, \dots, n$ span the Cartan subalgebra $\mathfrak{h}$. On the other hand, we know from 3.5 that $\mathfrak{g}_{\alpha+\beta} = [\mathfrak{g}_\alpha, \mathfrak{g}_\beta]$ for all $\alpha, \beta \in \Delta$, which together with the above easily implies that $\{E_j, F_j, H_j : 1 \leq j \leq n\}$ is a set of generators for the Lie algebra $\mathfrak{g}$. Such a set of generators is called a set of *standard generators* for $\mathfrak{g}$.

Next, there are some obvious relations. Since all H_j lie in $\mathfrak{h}$, we have $[H_i, H_j] = 0$ for all i, j. By definition and the fact that the difference of two positive roots is not a root, we further have $[E_i, F_j] = \delta_{ij} H_i$. Next, by definition of the Cartan matrix $A = (a_{ij})$ of $\mathfrak{g}$, we have $[H_i, E_j] = a_{ij} E_j$ and $[H_i, F_j] = -a_{ij} F_j$. Finally the formula for the length of the α_i–string through α_j from 3.5 implies that $\mathrm{ad}(E_i)^{-a_{ij}+1}(E_j) = 0$ and $\mathrm{ad}(F_i)^{-a_{ij}+1}(F_j) = 0$ for $i \neq j$. These six families of relations are called the *Serre relations* for $\mathfrak{g}$. The essential point for our questions now is that this is a complete set of relations.

To explain this, we need a short detour. There is a universal construction of free Lie algebras. This means that to any set X, one may associate a Lie algebra $\mathcal{L}(X)$, called the *free Lie algebra* generated by X. This Lie algebra comes with an injective set map $i : X \to \mathcal{L}(X)$ which has the following universal property: For any Lie algebra $\mathfrak{g}$ and any set map $f : X \to \mathfrak{g}$ there is a unique homomorphism $\tilde{f} : \mathcal{L}(X) \to \mathfrak{g}$ of Lie algebras such that $\tilde{f} \circ i = f$. We will describe a construction for free Lie algebras in 4.9. One may think of $\mathcal{L}(X)$ as the vector space spanned by all formal brackets of elements of X with the only equalities between brackets being those which are forced by skew symmetry and the Jacobi identity.

Now assume that $\mathfrak{g}$ is any Lie algebra and $X \subset \mathfrak{g}$ is a subset of elements which generates $\mathfrak{g}$. Then the inclusion of X into $\mathfrak{g}$ induces a Lie algebra homomorphism $\varphi : \mathcal{L}(X) \to \mathfrak{g}$. Since the image of this homomorphism is a Lie subalgebra containing X, the homomorphism φ is surjective. Hence putting $I = \ker(\varphi)$ we get an ideal in $\mathcal{L}(X)$ such that $\mathfrak{g} \cong \mathcal{L}(X)/I$. A complete set of relations for the generating set X by definition is a subset of I which generates I as an ideal. This can be used to obtain presentations of Lie algebras by generators and relations.

The result on completeness of the Serre–relations can now be more precisely stated as:

Theorem. *Let $A = (a_{ij})$ be the Cartan matrix associated to one of the Dynkin diagrams from Theorem 3.10 with n nodes. Let $\mathfrak{F}$ be the free complex Lie algebra generated by $3n$ elements E_j, F_j and H_j for $j = 1, \dots, n$, and let $\mathfrak{R}$ be the ideal generated by the Serre–relations. Then $\mathfrak{g} := \mathfrak{F}/\mathfrak{R}$ is a finite dimensional simple Lie algebra. The elements H_j span a Cartan subalgebra of $\mathfrak{g}$, and the functionals $\alpha_j \in \mathfrak{h}^*$ defined by $\alpha_j(H_i) = a_{ij}$ are a simple subsystem of the corresponding root system. In particular $\mathfrak{g}$ has the given Dynkin diagram.*

Proof. The proof is rather involved, see [**Knapp**, II.9–II.11]. Indeed, $\mathfrak{F}$ comes with a kind of root decomposition and it is easy to see that this descends to a root decomposition of $\mathfrak{F}/\mathfrak{R}$. The most difficult part of the proof is to show that there are only finitely many roots and any root space is finite dimensional, which in turn implies that $\mathfrak{F}/\mathfrak{R}$ is finite dimensional. $\square$

Corollary. *(1) Any irreducible reduced abstract root system is isomorphic to the root system of some finite dimensional complex simple Lie algebra.*
(2) Two complex simple Lie algebras are isomorphic if and only if their root systems are isomorphic, i.e. if and only if they have the same Dynkin diagram.

Proof. (1) is obvious from the theorem in view of the bijective correspondence between Cartan matrices and reduced irreducible abstract root systems.
(2) Let $\mathfrak{g}$ be any complex simple Lie algebra and let A be its Cartan matrix. Let $\mathfrak{F}/\mathfrak{R}$ be the Lie algebra constructed in the theorem. By the universal property of a free Lie algebra choosing a set of standard generator for $\mathfrak{g}$ gives a surjective homomorphism $\mathfrak{F} \to \mathfrak{g}$, which factors to $\mathfrak{F}/\mathfrak{R}$ since the Serre–relations hold in $\mathfrak{g}$. But from the theorem we know that $\mathfrak{F}/\mathfrak{R}$ is simple, which implies that this homomorphism must be injective, and thus an isomorphism. □

Remark. While the theorem asserts the existence of the exceptional complex simple Lie algebras, i.e. Lie algebras corresponding to the exceptional Dynkin diagrams of types E, F, and G, it does not offer a good description of these. Also, the dimensions of the exceptional algebras are high enough (14 for G_2, 52 for F_4, 78 for E_6, 133 for E_7, and 248 for E_8) to make explicit descriptions rather complicated.

There are various ways to explicitly construct the exceptional Lie algebras, which are related to other exceptional objects in mathematics, in particular to the eight dimensional non–associative division algebra $\mathbb{O}$ of octonions or Cayley numbers and the exceptional Jordan algebra of dimension 27, which can be constructed from $\mathbb{O}$.

The exceptional simple Lie algebras also provide input for other branches of mathematics. For example, the compact real form of G_2 is among the exceptional holonomy groups of Riemannian manifolds. On the other hand, the Dynkin diagram E_8 leads to a $\mathbb{Z}$–valued symmetric bilinear form on $\mathbb{Z}^8$, which turns out to be non–diagonalizable over $\mathbb{Z}$ (although it has to be diagonalizable over $\mathbb{R}$). This is related to the existence of topological manifolds of dimension four which do not admit any smooth structure and in turn to the existence of exotic smooth structures on $\mathbb{R}^4$.

CHAPTER 4

Representation theory of complex semisimple Lie algebras

Building on the structure theory of complex semisimple Lie algebras developed in the last chapter, we will next study complex representations of complex semisimple Lie algebras. Since we have enough background at hand, we will fairly quickly reach the fundamental result that any irreducible representations is determined up to isomorphism by its highest weight, which is a certain functional on a Cartan subalgebra. It is also easy to describe the properties of such highest weights. The basic tool for this is the representation theory of $\mathfrak{sl}(2,\mathbb{C})$ that we have developed in the last chapter. A more involved problem is to show that for any weight having these properties there actually exists an irreducible representation having the given weight as its highest weight. We will discuss two approaches to this problem. Both these approaches need some background from multilinear algebra which we will develop.

The theorem of the highest weight

4.1. The weights of a finite dimensional representation. Let us fix a complex semisimple Lie algebra $\mathfrak{g}$, a Cartan subalgebra $\mathfrak{h} \leq \mathfrak{g}$ and an order on $\mathfrak{h}_0^*$. Let Δ, Δ^+ and Δ^0 be the corresponding sets of roots, positive roots, and simple roots, respectively. Let $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of $\mathfrak{g}$ on a finite dimensional complex vector space V. From 3.2 we know that the elements of $\mathfrak{h}$ act simultaneously diagonalizable on V. The eigenvalues, called the weights of V, are given by linear functionals $\lambda : \mathfrak{h} \to \mathbb{C}$, and the corresponding eigenspace $V_\lambda = \{v \in V : H \cdot v = \lambda(H)v \quad \forall H \in \mathfrak{h}\}$ is called the weight space of weight λ in V. Denoting by $\mathrm{wt}(V)$ the set of all weights of V, we obtain the decomposition $V = \oplus_{\lambda\in\mathrm{wt}(V)} V_\lambda$. The dimension of V_λ is called the *multiplicity* of the weight λ in V. By definition of a weight space, for $v \in V_\lambda$ and $X \in \mathfrak{g}_\alpha$, we have $X \cdot v \in V_{\lambda+\alpha}$ (so in particular $X \cdot v = 0$ if $\lambda + \alpha$ is not a weight of V).

Consider a fixed positive root $\alpha \in \Delta^+$ and let $\mathfrak{s}_\alpha = \mathfrak{g}_{-\alpha} \oplus [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \oplus \mathfrak{g}_\alpha \cong \mathfrak{sl}(2,\mathbb{C})$ be the corresponding subalgebra. For a weight $\lambda \in \mathrm{wt}(V)$ consider the sum of all weight spaces of the form $\lambda + n\alpha$ for $n \in \mathbb{Z}$. From above we see that this subspace is invariant under the action of $\mathfrak{s}_\alpha$, so we may apply the representation theory of $\mathfrak{sl}(2,\mathbb{C})$ as developed in 3.4. Recall from 3.5 that the standard basis element in the Cartan subalgebra of $\mathfrak{s}_\alpha$ is given by $\frac{2}{\langle\alpha,\alpha\rangle} H_\alpha \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$. By Proposition 3.4 this has integral eigenvalues on any finite dimensional representation of $\mathfrak{s}_\alpha$. Thus, we conclude that $\frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle} \in \mathbb{Z}$ for any weight λ of a finite dimensional representation and any root $\alpha \in \Delta$. Elements of $\mathfrak{h}^*$ having this property are called *algebraically integral*. In particular, any weight of a finite dimensional representation of $\mathfrak{g}$ takes real values on the elements $H_\alpha \in \mathfrak{h}$, so it is contained in the real subspace $\mathfrak{h}_0^* \subset \mathfrak{h}^*$ spanned by Δ.

Now suppose that $\frac{2\langle\lambda,\alpha_j\rangle}{\langle\alpha_j,\alpha_j\rangle} \in \mathbb{Z}$ for some $\lambda \in \mathfrak{h}_0^*$ and all simple roots α_j. Consider the reflection s_{α_k} corresponding to the simple root α_k. Inserting the definition of the root reflection, we see that

$$\frac{2\langle s_{\alpha_k}(\lambda), \alpha_j\rangle}{\langle\alpha_j,\alpha_j\rangle} = \frac{2\langle\lambda,\alpha_j\rangle}{\langle\alpha_j,\alpha_j\rangle} - \frac{2\langle\lambda,\alpha_k\rangle}{\langle\alpha_k,\alpha_k\rangle}\frac{2\langle\alpha_k,\alpha_j\rangle}{\langle\alpha_j,\alpha_j\rangle},$$

so this is again an integer. Recalling from 3.9 that the Weyl group W of $\mathfrak{g}$ is generated by the reflections corresponding to simple roots, we conclude that $\frac{2\langle w(\lambda),\alpha_j\rangle}{\langle\alpha_j,\alpha_j\rangle}\in\mathbb{Z}$ for all $w\in W$ and all $\alpha_j\in\Delta^0$. But now for an arbitrary root $\alpha\in\Delta$ we know from 3.9 that there exists a simple root α_j and an element $w\in W$ such that $\alpha=w(\alpha_j)$. Since w is an orthogonal map, $\frac{2\langle w^{-1}(\lambda),\alpha_j\rangle}{\langle\alpha_j,\alpha_j\rangle}\in\mathbb{Z}$ immediately implies $\frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle}\in\mathbb{Z}$. Thus we see that $\lambda\in\mathfrak{h}_0^*$ is algebraically integral if and only if $\frac{2\langle\lambda,\alpha_j\rangle}{\langle\alpha_j,\alpha_j\rangle}$ is an integer for all $\alpha_j\in\Delta^0$.

If $\Delta^0=\{\alpha_1,\dots,\alpha_n\}$, then we define the *fundamental weights* $\omega_1,\dots,\omega_n\in\mathfrak{h}_0^*$ for $\mathfrak{g}$ by $\frac{2\langle\omega_i,\alpha_j\rangle}{\langle\alpha_j,\alpha_j\rangle}=\delta_{ij}$. Then the set Λ_W of all algebraically integral elements of $\mathfrak{h}_0^*$ coincides with the set of integral linear combinations of the fundamental weights. Since the simple roots form a basis for $\mathfrak{h}_0^*$, the fundamental weights form a basis, too. Thus Λ_W is a lattice (i.e. a subgroup isomorphic to $\mathbb{Z}^n$), the *weight lattice* of $\mathfrak{g}$.

From the above we conclude that any element $\lambda\in\mathfrak{h}_0^*$ is uniquely determined by the n numbers $\frac{2\langle\lambda,\alpha_j\rangle}{\langle\alpha_j,\alpha_j\rangle}$ for $j=1,\dots,n$, which are exactly the coefficients in a presentation of λ as a linear combination of the fundamental weights. Moreover, λ is algebraically integral if and only if all these numbers are integers.

Let us return to the representation of $\mathfrak{s}_\alpha\cong\mathfrak{sl}(2,\mathbb{C})$ on the subspace $\oplus_{n\in\mathbb{Z}}V_{\lambda+n\alpha}$. From 3.4 we know that any finite dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$ decomposes into a direct sum of irreducible representations, and for these the eigenvalues form an unbroken string of integers of difference two, which is symmetric around zero. So in particular, since $\frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle}$ is an eigenvalue, also its negative has to show up. Hence we conclude that choosing $k\in\mathbb{Z}$ in such a way that $\frac{2\langle\lambda-k\alpha,\alpha\rangle}{\langle\alpha,\alpha\rangle}=-\frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle}$, then $\lambda-k\alpha$ must be a weight of V. But the left hand side equals $\frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle}-2k$, so we conclude that $k=\frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle}$, and thus $\lambda-k\alpha=s_\alpha(\lambda)$, where $s_\alpha:\mathfrak{h}_0^*\to\mathfrak{h}_0^*$ is the root reflection corresponding to the root α.

Replacing α by $-\alpha$ if necessary, we may assume that $\frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle}=k\geq 0$, and conclude from the fact that the eigenvalues form an unbroken string that $\lambda-2\ell\alpha$ is a weight of V for all $0\leq\ell\leq k$. Moreover, the multiplicity of this weight is at least as big as the multiplicity of λ. In particular, we conclude that the weights λ and $s_\alpha(\lambda)$ occur with the same multiplicity. Of course, this immediately implies that for any element w of the Weyl group W of $\mathfrak{g}$ and any weight λ also $w(\lambda)$ is a weight of $\mathfrak{g}$, which has the same multiplicity as λ. Hence the set $\mathrm{wt}(V)$ of weights (including multiplicities) is invariant under the action of the Weyl group. Let us collect the information on the weights of finite dimensional representations of $\mathfrak{g}$:

PROPOSITION. *Let V be a finite dimensional representation of a complex semisimple Lie algebra $\mathfrak{g}$. Then we have:*
(1) Any weight λ of V is algebraically integral, i.e. $\frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle}\in\mathbb{Z}$ for any root α of $\mathfrak{g}$.
(2) If λ is a weight of V and w is any element of the Weyl group W of $\mathfrak{g}$, then $w(\lambda)$ is a weight of V which occurs with the same multiplicity as λ.
(3) If λ is a weight of V and α is a root such that $k:=\frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle}\geq 0$, then for any $\ell\in\mathbb{N}$ such that $0\leq\ell\leq k$ the functional $\lambda-2\ell\alpha$ is a weight of V whose multiplicity is at least as big as the multiplicity of λ.

4.2. Highest weight vectors. Let V be any representation of $\mathfrak{g}$ such that the action of any element of the Cartan subalgebra $\mathfrak{h}$ is diagonalizable. A *highest weight vector* in V is a weight vector $v\in V$ such that $X\cdot v=0$ for any element X lying in a root space $\mathfrak{g}_\alpha$ with $\alpha\in\Delta^+$.

Consider a set $\{E_i, F_i, H_i : i = 1, \dots, n\}$ of standard generators for $\mathfrak{g}$, see 3.11. For any highest weight vector $v \in V$ and each i we have $E_i \cdot v = 0$ since E_i is an element of the root space of a simple root. Conversely, recall from 3.5 that for roots α and β such that also $\alpha + \beta$ is a root, we have $\mathfrak{g}_{\alpha+\beta} = [\mathfrak{g}_\alpha, \mathfrak{g}_\beta]$. This immediately implies that any weight vector $v \in V$ such that $E_i \cdot v = 0$ for all i is a highest weight vector.

If $v \in V$ is a highest weight vector, consider the subspace W of V generated by all elements of the form $F_{i_1} \cdot F_{i_2} \cdot \dots \cdot F_{i_\ell} \cdot v$. By definition, W is invariant under the action of all F_j, and any H_j acts by a scalar on v. From the Serre relations in 3.11 we conclude that $H_j \cdot (F_{i_1} \cdots F_{i_\ell} \cdot v)$ is a linear combination of $F_{i_1} \cdot H_j \cdot F_{i_2} \cdots F_{i_\ell} \cdot v$ and $F_{i_1} \cdots F_{i_\ell} \cdot v$. By induction, we conclude that W is invariant under the action of all H_j. Similarly, $E_j \cdot (F_{i_1} \cdots F_{i_\ell} \cdot v)$ is a linear combination of $F_{i_1} \cdot E_j \cdot F_{i_2} \cdots F_{i_\ell} \cdot v$ and $H_j \cdot F_{i_2} \cdots F_{i_\ell} \cdot v$. Again by induction, we see that W is invariant under the action of all the standard generators, and thus it is invariant under the action of the whole Lie algebra $\mathfrak{g}$. In particular, if V is irreducible, then $W = V$, so any element of V may be written in the above form. This implies on the one hand that the highest weight vector is unique up to nonzero complex multiples. On the other hand, if λ is the weight of v then $F_{i_1} \cdot F_{i_2} \cdot \dots \cdot F_{i_\ell} \cdot v$ is a weight vector of weight $\lambda - \alpha_{i_1} - \dots - \alpha_{i_\ell}$, so any weight of V is obtained from λ by subtracting a linear combination of simple roots with non–negative integral coefficients.

There is an important observation on invariant subspaces of a representation ρ of $\mathfrak{g}$ on V as considered above, which holds even in the infinite dimensional case: We have assumed that $\mathfrak{h}$ acts diagonalizably on V which means that V is a direct sum of weight spaces. Thus any element $v \in V$ may be written as a finite sum of weight vectors. Fix $H \in \mathfrak{h}$, let $v = v_1 + \dots + v_k$ be the decomposition of v into eigenvectors with different eigenvalues for the action of $\rho(H)$ and let a_i be the eigenvalue of v_i. Consider the operator $\prod_{j \neq i} \frac{1}{(a_i - a_j)}(\rho(H) - a_j \operatorname{id})$. Visibly, this is a polynomial in $\rho(H)$ and it maps v to v_i. Doing this for all elements of a basis of $\mathfrak{h}$ we conclude that if we write v as a sum of weight vectors with different weights, then each of these weight vectors can be obtained by applying a combination of actions of elements of $\mathfrak{h}$ to v. The upshot of this is that if $V' \subset V$ is an $\mathfrak{h}$–invariant subspace, $v \in V'$ is any element and $v = v_1 + \dots + v_\ell$ is the decomposition of v into weight vectors of different weights, then each v_i is contained in V'.

Using this, we can now show that the subrepresentation $V' \subset V$ generated by a highest weight vector $v_0 \in V$ of weight λ as above is indecomposable. Suppose that $V' = V_1' \oplus V_2'$ as a $\mathfrak{g}$–representation. Then we can decompose $v_0 = v_1 + v_2$, and we look at the decomposition of v_1 and v_2 into weight vectors of different weights. Of course, one of the two elements must contain a nonzero component of weight λ, so let us assume that v_1 has this property. From above we see that this weight component itself lies in V_1'. But by construction the λ weight space of V' is one–dimensional and thus contained in V_1', which implies $v_0 \in V_1'$ and hence $V' \subset V_1'$, since V' is generated by v_0. Consequently, $V_2' = \{0\}$ which implies that V' is indecomposable. Of course, if V is finite dimensional, the by complete reducibility (Theorem 2.9) V' has to be irreducible.

Assume that V is finite dimensional. Recall that in order to define positive roots we have chosen a total ordering on the space $\mathfrak{h}_0^*$ (in which all the weights of V lie), see 3.7. Since V has only finitely many weights, there is a maximal weight $\lambda_0 \in \operatorname{wt}(V)$, i.e. $\lambda \leq \lambda_0$ for all $\lambda \in \operatorname{wt}(V)$. This is often called the *highest weight* of V, in particular if V is irreducible. If v is any nonzero element of the weight space V_{λ_0}, then v must be a highest weight vector. Indeed, if $E_i \cdot v \neq 0$, then it is a weight vector of weight $\lambda_0 + \alpha_i$, which is strictly bigger than λ_0.

Finally, we claim that the highest weight λ_0 must be *dominant*, i.e. that $\langle \lambda, \alpha_i\rangle \geq 0$ for all simple roots α_i. Indeed, if $\langle \lambda_0, \alpha_j\rangle < 0$ for some simple root α_j, then $s_{\alpha_j}(\lambda_0) = \lambda_0 - \frac{2\langle \lambda_0,\alpha_j\rangle}{\langle \alpha_j,\alpha_j\rangle}\alpha_j$ is strictly bigger than λ_0, but from above we know that this is a weight of V which contradicts maximality of λ_0. In particular we see that the (unique) highest weight of a finite dimensional irreducible representation is a linear combination of the fundamental weights in which all coefficients are non–negative integers. Collecting our information we get

THEOREM. *Let $\mathfrak{g}$ be a finite dimensional complex semisimple Lie algebra, $\mathfrak{h} \leq \mathfrak{g}$ a fixed Cartan subalgebra and Δ the corresponding set of roots. Fix an order on the real span $\mathfrak{h}_0^*$ of Δ, let Δ^+ and Δ^0 be the sets of positive and simple roots with respect to this order, and let $\{E_i, F_i, H_i : i = 1, \dots, n\}$ be a set of standard generators for $\mathfrak{g}$. Let V be a representation of $\mathfrak{g}$ such that any element of $\mathfrak{h}$ acts diagonalizably.*
(1) If $v \in V$ is a highest weight vector of weight λ_0 then the elements of the form $F_{i_1} \cdots F_{i_\ell} \cdot v$ span a subrepresentation of V, and the weights occurring in this subrepresentation have the form $\lambda_0 - \sum n_i\alpha_i$ for $\alpha_i \in \Delta^0$ and non–negative integers n_i.
(2) If V is finite dimensional, then it has at least one highest weight vector, and the weight of any highest weight vector is dominant and algebraically integral. Moreover, the subrepresentation generated by a highest weight vector as in (1) is irreducible. If V itself is irreducible, then up to nonzero complex multiples there is only one highest weight vector in V.

Let us say a few more words on dominant weights: By definition, $\lambda \in \mathfrak{h}_0^*$ is dominant if $\langle \lambda, \alpha\rangle \geq 0$ for all simple roots α. Of course, this is equivalent to the same condition for all positive roots α. The set of all dominant elements of $\mathfrak{h}_0^*$ is called the closed *dominant Weyl chamber*. In general, one defines an open Weyl chamber to be a connected component of the complement of the hyperplanes perpendicular to the roots. Otherwise put, a functional lies in some open Weyl chamber if and only if its inner product with all roots are nonzero, and two such functionals lie in the same open Weyl chamber if and only if all these inner products have the same signs for both functionals. By definition, the open Weyl chambers are disjoint and the intersection of the closures of two Weyl chambers is contained in the union of the hyperplanes perpendicular to the roots.

It is easy to see that the set of open Weyl chambers is in bijective correspondence with the set of positive subsystems of Δ and thus also with the set of simple subsystems of Δ. This is achieved by assigning to an open Weyl chamber the positive subsystem consisting of those roots whose inner products with the elements of the given chamber are positive and then passing to the simple subsystem as usual. Finally, if $\lambda \in \mathfrak{h}_0^*$ is perpendicular to $\alpha \in \Delta$ and w is any element of the Weyl group W, then $w(\lambda)$ is perpendicular to the root $w(\alpha)$. This immediately implies that the action of the Weyl group maps open Weyl chambers to open Weyl chambers. From 3.9 we know that the Weyl group is in bijective correspondence with the set of simple subsystems of Δ and thus we conclude that W is also in bijective correspondence with the set of open Weyl chambers. Otherwise put, for any functional $\lambda \in \mathfrak{h}_0^*$ there is an element $w \in W$ such that $w(\lambda)$ is dominant.

4.3. The theorem of the highest weight. If V is a finite dimensional irreducible representation of a semisimple Lie algebra $\mathfrak{g}$, then from above we know that there is exactly one weight space V_λ which contains a highest weight vector (and the highest weight vectors of V are exactly the nonzero elements of V_λ), and the weight λ is dominant and algebraically integral, i.e. $\frac{2\langle \lambda,\alpha_j\rangle}{\langle \alpha_j,\alpha_j\rangle}$ is a non–negative integer for any simple root α_j.

Next, if V and V' are irreducible representations with the same highest weight λ, then we choose highest weight vectors $v \in V$ and $v' \in V'$. Then (v, v') is a highest weight vector in the finite dimensional representation $V \oplus V'$ of $\mathfrak{g}$, so it generates an irreducible subrepresentation $\tilde{V} \subset V \oplus V'$. The restrictions of the two projections to $\tilde{V}$ define homomorphisms $\tilde{V} \to V$ and $\tilde{V} \to V'$. The homomorphism $\tilde{V} \to V$ is nonzero, since v lies in the image. By irreducibility it must be an isomorphism, since the kernel is an invariant subspace of $\tilde{V}$ and the image is an invariant subspace of V. Similarly, the other homomorphism $\tilde{V} \to V'$ is an isomorphism, so $V \cong V'$. Conversely, isomorphic irreducible representations obviously have the same highest weight.

Thus, to get a complete hand on the finite dimensional irreducible representations (and thus by complete reducibility, see 2.9, on all finite dimensional representations), the remaining question is for which dominant algebraically integral weights λ there exists a finite dimensional irreducible representation with highest weight λ:

THEOREM (Theorem of the highest weight). *If $\mathfrak{g}$ is a finite dimensional complex semisimple Lie algebra, then for any dominant algebraically integral weight $\lambda \in \mathfrak{h}_0^*$ there is a (up to isomorphism) unique finite dimensional irreducible representation with highest weight λ.*

We will next discuss two ways to prove the existence part of this theorem. One of these ways provides a universal construction for all the representations but is technically more involved. The other way is on a case by case basis, but has the advantage that it offers a more concrete description of the representations, which is useful for dealing with examples. Both these approaches need a bit of background from multilinear algebra, which we will discuss next.

Some multilinear algebra

4.4. Tensor products. Let V and W be two vector spaces over some field $\mathbb{K}$. A *tensor product* of V and W is a vector space Z together with a bilinear map $b : V \times W \to Z$ such that for any vector space U and any bilinear map $\varphi : V \times W \to U$ there exists a unique linear map $\tilde{\varphi} : Z \to U$ such that $\varphi = \tilde{\varphi} \circ b$. Hence a tensor product can be thought of as a universal way to "linearize" bilinear mappings.

It follows immediately from the definition by a universal property that if tensor products exist, then they are unique up to canonical isomorphism. Indeed, suppose that (Z, b) and (Z', b') are two tensor products. Applying the universal property of (Z, b) to $b' : V \times W \to Z'$ we obtain a unique linear map $\widetilde{b'} : Z \to Z'$ such that $b' = \widetilde{b'} \circ b$. Exchanging the roles of Z and Z' we obtain a unique linear map $\tilde{b} : Z' \to Z$ such that $b = \tilde{b} \circ b'$. Then $\tilde{b} \circ \widetilde{b'} : Z \to Z$ is a linear map such that $\tilde{b} \circ \widetilde{b'} \circ b = \tilde{b} \circ b' = b$. Since the identity id_Z has the same property, we conclude from the uniqueness part of the universal property that $\tilde{b} \circ \widetilde{b'} = \mathrm{id}_Z$. In the same way $\widetilde{b'} \circ \tilde{b} = \mathrm{id}_{Z'}$, so these are inverse isomorphisms which are compatible with the maps b and b'.

In view of this uniqueness, we can speak of "the tensor product" of V and W (if it exists) and we denote the space by $V \otimes W$ and the bilinear map $V \times W \to V \otimes W$ by $(v, w) \mapsto v \otimes w$. Once we find a way to construct a tensor product, we may then forget about the construction and view the tensor product simply as being defined by the universal property.

There is a (rather boring) general construction of a tensor product as follows: Take the free vector space $\mathcal{F}$ generated by the set $V \times W$, which can be thought of as the set of all formal linear combinations of elements (v, w) with $v \in V$ and $w \in W$. This is endowed with an obvious set map $i : V \times W \to \mathcal{F}$ and it has the universal property that any set map $f : V \times W \to U$ to some vector space U induces a unique linear map $\tilde{f} : \mathcal{F} \to U$ such that $\tilde{f} \circ i = f$. Then consider the linear subspace of $\mathcal{F}$ generated by all elements of the form $(v_1+tv_2, w)-(v_1, w)-t(v_2, w)$ and $(v, w_1+tw_2)-(v, w_1)-t(v, w_2)$. Define $V\otimes W$ as the quotient of $\mathcal{F}$ by this subspace, and $v\otimes w$ as the class of (v, w) in that quotient. Of course, the subspace is chosen precisely in such a way that $(v, w) \mapsto v \otimes w$ becomes bilinear. Given a bilinear map $\varphi : V \times W \to U$, we obtain a unique linear map $\tilde{\varphi} : \mathcal{F} \to U$ which factorizes to the quotient since φ is bilinear. From this it follows immediately that we have really constructed a tensor product. One important thing to note is that if $\{v_i : i \in I\}$ is a basis for V and $\{w_j : j \in J\}$ is a basis of W, then visibly $\{v_i \otimes w_j : (i, j) \in I \times J\}$ is a basis for $V \otimes W$. In particular, if V and W are finite dimensional, then $\dim(V \otimes W) = \dim(V) \dim(W)$.

The tensor product has simple functorial properties. Consider linear maps $\varphi : V \to V'$ and $\psi : W \to W'$, and the map $V \times W \to V' \otimes W'$ defined by $(v, w) \mapsto \varphi(v) \otimes \psi(w)$. This map is obviously bilinear, so by the universal property of the tensor product we get a linear map $\varphi \otimes \psi : V \otimes W \to V' \otimes W'$ which is characterized by $(\varphi \otimes \psi)(v \otimes w) = \varphi(v) \otimes \psi(w)$. One immediately verifies that this construction is compatible with compositions, i.e. $(\varphi_2 \circ \varphi_1) \otimes (\psi_2 \circ \psi_1) = (\varphi_2 \otimes \psi_2) \circ (\varphi_1 \otimes \psi_1)$.

For any vector space U, one may identify bilinear maps $V \times W \to U$ with linear maps $V \to L(W, U)$, so we may rewrite the universal property of the tensor product as having given a natural isomorphism $L(V \otimes W, U) \cong L(V, L(W, U))$ for any space U. Taking $U = \mathbb{K}$, this immediately gives $(V \otimes W)^* \cong L(V, W^*)$.

Next, for any V and W, we have an obvious bilinear map $V^* \times W \to L(V, W)$, which maps (λ, w) to the map $v \mapsto \lambda(v)w$. This induces a linear map $V^* \otimes W \to L(V, W)$. We claim that this map is always injective and it is a linear isomorphism provided that V is finite dimensional. Indeed, any element of $V^* \otimes W$ can be written as a finite sum of elements of the form $\lambda_i \otimes w_i$, and we may assume that the element w_i showing up in this sum are linearly independent. But then we can obviously only get the zero map if $\lambda_i(v) = 0$ for all i and all $v \in V$, so injectivity follows. On the other hand, if V is finite dimensional, then take a basis $\{v_1, \dots, v_n\}$ of V and the dual basis $\{\lambda_1, \dots, \lambda_n\}$ of V^*. By definition, any $v \in V$ can be written as $v = \sum \lambda_i(v)v_i$, so for a linear map $f : V \to W$ we get $f(v) = \sum \lambda_i(v)f(v_i)$, which shows that f lies in the image of our map. Thus $V^* \otimes W \cong L(V, W)$ if V is finite dimensional. Hence $L(V, W^*) \cong V^* \otimes W^*$, so from above we conclude that $(V \otimes W)^* \cong V^* \otimes W^*$ if V is finite dimensional.

Let V be any vector space with dual V^*. Then we have an obvious bilinear map $V^* \times V \to \mathbb{K}$ defined by $(\lambda, v) \mapsto \lambda(v)$. This induces a linear map $V^* \otimes V \to \mathbb{K}$ called the *contraction*. If V is finite dimensional, then $V^* \otimes V \cong L(V, V)$, and one immediately sees that under this isomorphism the contraction corresponds to the trace. In this case, we also have a natural map $\mathbb{K} \to V^* \otimes V$, namely the one corresponding to $t \mapsto t\,\mathrm{id}_V$. This may be realized explicitly by $t \mapsto \sum \lambda_i \otimes v_i$, where $\{v_i\}$ is any basis of V and $\{\lambda_i\}$ is the dual basis of V^*.

From the universal property it follows easily that the tensor product is compatible with direct sums, i.e. $V \otimes (W_1 \oplus W_2)$ is naturally isomorphic to $(V \otimes W_1) \oplus (V \otimes W_2)$. On the other hand, since $L(\mathbb{K}, U) = U$ for any vector space, we immediately conclude that $\mathbb{K} \otimes W \cong W$. Exchanging the two factors defines a natural isomorphism $V \otimes W \cong W \otimes V$.

4.5. Tensor powers, symmetric and exterior powers. Of course, we may form iterated tensor products, and one easily shows that $(U \otimes V) \otimes W$ is naturally isomorphic to $U \otimes (V \otimes W)$, so one may forget about brackets. The tensor product $V_1 \otimes \cdots \otimes V_k$

of k vector spaces is characterized by a universal property for maps $V_1 \times \ldots \times V_k \to U$ which are linear in each entry. In particular, for any space V and any $k \in \mathbb{N}$ we may form the kth tensor power $\otimes^k V$ defined as the tensor product of k copies of V. It is convenient to define $\otimes^0 V := \mathbb{K}$ and $\otimes^1 V := V$.

Next, suppose that $\sigma \in \mathfrak{S}_k$ is a permutation of the set $\{1, \ldots, k\}$. Then for any vector space V we can take the kth tensor power $\otimes^k V$ and consider the map $V^k \to \otimes^k V$ defined by $(v_1, \ldots, v_k) \mapsto v_{\sigma_1} \otimes \cdots \otimes v_{\sigma_k}$. By the universal property of the tensor power we obtain a linear map $f_\sigma : \otimes^k V \to \otimes^k V$, which is characterized by $f_\sigma(v_1 \otimes \cdots \otimes v_k) = v_{\sigma_1} \otimes \cdots \otimes v_{\sigma_k}$.

Recall that to any permutation $\sigma \in \mathfrak{S}_k$ one may associate the sign $\operatorname{sgn}(\sigma) = \pm 1$, which may be defined as the determinant of the linear map $\varphi : \mathbb{R}^k \to \mathbb{R}^k$ defined by $\varphi(e_i) = e_{\sigma_i}$, where $\{e_1, \ldots, e_k\}$ is the standard basis of $\mathbb{R}^k$. Alternatively, one may write $\sigma \in \mathfrak{S}_k$ as a composition of transpositions (i.e. permutations which exchange two elements and leave all others fixed) and define the sign to be 1 even the number of factors is even and -1 if the number of factors is odd. Using this we now define:

Definition. Let V be a vector space over $\mathbb{K}$.
(1) The kth *symmetric power* $S^k V$ of V is defined as the linear subspace formed by all elements $x \in \otimes^k V$ such that $f_\sigma(x) = x$ for all $\sigma \in \mathfrak{S}_k$.
(2) The kth *exterior power* $\Lambda^k V$ of V is defined as the linear subspace formed by all elements $x \in \otimes^k V$ such that $f_\sigma(x) = \operatorname{sgn}(\sigma) x$ for all $\sigma \in \mathfrak{S}_k$.

The fact that $S^k V$ and $\Lambda^k V$ are really linear subspaces of $\otimes^k V$ follows immediately from the fact that they are the common kernels of the maps $f_\sigma - \mathrm{id}$ (respectively $f_\sigma - \operatorname{sgn}(\sigma)\,\mathrm{id}$) for all $\sigma \in \mathfrak{S}_k$.

Next, we can construct natural projections onto the subspaces $S^k V$ and $\Lambda^k V$ of $\otimes^k V$, which in turn allow us to view these two spaces also as quotients of $\otimes^k V$. Define $\mathrm{Symm} = \mathrm{Symm}_k : \otimes^k V \to \otimes^k V$ to be the linear map induced by $(v_1, \ldots, v_k) \mapsto \frac{1}{k!} \sum_{\sigma \in \mathfrak{S}_k} v_{\sigma_1} \otimes \cdots \otimes v_{\sigma_k}$. Otherwise put, $\mathrm{Symm} = \frac{1}{k!} \sum_{\sigma \in \mathfrak{S}_k} f_\sigma$. Obviously, $\mathrm{Symm}(x) = x$ for $x \in S^k V$, while $f_\tau \circ \mathrm{Symm} = \mathrm{Symm}$ for all $\tau \in \mathfrak{S}_k$, so we may view Symm as a projection $\otimes^k V \to S^k V$. On the one hand, this shows that $S^k V \cong V / \ker(\mathrm{Symm})$. On the other hand, for $v_1, \ldots, v_k \in V$, we define $v_1 \vee \cdots \vee v_k := \mathrm{Symm}(v_1 \otimes \cdots \otimes v_k) \in S^k V$.

Now suppose that V is finite dimensional and $\{v_1, \ldots, v_n\}$ is a basis for V. Then the elements $v_{i_1} \otimes \cdots \otimes v_{i_k}$ for $i_1, \ldots, i_k \in \{1, \ldots, n\}$ form a basis for $\otimes^k V$. Since Symm is a projection onto $S^k V$, we conclude that the elements $v_{i_1} \vee \cdots \vee v_{i_k}$ generate $S^k V$. By construction, permuting the elements does not change this product, so we conclude that the elements of that form with $i_1 \leq i_2 \leq \cdots \leq i_k$ still generate $S^k V$. It is also easy to show that the set $\{v_{i_1} \vee \cdots \vee v_{i_k} : i_1 \leq \cdots \leq i_k\}$ is linearly independent, so this forms a basis for $S^k V$.

Another way to enumerate the above basis elements is as $v_1^{i_1} \ldots v_n^{i_n}$ such that $i_1 + \cdots + i_n = k$, where this elements denotes the $\vee$–product of i_1 copies of v_1, i_2 copies of v_2, and so on. One easily verifies that the set of these basis elements is in bijective correspondence with the set of all $n-1$–element subsets of $\{1, \ldots, n+k-1\}$. Explicitly, this bijection associates to the n–tuple $(i_1, \ldots, i_n)$ with $i_1 + \cdots + i_n = k$ the subset $\{i_1 + 1, i_1 + i_2 + 2, \ldots, i_1 + \cdots + i_{n-1} + n - 1\}$. Conversely, to the subset $\{j_1, \ldots, j_{n-1}\}$ with $j_1 < \cdots < j_{n-1}$, one associates the n–tuple

$$(j_1 - 1, j_2 - j_1 - 1, \ldots, j_{n-1} - j_{n-2} - 1, n + k - j_{n-1} - 1).$$

Thus we see that $\dim(V) = n$ implies $\dim(S^k V) = \binom{k+n-1}{n-1}$.

The analogous construction for $\Lambda^k V$ is very similar. One defines $\mathrm{Alt} = \mathrm{Alt}_k : \otimes^k V \to \otimes^k V$ by $\mathrm{Alt}_k = \sum_{\sigma\in\mathfrak{S}_k} \mathrm{sgn}(\sigma) f_\sigma$. Obviously, Alt_k restricts to the identity on $\Lambda^k V$, and one immediately verifies that it has values in $\Lambda^k V$. Similarly as above, we define $v_1\wedge\dots\wedge v_k = \mathrm{Alt}(v_1\otimes\dots\otimes v_k)$ for $v_j \in V$. It is easy to see that $v_1\wedge\dots\wedge v_k = 0$ if two of the v_i are equal. This in turn implies that $v_1\wedge\dots\wedge v_k$ vanishes if the vectors are linearly dependent. Consequently, if V is finite dimensional, then $\Lambda^k V = 0$ for $k > \dim(V)$. Given a basis $\{v_1,\dots,v_n\}$ of V one concludes as above that the elements $v_{i_1}\wedge\dots\wedge v_{i_k}$ with $i_1 < i_2 < \dots < i_k$ form a basis of $\Lambda^k V$. In particular, $\dim(V) = n$ implies $\dim(\Lambda^k V) = \binom{n}{k}$.

For $k = 2$, we have $v_1\otimes v_2 = v_1\vee v_2 + v_1\wedge v_2$, which easily implies $\otimes^2 V = V\otimes V = S^2V\oplus\Lambda^2 V$. For $k > 2$, we still have $S^kV\oplus\Lambda^kV \subset \otimes^k V$, but this is a proper subspace. Describing further subspaces in $\otimes^k V$ is an important ingredient for understanding representations of $GL(V)$ and $SL(V)$.

The symmetric and exterior powers can also be characterized by universal properties. Suppose that $\varphi : V^k \to W$ is a k–linear map which is symmetric, i.e. has the property that $\varphi(v_{\sigma_1},\dots,v_{\sigma_k}) = \varphi(v_1,\dots,v_k)$ for all $v_j\in V$ and all $\sigma\in\mathfrak{S}_k$. By the universal property of the tensor power, we obtain a linear map $\otimes^k V \to W$. This map visibly vanishes on the kernel of Symm_k, so it factorizes to a unique linear map $\tilde{\varphi} : S^kV \to W$, which is characterized by $\tilde{\varphi}(v_1\vee\dots\vee v_k) = \varphi(v_1,\dots,v_k)$. Similarly, if φ is alternating, i.e. $\varphi(v_{\sigma_1},\dots,v_{\sigma_k}) = \mathrm{sgn}(\sigma)\varphi(v_1,\dots,v_k)$ for all $v_j \in V$ and all $\sigma\in\mathfrak{S}_k$, then we get a unique linear map $\tilde{\varphi} : \Lambda^kV \to W$ such that $\tilde{\varphi}(v_1\wedge\dots\wedge v_k) = \varphi(v_1,\dots,v_k)$.

In particular, given a linear map $f : V \to W$, we may use that k–linear symmetric map $(v_1,\dots,v_k)\mapsto f(v_1)\vee\dots\vee f(v_k)$, to obtain a linear map $S^kf : S^kV \to S^kW$ characterized by $S^kf(v_1\vee\dots\vee v_k) = f(v_1)\vee\dots\vee f(v_k)$. Clearly, $S^k(g\circ f) = S^kg\circ S^kf$, so this has functorial properties. In the same way, we obtain $\Lambda^kf : \Lambda^kV\to\Lambda^kW$, which also is functorial. An interesting special case of the latter construction is to consider an endomorphism $f : V \to V$ of an n–dimensional vector space V, and the induced map $\Lambda^nf : \Lambda^nV \to \Lambda^nV$. From above, we know that Λ^nV is one–dimensional, so Λ^nf is given by multiplication by an element of $\mathbb{K}$, which is easily seen to equal the determinant of f. More general, the expressions $\mathrm{tr}(\Lambda^kf)$ turns out to be up to a sign the coefficient of t^{n-k} in the characteristic polynomial of f.

Finally, we note that forming symmetric and exterior powers is nicely compatible with dualities, i.e. for a vector space V with dual V^* we have natural isomorphisms $S^kV^* \cong (S^kV)^*$ and $\Lambda^kV^*\cong(\Lambda^kV)^*$ for all k.

4.6. Tensor products and powers of representations. From the functorial properties of tensor products and symmetric and exterior powers it follows that these constructions can be carried out for representations. Indeed, suppose that we have given representations of a group G on two vector spaces V and W. Then we define a representation of G on $V\otimes W$ by $g\cdot(v\otimes w) = (g\cdot v)\otimes(g\cdot w)$. Otherwise put, if we denote the representations by $\rho : G\to GL(V)$ and $\rho' : G \to GL(W)$, then the action of g on $V\otimes W$ is given by $\rho(g)\otimes\rho'(g)$. Therefore, the tensor product representation is often denoted by $\rho\otimes\rho'$. That this is indeed a representation follows immediately from the functorial properties of the tensor product.

If V and W are finite dimensional, then we can choose bases of the two spaces and use them to construct a basis of $V\otimes W$. For linear maps $f : V\to V$ and $g : W\to W$ we see immediately that the matrix coefficients of $f\otimes g : V\otimes W\to V\otimes W$ depend smoothly (even polynomially) on the matrix coefficients of f and g. Hence if we start with a Lie group G and smooth homomorphisms $\rho : G\to GL(V)$ and $\rho' : G\to GL(W)$, then the homomorphism $\rho\otimes\rho' : G\to GL(V\otimes W)$ is smooth, too. Hence we can also form tensor products of (finite dimensional) representations of a Lie group.

To pass to the Lie algebra, one has to differentiate homomorphisms and concludes that given representations of a Lie algebra $\mathfrak{g}$ on V and W, then the natural action on $V\otimes W$ is given by $X\cdot(v\otimes w)=(X\cdot v)\otimes w+v\otimes X\cdot w$. Alternatively, one may use this as the definition of the tensor product of two representations and verify directly that it indeed defines a representation. This follows immediately from the fact that

$$X\cdot(Y\cdot(v\otimes w))=(X\cdot(Y\cdot v))\otimes w+(Y\cdot v)\otimes(X\cdot w)+(X\cdot v)\otimes(Y\cdot w)+v\otimes(X\cdot(Y\cdot w)),$$

and the middle two terms cancel with the corresponding terms in $-Y\cdot(X\cdot(v\otimes w))$.

If we have given a third representation U, then the isomorphism $L(V\otimes W,U)\cong L(V,L(W,U))$ from 4.4 is actually an isomorphism of representations. Let us verify this for the Lie algebra representations: For $\varphi:V\otimes W\to U$ and $X\in\mathfrak{g}$, we by definition have $(X\cdot\varphi)(v\otimes w)=X\cdot(\varphi(v\otimes w))-\varphi(X\cdot(v\otimes w))$, and the second summand gives $-\varphi((X\cdot v)\otimes w)-\varphi(v\otimes(X\cdot w))$. The linear map $\Phi:V\to L(W,U)$ is given by $\Phi(v)(w)=\varphi(v\otimes w)$. The action there is given by $(X\cdot\Phi)(v)=X\cdot(\Phi(v))-\Phi(X\cdot v)$. Moreover, $(X\cdot(\Phi(v)))(w)=X\cdot(\Phi(v)(w))-\Phi(v)(X\cdot w)$, which proves our claim. As in 4.4 we can put $U=\mathbb{K}$ with the trivial representation to conclude that $(V\otimes W)^*\cong L(V,W^*)$ as representations of $\mathfrak{g}$.

Similarly, the map $V^*\otimes W\to L(V,W)$ constructed in 4.4 is a homomorphism of representations. For $X\in\mathfrak{g}$ we have $X\cdot(\lambda\otimes w)=(X\cdot\lambda)\otimes w+\lambda\otimes(X\cdot w)$. Since by definition $(X\cdot\lambda)(v)=-\lambda(X\cdot v)$ we see that $X\cdot(\lambda\otimes w)$ corresponds to the map $v\mapsto-\lambda(X\cdot v)w+\lambda(v)(X\cdot w)$. Again by definition, this equals the action of X on the linear map $v\mapsto\lambda(v)w$. Thus we see that if V is finite dimensional, then $V^*\otimes W\cong L(V,W)$ as representations of $\mathfrak{g}$. Together with the above, we see that for finite dimensional V we have $(V\otimes W)^*\cong V^*\otimes W^*$ as representations.

Of course, we can also form iterated tensor products of representations. In particular, given a representation of $\mathfrak{g}$ on V, we obtain a representation on $\otimes^k V$, which is given by

$$X\cdot(v_1\otimes\dots\otimes v_k)=\sum_{i=1}^k v_1\otimes\dots\otimes X\cdot v_i\otimes\dots\otimes v_k.$$

From this it is obvious that the action of X commutes with the map f_σ for any $\sigma\in\mathfrak{S}_k$, and thus in particular with the maps Symm_k and Alt_k. Hence we conclude that for $a\in S^kV$ and $X\in\mathfrak{g}$ we have $X\cdot a=X\cdot\mathrm{Symm}_k(a)=\mathrm{Symm}_k(X\cdot a)\in S^kV$, and similarly for Λ^kV. Thus we conclude that S^kV and Λ^kV are invariant subspaces of $\otimes^kV$, so we can also form symmetric and exterior powers of representations. Again, the isomorphisms $S^kV^*\cong(S^kV^*)$ and $\Lambda^kV^*\cong(\Lambda^kV)^*$ are isomorphisms of representations.

Existence of irreducible representations

4.7. Fundamental representations. Having the construction of tensor products of representations at hand, we can now describe the case by case approach to the proof of existence of finite dimensional irreducible representations. This is based on the following simple observation: Suppose that V and W are irreducible representations with highest weights λ and μ respectively. Take highest weight vectors $v_0\in V$ and $w_0\in W$, and consider $v_0\otimes w_0\in V\otimes W$. For any element X of a positive root space $\mathfrak{g}_\alpha$ we by definition have $X\cdot v_0=0$ and $X\cdot w_0=0$, and thus $X\cdot(v_0\otimes w_0)=0$, so this is again a highest weight vector. For H in the Cartan subalgebra $\mathfrak{h}$ we have by definition $H\cdot(v_0\otimes w_0)=(\lambda+\mu)(H)v_0\otimes w_0$, so this highest weight vector has weight $\lambda+\mu$. Moreover, since v_0 and w_0 are the unique weight vectors of weight λ and μ up to scale, we immediately conclude that the weight space $(V\otimes W)_{\lambda+\mu}$ has dimension one. In view of Theorem 4.2 this implies that $V\otimes W$ contains a unique irreducible subrepresentation with highest weight $\lambda+\mu$.

Similarly, the vector $v_0\vee\dots\vee v_0\in S^kV$ is a highest weight vector of weight $k\lambda$ for each $k\in\mathbb{N}$, which is unique up to scale. Together with the above, this implies that $S^kV\otimes S^\ell W$ contains a unique irreducible subrepresentation of highest weight $k\lambda+\ell\mu$.

Now recall from 4.1 that we have the fundamental weights $\omega_1,\dots,\omega_n$ for $\mathfrak{g}$ (where n is the rank of $\mathfrak{g}$). Any dominant integral weight λ for $\mathfrak{g}$ can be written as $a_1\omega_1+\dots+a_n\omega_n$, where the a_i are non–negative integers. Suppose that for each $i=1,\dots,n$ we are able to construct an irreducible representation V_i of $\mathfrak{g}$ with highest weight ω_i. Then this is called the ith *fundamental representation* of $\mathfrak{g}$, or the fundamental representation corresponding to the simple root α_i. In view of the above, the representation $S^{a_1}V_1\otimes\dots\otimes S^{a_n}V_n$ then contains a unique irreducible subrepresentation of highest weight $\lambda=a_1\omega_1+\dots+a_n\omega_n$. Thus constructing the fundamental representations leads to a proof of existence of finite dimensional irreducible representations with arbitrary dominant integral highest weight.

Let us carry this out in the case $\mathfrak{g}=\mathfrak{sl}(n,\mathbb{C})$. Recall that in this case we have chosen the Cartan subalgebra $\mathfrak{h}$ to be the space of tracefree diagonal $n\times n$–matrices. Thus $\mathfrak{g}$ has rank $n-1$ and therefore we have to construct $n-1$ fundamental representations. To do this, we first have to determine the fundamental weights. From 3.6 we know that $\langle e_i,\alpha_{i-1}\rangle=-1$, $\langle e_i,\alpha_i\rangle=1$ and $\langle e_i,\alpha_j\rangle=0$ for $j\neq i-1,i$. This immediately implies that the fundamental weights are given by $\omega_1=e_1$, $\omega_2=e_1+e_2$, $\dots$, $\omega_{n-1}=e_1+\dots+e_{n-1}$.

Now consider the standard representation of $\mathfrak{g}$ on $\mathbb{C}^n$ and let $\{v_1,\dots,v_n\}$ be the standard basis of $\mathbb{C}^n$. A diagonal matrix acts on v_i by multiplication with the ith diagonal entry, so we see that each v_i is a weight vector of weight e_i. In the ordering introduced in 3.8, we have $e_1>\dots>e_n$, so the highest weight of $\mathbb{C}^n$ is e_1. Since $\mathbb{C}^n$ is obviously irreducible, we have found the first fundamental representation.

Getting the other fundamental representations is not difficult either: For $2\leq k\leq n-1$ the exterior power $\Lambda^k\mathbb{C}^n$ contains the vector $v_1\wedge\dots\wedge v_k$, which is a weight vector of weight $e_1+\dots+e_k$. The positive root spaces of $\mathfrak{g}$ are generated by the elementary matrices E_{ij} with $i<j$, so all these matrices are strictly upper triangular. This means that they map each v_i to a linear combination of the v_j for $j<i$. This immediately implies that $v_1\wedge\dots\wedge v_k$ is a highest weight vector, so $\Lambda^k\mathbb{C}^n$ contains an irreducible subrepresentation isomorphic to the kth fundamental representation. It is however easy to show that the subrepresentation of $\Lambda^k\mathbb{C}^n$ generated by the highest weight vector is all of $\Lambda^k\mathbb{C}^n$, so this representation is irreducible and it gives the required fundamental representation: Indeed, acting on $v_1\wedge\dots\wedge v_k$ with an element E_{i_kk} with $i_k>k$, we obtain $v_1\wedge\dots\wedge v_{k-1}\wedge v_{i_k}$. Acting on this with $E_{i_{k-1},k-1}$ for $k-1<i_{k-1}<i_k$ we obtain $v_1\wedge\dots\wedge v_{k-2}\wedge v_{i_{k-1}}\wedge v_{i_k}$. Iterating this, we see that we obtain all elements of the standard basis $\{v_{i_1}\wedge\dots\wedge v_{i_k}:i_1<\dots<i_k\}$ of $\Lambda^k\mathbb{C}^n$.

Thus we have proved existence of finite dimensional irreducible representations with arbitrary dominant integral highest weights for $\mathfrak{g}=\mathfrak{sl}(n,\mathbb{C})$. We moreover know that the irreducible representation of highest weight $a_1\omega_1+\dots+a_{n-1}\omega_{n-1}$ can be realized as a subrepresentation in $S^{a_1}\mathbb{C}^n\otimes\dots\otimes S^{a_{n-1}}(\Lambda^{n-1}\mathbb{C}^n)$, and thus also as a subrepresentation in $\otimes^\ell\mathbb{C}^n$, where $\ell=a_1+2a_2+\dots+(n-1)a_{n-1}$. We will return to a more explicit description of these representations later on.

Of course, we can also consider the dual $\mathbb{C}^{n*}$ of the standard representation and its exterior powers $\Lambda^k\mathbb{C}^{n*}$. Consider the basis of $\mathbb{C}^{n*}$ which is dual to the standard basis of

$\mathbb{C}^n$. From the definition of the dual action of a Lie algebra it follows immediately that this is a basis consisting of weight vectors of weights $-e_1, \dots, -e_n$, so the highest weight is $-e_n$. Recall however, that as functionals on $\mathfrak{h}$, we have the relation $e_1 + \dots + e_n = 0$, so the highest weight is actually given by $e_1 + \dots + e_{n-1}$. Hence we conclude that $\mathbb{C}^{n*} \cong \Lambda^{n-1}\mathbb{C}^n$. Similarly one sees that $\Lambda^k\mathbb{C}^{n*}$ has the same highest weight as $\Lambda^{n-k}\mathbb{C}^n$, so these two representations are isomorphic.

These isomorphisms are actually easy to see directly: Since $\Lambda^n\mathbb{C}^n$ is a one–dimensional representation of the Lie algebra $\mathfrak{g}$, it must be trivial. (Alternatively, one may observe that the Lie algebra action of a matrix $A \in M_n(\mathbb{C})$ on $\Lambda^n\mathbb{C}^n$ is given by multiplication with $\operatorname{tr}(A)$.) Now for $k = 1, \dots, n-1$ one defines a bilinear map $\Lambda^k\mathbb{C}^n \times \Lambda^{n-k}\mathbb{C}^n \to \Lambda^n\mathbb{C}^n$ by $(a_1 \wedge \dots \wedge a_k, b_1 \wedge \dots \wedge b_{n-k}) \mapsto a_1 \wedge \dots \wedge a_k \wedge b_1 \wedge \dots \wedge b_{n-k}$ for all $a_i, b_j \in V$. This is easily seen to be well defined, non–degenerate and $\mathfrak{g}$–equivariant, and thus identifies $\Lambda^{n-k}\mathbb{C}^n$ with the dual of $\Lambda^k\mathbb{C}^n$ as a representation of $\mathfrak{g}$.

Let us proceed a bit further in the discussion of representations of $\mathfrak{g}$. For any $k \geq 2$ we can form the kth symmetric power $S^k\mathbb{C}^n$ of the standard representation. Of course, $v_1 \vee \dots \vee v_1 = v_1^k$ is a highest weight vector of weight ke_1 in this representation. Hence $S^k\mathbb{C}^n$ contains an irreducible component of highest weight ke_1. Similar arguments as for the exterior powers show that $S^k\mathbb{C}^n$ is irreducible. By definition, the action of $E_{i_k 1}$ on this element gives $k(v_1^{k-1} \vee v_{i_k})$. Acting once more with the same element we obtain $k(k-1)v_1^{k-1} \vee v_{i_k}^2$. Iterating this, we see that we again obtain all elements of the standard basis from v_1^k.

In particular, in the case of $\mathfrak{sl}(2, \mathbb{C})$ the representation $S^k\mathbb{C}^2$ is the irreducible representation of dimension $k+1$, so this gives a non–computational way to explicitly describe all the irreducible representations of $\mathfrak{sl}(2, \mathbb{C})$.

Another simple consequence of this is that $\otimes^2\mathbb{C}^n = S^2\mathbb{C}^n \oplus \Lambda^2\mathbb{C}^n$ is the decomposition into irreducible representations. Suppose that V is an irreducible representation with dual V^*. Then also V^* is irreducible, since for an invariant subspace $W \subset V^*$ the subspace $\{v \in V : \varphi(v) = 0 \quad \forall \varphi \in W\}$ is invariant in V. Thus we see that $\otimes^2\mathbb{C}^{n*} = S^2\mathbb{C}^{n*} \oplus \Lambda^2\mathbb{C}^{n*}$ is the decomposition into irreducibles.

We can use this, to describe a very simple but typical application of representation theory: The elements of $\otimes^2\mathbb{C}^{n*}$, $S^2\mathbb{C}^{n*}$ and $\Lambda^2\mathbb{C}^{n*}$ are by definition bilinear, symmetric bilinear, and skew symmetric bilinear maps $\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$, respectively. Given a bilinear map $b : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$, we of course can form the symmetric part $b_s := \frac{1}{2}(b + \tilde{b})$ and the alternating part $b_a := \frac{1}{2}(b - \tilde{b})$, where $\tilde{b}(x, y) := b(y, x)$, and $b = b_s + b_a$. Now suppose that $\Phi : \otimes^2\mathbb{C}^{n*} \to \otimes^2\mathbb{C}^{n*}$ is an $\mathfrak{sl}(n, \mathbb{C})$–equivariant map. Since $S^2\mathbb{C}^{n*}$ and $\Lambda^2\mathbb{C}^{n*}$ are non–isomorphic irreducible representations of $\mathfrak{g}$, we conclude that Φ has to preserve the decomposition $\otimes^2\mathbb{C}^{n*} = S^2\mathbb{C}^{n*} \oplus \Lambda^2\mathbb{C}^{n*}$. Moreover, by Schur's lemma from 2.4, the restriction of Φ to each of the irreducible components must be a multiple of the identity. Hence we conclude that there are unique numbers $z, w \in \mathbb{C}$ such that $\Phi(b) = \Phi(b_s + b_a) = zb_s + wb_a$. Inserting the definitions, we obtain $\Phi(b) = \frac{z+w}{2}b + \frac{z-w}{2}\tilde{b}$. Hence we conclude that the space of $\mathfrak{g}$–homomorphisms Φ as above is two dimensional, with the identity and $b \mapsto \tilde{b}$ as a basis. Otherwise put, the only basis–independent linear maps on bilinear forms are given by linear combinations of the identity and exchanging the arguments.

4.8. Fundamental representations for the other classical simple Lie algebras. For the other classical simple Lie algebras one may proceed similarly, and we just sketch briefly how this looks. Lots of details about this approach can be found in [**Fulton-Harris**].

Let us first discuss the symplectic Lie algebra $\mathfrak{g} = \mathfrak{sp}(2n, \mathbb{C})$ from part (4) of 3.8. We have the basic functionals e_i (which this time are linearly independent) and in terms of these we have $\Delta = \{\pm e_i \pm e_j\} \cup \{\pm 2e_i\}$. The simple roots are $\alpha_j = e_j - e_{j+1}$ for $j < n$ and $\alpha_n = 2e_n$. Up to a multiple, the e_i are orthonormal for the inner product $\langle\ ,\ \rangle$. As in the case of $\mathfrak{sl}(n, \mathbb{C})$, this implies that the fundamental weight ω_j equals $e_1 + \dots + e_j$ for $j = 1, \dots, n-1$. Finally, $e_1 + \dots + e_n$ has trivial inner product with the α_j for $j < n$, while with α_n we get inner product 2, which is exactly half of $\langle \alpha_n, \alpha_n \rangle$, so $\omega_n = e_1 + \dots + e_n$.

As before, the standard basis $\{v_1, \dots, v_{2n}\}$ of $\mathbb{C}^{2n}$ consists of weight vectors, and the weight of v_i is e_i for $i \leq n$ and $-e_{i-n}$ for $i > n$. Hence the highest weight vector of the standard representation is v_1, and the standard representation is the first fundamental representation. From the description in 3.8 one sees immediately that elements of positive root spaces map each v_i with $i \leq n$ to a linear combination of v_j with $j < i$. This implies as above that for $k = 2, \dots, n$ the element $v_1 \wedge \dots \wedge v_k$ is a highest weight vector in $\Lambda^k \mathbb{C}^{2n}$ of weight $e_1 + \dots + e_k$. Hence the kth fundamental representation is contained as an irreducible subrepresentation in $\Lambda^k \mathbb{C}^{2n}$ for all $k = 1, \dots, n$.

The main difference to the $\mathfrak{sl}(n, \mathbb{C})$ case is that the exterior powers of the standard representation are not irreducible any more. Indeed, this is clear already for $\Lambda^2 \mathbb{C}^{2n}$: By definition of $\mathfrak{g}$, there is a skew symmetric $\mathfrak{g}$–invariant bilinear form b on $\mathbb{C}^{2n}$. Of course, this corresponds to a $\mathfrak{g}$–homomorphism $b : \Lambda^2 \mathbb{C}^{2n} \to \mathbb{C}$, whose kernel we denote by $\Lambda^2_0 \mathbb{C}^{2n}$. More generally, for $2 < k \leq n$ we can consider the map $(\mathbb{C}^{2n})^k \to \Lambda^{k-2} \mathbb{C}^n$ defined by

$$(a_1, \dots, a_k) \mapsto \sum_{\sigma \in \mathfrak{S}_k} \operatorname{sgn}(\sigma) b(a_{\sigma_1}, a_{\sigma_2}) a_{\sigma_3} \wedge \dots \wedge a_{\sigma_k}.$$

This is immediately seen to be skew symmetric and $\mathfrak{g}$–equivariant, so it induces a $\mathfrak{g}$–homomorphism $\Lambda^k \mathbb{C}^{2n} \to \Lambda^{k-2} \mathbb{C}^{2n}$. The kernel $\Lambda^k_0 \mathbb{C}^{2n}$ is a subrepresentation which visibly contains the highest weight vector $v_1 \wedge \dots \wedge v_k$. With a bit more effort than in the $\mathfrak{sl}(n, \mathbb{C})$–case one shows that $\Lambda^k_0 \mathbb{C}^{2n}$ is irreducible for $2 \leq k \leq n$ and thus the kth fundamental representation of $\mathfrak{g}$.

Let us note two more facts on the symplectic Lie algebras. On the one hand, the $\mathfrak{g}$–invariant non–degenerate bilinear form b on $\mathbb{C}^{2n}$ induces an isomorphism $\mathbb{C}^{2n} \cong \mathbb{C}^{2n*}$ of representations of $\mathfrak{g}$. Hence for any k we obtain $\Lambda^k \mathbb{C}^{2n} \cong \Lambda^k \mathbb{C}^{2n*}$ as $\mathfrak{g}$–representations. Since $\mathfrak{g} \subset \mathfrak{sl}(2n, \mathbb{C})$, we also get an isomorphism $\Lambda^k \mathbb{C}^{2n} \cong \Lambda^{2n-k} \mathbb{C}^{2n*}$ and the latter representation is isomorphic to $\Lambda^{2n-k} \mathbb{C}^{2n}$. Hence, we also get a description of the remaining exterior powers. On the other hand, the space $S^2 \mathbb{C}^{2n}$ contains the vector $v_1 \vee v_1$, which is a highest weight vector of weight $2e_1$, which coincides with the highest root of $\mathfrak{g}$. Since $\mathfrak{g}$ is simple, the adjoint representation is irreducible. Consequently, $S^2 \mathbb{C}^{2n}$ contains the adjoint representation $\mathfrak{g}$ as an irreducible subrepresentation. From the explicit description of $\mathfrak{g}$, we see that $\dim(\mathfrak{g}) = n^2 + n(n+1) = 2n^2 + n$, which coincides with the dimension $\binom{2n+1}{2n-1}$ of $S^2 \mathbb{C}^{2n}$. Thus we conclude that $S^2 \mathbb{C}^{2n}$ is irreducible and isomorphic to the adjoint representation. This isomorphism is given explicitly by mapping $w_1 \vee w_2$ to the map $v \mapsto b(v, w_1) w_2 + b(v, w_2)$.

For the orthogonal algebras, a new phenomenon arises. Let us start with the odd case, i.e. $\mathfrak{g} = \mathfrak{so}(2n+1, \mathbb{C})$ as treated in part (3) of 3.8. As before, we have the basic functionals e_i for $i = 1, \dots, n$, which are orthonormal with respect to (a multiple of) the Killing form. The set of roots is given by $\Delta = \{\pm e_i \pm e_j\} \cup \{\pm e_i\}$, and the simple roots are $\alpha_j = e_j - e_{j+1}$ for $1 \leq j < n$ and $\alpha_n = e_n$. Hence one may compute the fundamental weights as in the case of $\mathfrak{sp}(2n, \mathbb{C})$ and $\omega_j = e_1 + \dots + e_j$ for $j < n$. However, this time the last fundamental weight ω_n is given by $\frac{1}{2}(e_1+\dots+e_n)$. The weights of the standard representation $\mathbb{C}^{2n+1}$ in this case are $\pm e_i$ for $i=1,\dots,n$ and 0.

As before, this shows that for $k=1,\dots,n-1$ the exterior power $\Lambda^k\mathbb{C}^{2n+1}$ contains the kth fundamental representation. One proves that these exterior powers are themselves irreducible, and thus give all but the last fundamental representations. In particular, the representation $\Lambda^2\mathbb{C}^{2n+1}$ has highest weight e_1+e_2, which coincides with the highest root of $\mathfrak{g}$. Since both representations are irreducible (or by observing that $\mathfrak{g}$ is irreducible and has the same dimension as $\Lambda^2\mathbb{C}^{2n+1}$), we conclude that $\Lambda^2\mathbb{C}^{2n+1}$ is isomorphic to the adjoint representation. Explicitly, this isomorphism is given by assigning to $w_1\wedge w_2$ the map $v\mapsto b(v,w_1)w_2-b(v,w_2)w_1$.

The last fundamental representation however is more mysterious. The point here is that visibly the fundamental weight ω_n cannot be a weight of any tensor power of the standard representation. Indeed, it turns out that this last fundamental representation does not come from a representation of the group $SO(2n+1,\mathbb{C})$. The point here is that this group is not simply connected, so there are Lie algebra homomorphism on $\mathfrak{g}$ which do not integrate to group homomorphisms on $SO(2n+1,\mathbb{C})$.

The fundamental representation ω_n of $\mathfrak{g}$ is called the *spin–representation* S. Correspondingly, the simply connected group with Lie algebra $\mathfrak{g}$ is called the spin group $\mathrm{Spin}(2n+1,\mathbb{C})$. It turns out that this spin group is an extension of $SO(2n+1,\mathbb{C})$ with kernel $\mathbb{Z}_2=\{\pm1\}$, i.e. there is a surjective homomorphism $\mathrm{Spin}(2n+1,\mathbb{C})\to SO(2n+1,\mathbb{C})$ whose kernel consists of two elements. One can construct both the spin representation S and the spin group $\mathrm{Spin}(2n+1,\mathbb{C})$ using the *Clifford algebra* $Cl(2n+1,\mathbb{C})$ of $\mathbb{C}^{2n+1}$. This is an associative algebra of dimension 2^{2n+1}, which is canonically associated to the non–degenerate complex bilinear form on $\mathbb{C}^{2n+1}$. One shows that $\mathfrak{g}$ may be realized as a subspace in $Cl(2n+1,\mathbb{C})$ with the bracket given by the commutator in the algebra, while the spin group may be realized as a subgroup of the group of invertible elements in $Cl(2n+1,\mathbb{C})$. Both objects turn out to be contained in a subalgebra which has a natural representation S with $\dim(S)=2^n$, which restricts to the last fundamental representation of $\mathfrak{g}$ as well as to a representation of the spin group.

What we can read off from the highest weights immediately is that the tensor product $S\otimes S$ contains a unique irreducible subrepresentation with highest weight $e_1+\dots+e_n$ and this is isomorphic to $\Lambda^n\mathbb{C}^{2n+1}$. More precisely, one can show that the decomposition of $S\otimes S$ into irreducibles is given by $S\otimes S=\oplus_{k=0}^n\Lambda^k\mathbb{C}^{2n+1}$. Details about the constructions of the spin groups and spin representations can be found in [**Fulton-Harris**, chapter 20]. It should also be mentioned that the spin representations play an important role in theoretical physics and in differential geometry.

Let us finally turn to the even orthogonal algebra $\mathfrak{g}=\mathfrak{so}(2n,\mathbb{C})$ with $n\geq 4$. From example (2) of 3.8 we again have the basic functional e_i which are orthonormal, and $\Delta=\{\pm e_i\pm e_j : i\neq j\}$. The simple roots $\alpha_1,\dots,\alpha_n$ are given by $\alpha_j=e_j-e_{j+1}$ for $j<n$ and $\alpha_n=e_{n-1}+e_n$. As before, we see that the first fundamental weights are given by $\omega_j=e_1+\dots+e_j$ for $j\leq n-2$, while $\omega_{n-1}=\frac{1}{2}(e_1+\dots+e_{n-1}-e_n)$ and $\omega_n=\frac{1}{2}(e_1+\dots+e_n)$. The weights of the standard representation $\mathbb{C}^{2n}$ are $\pm e_i$. Similarly as before we conclude that the first $n-2$ fundamental representations are given by the exterior powers $\Lambda^k\mathbb{C}^{2n}$ for $k=1,\dots,n-2$, and these again turn out to be irreducible. The last two fundamental representations are realized by two non–isomorphic spin representations S_+ and S_- which both turn out to have dimension 2^{n-1}. As in the odd case, these representations and the spin group (which is again the universal covering of $SO(2n,\mathbb{C})$) can be constructed using Clifford algebras.

As before, we have $\mathbb{C}^{2n*} \cong \mathbb{C}^{2n}$ as a $\mathfrak{g}$–representation and $\Lambda^k\mathbb{C}^{2n*} \cong \Lambda^{2n-k}\mathbb{C}^{2n}$, so it remains to understand the exterior powers $\Lambda^{n-1}\mathbb{C}^{2n}$ and $\Lambda^n\mathbb{C}^{2n}$. The representation $\Lambda^{n-1}\mathbb{C}^{2n}$ turns out to be irreducible, and of course its highest weight is $e_1+\dots+e_{n-1} = \omega_{n-1}+\omega_n$. On the other hand, $\Lambda^n\mathbb{C}^{2n}$ is not irreducible, but a direct sum of two non–isomorphic irreducible representations $\Lambda^n_+\mathbb{C}^{2n}$ and $\Lambda^n_-\mathbb{C}^{2n}$, called the *self–dual* and the *anti–self–dual* part. This can be seen as follows: Looking at the explicit description of $\mathfrak{g}$ in 3.8 we see that any element in a positive root space has the property that it maps each element v_i of the standard basis of $\mathbb{C}^{2n}$ with $i \leq n$ to a linear combination of elements v_j with $j < i$. Therefore, $v_1 \wedge \dots \wedge v_n \in \Lambda^n\mathbb{C}^{2n}$ is a highest weight vector of weight $e_1+\dots+e_n = 2\omega_n$. But such an element also has the property that it maps v_{2n} to a linear combination of $v_1,\dots,v_{n-1}$. Thus we conclude that also $v_1\wedge\dots\wedge v_{n-1}\wedge v_{2n}$ is a highest weight vector, and its weight is $e_1+\dots+e_{n-1}-e_n = 2\omega_{n-1}$. Hence we have found two irreducible subrepresentations of $\Lambda^n\mathbb{C}^{2n}$, which turn out to be irreducible and each have half the dimension of $\Lambda^n\mathbb{C}^{2n}$. From the highest weights we can read off that $S_+\otimes S_-$ contains an irreducible subrepresentation isomorphic to $\Lambda^{n-1}\mathbb{C}^{2n}$, while $\Lambda^n_\pm\mathbb{C}^{2n}$ are isomorphic to irreducible subrepresentations of S^2S_+ respectively S^2S_-.

The counterpart of the fact that the exterior powers of the standard representation are not irreducible as representations of the symplectic algebra is that symmetric powers of the standard representation are not irreducible as representations of the orthogonal algebras. The situation is the same for the even and odd orthogonal algebras, so let us look at the standard representation $\mathbb{C}^n$ of $\mathfrak{g} = \mathfrak{so}(n,\mathbb{C})$. Of course, the $\mathfrak{g}$–invariant symmetric bilinear form b induces a homomorphism $S^2\mathbb{C}^n \to \mathbb{C}$, whose kernel we denote by $S^2_0\mathbb{C}^n$. This is called the tracefree part of $S^2\mathbb{C}^n$. More generally, for each k we may consider the map $(\mathbb{C}^n)^k \to S^{k-2}\mathbb{C}^n$ defined by

$$(v_1,\dots,v_k) \mapsto \sum_{\sigma\in\mathfrak{S}_k} b(v_{\sigma_1},v_{\sigma_2})v_{\sigma_3}\vee\dots\vee v_{\sigma_k}.$$

Of course, this induces a linear map $S^k\mathbb{C}^n \to S^{k-2}\mathbb{C}^n$, which is easily seen to be a surjective $\mathfrak{g}$–homomorphism. The kernel of this map is a subrepresentation $S^k_0\mathbb{C}^n \subset S^k\mathbb{C}^n$ which is called the *totally tracefree part* of $S^k\mathbb{C}^n$. It turns out that these representations are all irreducible, so they are the irreducible representations with highest weight ke_1.

It is also possible to describe the fundamental representations of the exceptional simple Lie algebras, but for the larger ones, this becomes rather difficult. Indeed, the fundamental representation of lowest dimension of E_8 turns out to be the adjoint representation, which has dimension 248. The largest among the dimensions of the fundamental representations of E_8 happens to be 6899079264.

The universal enveloping algebra and Verma modules

To conclude this chapter, we briefly discuss an approach which allows to prove the existence of irreducible representations with any dominant integral weight for arbitrary complex semisimple Lie algebras. This is based on the universal enveloping algebra of a Lie algebra, which is an important tool in deeper aspects of representation theory.

4.9. The universal enveloping algebra. Recall that a *associative algebra* over a field $\mathbb{K}$ is a $\mathbb{K}$–vector space A endowed with a bilinear associative multiplication $\mu : A\times A \to A$. The algebra (A,μ) is called *unital* if there is a unit element $1 \in A$ for μ. We will usually simply write ab for $\mu(a,b)$. A homomorphism of unital associative algebras is simply a linear map which is compatible with the multiplication and maps the unit element to the unit element. A representation of a unital associative algebra A on a vector space V is a homomorphism $A \to L(V,V)$ of unital associative algebras.

Given an associative algebra A, we define the commutator $[a,b] := ab - ba$. Of course, this is skew symmetric and bilinear, and one immediately verifies that it satisfies the Jacobi identity, so $(A, [\ ,\ ])$ is a Lie algebra over $\mathbb{K}$. In particular, for any $\mathbb{K}$–vector space V, the basic example $\mathfrak{gl}(V)$ of a Lie algebra is obtained by this construction starting from the unital associative algebra $(L(V,V), \circ)$. Hence studying representations of a Lie algebra $\mathfrak{g}$ is a special case of studying Lie algebra homomorphisms from $\mathfrak{g}$ to unital associative algebras, and the universal enveloping algebra of $\mathfrak{g}$ is a universal homomorphism of this type.

To construct the universal enveloping algebra, we have to return to multilinear algebra and first construct the tensor algebra $\mathcal{T}(V)$ of a $\mathbb{K}$–vector space V. For $k \in \mathbb{N}$ define $\mathcal{T}_k(V)$ to be the kth tensor power $\otimes^k V$, so in particular, $\mathcal{T}_0(V) = \mathbb{K}$ and $\mathcal{T}_1(V) = V$. Then define $\mathcal{T}(V) := \oplus_{k\in\mathbb{N}} \mathcal{T}_k(V)$, so this is an infinite dimensional $\mathbb{K}$–vector space, whose elements are finite linear combinations of elements of some $\otimes^k V$. Next, one defines a bilinear map $\otimes : \mathcal{T}_k(V) \times \mathcal{T}_\ell(V) \to \mathcal{T}_{k+\ell}(V)$ simply by

$$(v_1 \otimes \cdots \otimes v_k) \otimes (v_{k+1} \otimes \cdots \otimes v_{k+\ell}) := v_1 \otimes \cdots \otimes v_{k+\ell}.$$

These piece together to a bilinear map $\mathcal{T}(V) \times \mathcal{T}(V) \to \mathcal{T}(V)$ which is easily seen to make $\mathcal{T}(V)$ into a (non–commutative) associative algebra with unit element $1 \in \mathbb{K} = \mathcal{T}_0(V)$.

The tensor algebra $\mathcal{T}(V)$ has a simple universal property. Note first that we have an obvious inclusion $i : V = \mathcal{T}_1(V) \to \mathcal{T}(V)$. Now suppose that A is a unital associative algebra and that $\varphi : V \to A$ is any linear map. Then we define a linear map $\tilde{\varphi} : \mathcal{T}(V) \to A$ as follows: On $\mathcal{T}_0(V)$ we fix $\tilde{\varphi}$ by requiring that it maps $1 \in \mathbb{K}$ to the unit element of A, and on $\mathcal{T}_1(V) = V$ we put $\tilde{\varphi} = \varphi$. For $k > 1$, we consider the map $(v_1, \dots, v_k) \mapsto \varphi(v_1)\cdots\varphi(v_k)$, where on the right hand side we use the multiplication in A. Of course, this map is k–linear, so is induces a linear map $\tilde{\varphi} : \mathcal{T}_k(V) \to A$. All these maps together define $\tilde{\varphi} : \mathcal{T}(V) \to A$, and of course, $\tilde{\varphi} \circ i = \varphi$. From the construction it follows immediately that $\tilde{\varphi}$ is a homomorphism of unital associative algebras. On the other hand, any element of $\mathcal{T}(V)$ can be written as a linear combination of elements from some $\mathcal{T}_k(V)$ which in turn can be written as linear combinations of products of elements of $\mathcal{T}_1(V)$. Thus we conclude that any homomorphism from $\mathcal{T}(V)$ to a unital associative algebra is determined by its composition with i.

Hence we see that $\mathcal{T}(V)$ has the universal property that for any linear map $\varphi : V \to A$ into a unital associative algebra, there is a unique algebra homomorphism $\tilde{\varphi} : \mathcal{T}(V) \to A$ such that $\tilde{\varphi} \circ i = \varphi$. Similarly as in the case of the tensor product, it is easy to see that the pair $\mathcal{T}(V)$ is uniquely determined up to natural isomorphism by this universal property.

Next, recall that a (two–sided) ideal in an associative algebra A is a linear subspace $I \subset A$ such that for each $a \in A$ and $b \in I$ the products ab and ba lie in I. Having given an ideal $I \subset A$, the multiplication on A descends to a well defined associative multiplication on the quotient space A/I. If A has a unit element 1 then the class of 1 in A/I is a unit element for the multiplication on A/I. Note further that the intersection of an arbitrary family of ideals is again an ideal, so for any subset $C \subset A$ there is a smallest ideal of A which contains C, namely the intersection of all ideals containing C. This is called the ideal generated by C.

Suppose now that $\mathfrak{g}$ is a Lie algebra, and let $\mathcal{T}(\mathfrak{g})$ be the tensor algebra of the vector space $\mathfrak{g}$. Let I be the ideal generated by all elements of the form $X \otimes Y - Y \otimes X - [X, Y]$

for $X, Y \in \mathfrak{g}$, and define the *universal enveloping algebra* $\mathcal{U}(\mathfrak{g})$ as the quotient $\mathcal{T}(\mathfrak{g})/I$. By construction, this is a unital associative algebra and the inclusion $\mathfrak{g} \to \mathcal{T}(V)$ induces a linear map $i : \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$, which is easily seen to be injective. Moreover, for $X, Y \in \mathfrak{g}$ the commutator of $i(X)$ and $i(Y)$ in $\mathcal{U}(\mathfrak{g})$ is simply the class of $X \otimes Y - Y \otimes X$ in $\mathcal{U}(\mathfrak{g})$. By construction, this coincides with the class of $[X, Y]$, so we see that $i : \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$ is a homomorphism.

Consider a homomorphism $\varphi : \mathfrak{g} \to A$, where A is a unital associative algebra. Then we have the induced algebra homomorphism $\mathcal{T}(\mathfrak{g}) \to A$, whose kernel of course is an ideal in $\mathcal{T}(\mathfrak{g})$. By definition, for $X, Y \in \mathfrak{g}$, the induced homomorphism maps $X \otimes Y - Y \otimes X - [X, Y]$ to $\varphi(X)\varphi(Y) - \varphi(Y)\varphi(X) - \varphi([X, Y]) = 0$. Thus the ideal I is contained in the kernel, so we get a homomorphism $\tilde{\varphi} : \mathcal{U}(\mathfrak{g}) \to A$ such that $\tilde{\varphi} \circ i = \varphi$. Hence $\mathcal{U}(\mathfrak{g})$ has the universal property that for any homomorphism φ from $\mathfrak{g}$ to a unital associative algebra A there is a unique homomorphism $\tilde{\varphi} : \mathcal{U}(\mathfrak{g}) \to A$ of unital associative algebras such that $\tilde{\varphi} \circ i = \varphi$. In particular, any representation of $\mathfrak{g}$ uniquely extends to a representation of $\mathcal{U}(\mathfrak{g})$. Given a representation of $\mathfrak{g}$ on a vector space V, a subspace $W \subset V$ is $\mathcal{U}(\mathfrak{g})$–invariant if an only if it is $\mathfrak{g}$–invariant, and so on.

In the sequel, we will suppress the inclusion $i : \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$ and the symbol for the multiplication in $\mathcal{U}(\mathfrak{g})$. In particular, for $X_i \in \mathfrak{g}$, we can consider the element $X_1 \dots X_k \in \mathcal{U}(\mathfrak{g})$. Take a basis $\{X_1, \dots, X_n\}$ for $\mathfrak{g}$. For arbitrary indices $i_1, \dots, i_n \in \mathbb{N}$ we can then consider the monomial $X_1^{i_1} \cdots X_n^{i_n} \in \mathcal{U}(\mathfrak{g})$. Observe that in $\mathcal{U}(\mathfrak{g})$ we have $X_i X_j = X_j X_i + [X_i, X_j]$, and we can expand the bracket as a linear combination of the X_ℓ. Inductively, this implies that any product of basis elements which contains i_j occurrences of X_j can be written as a linear combination of $X_1^{i_1} \cdots X_n^{i_n}$ and of products containing less that $i_1 + \dots + i_n$ elements. Thus we conclude that the monomials $X_1^{i_1} \cdots X_n^{i_n}$ linearly generate $\mathcal{U}(\mathfrak{g})$. The Poincaré–Birkhoff–Witt theorem (usually referred to as PBW–theorem) states that these elements are also linearly independent, and thus form a basis of $\mathcal{U}(\mathfrak{g})$. A proof of this result can be found in [**Knapp**, III.2].

Remark. The tensor algebra $\mathcal{T}(V)$ over a vector space V can be used to construct several other important associative algebras. For example, there is the *symmetric algebra* $S(V)$ which is the quotient of $\mathcal{T}(V)$ by the ideal generated by all elements of the form $v_1 \otimes v_2 - v_2 \otimes v_1$ with $v_i \in V$. It is easy to see that $S(V) = \oplus_{k=0}^{\infty} S^k V$, and the product is given by $(v_1 \vee \dots \vee v_k) \vee (v_{k+1} \vee \dots \vee v_{k+\ell}) = v_1 \vee \dots \vee v_{k+\ell}$. Clearly, $S(V)$ is commutative. From the universal property of $\mathcal{T}(V)$ one immediately concludes that $S(V)$ has the universal property that any linear map $\varphi : V \to A$ into a commutative associative unital algebra uniquely extends to an algebra homomorphism $S(V) \to A$. For any vector space V the symmetric algebra $S(V^*)$ of the dual can be identified with the algebra of polynomials on V (with the pointwise multiplication).

Similarly, one has the exterior algebra $\Lambda(V)$, the quotient of $\mathcal{T}(V)$ by the ideal generated by all elements of the form $v_1 \otimes v_2 + v_2 \otimes v_1$ with $v_i \in V$. One shows that $\Lambda(V) = \oplus_{k=0}^{\dim(V)} \Lambda^k V$ of V, and the product is given in terms of the wedge product similarly as above. This algebra is *graded commutative*, i.e. $ab = (-1)^{k\ell} ba$ for $a \in \Lambda^k V$ and $b \in \Lambda^\ell V$. The exterior algebra also has a universal property which can be stated in terms of graded commutative algebras.

Given a vector space V with a symmetric bilinear form $b : V \times V \to \mathbb{K}$ one can define the Clifford algebra $Cl(V, b)$ which was mentioned in the context of spin representations in 4.8. This is defined as the quotient of the tensor algebra $\mathcal{T}(V)$ by the ideal I generated by all elements of the form $v \otimes w + w \otimes v - 2b(v, w)1$.

Finally, the tensor algebra can also be used to construct the free Lie algebra on a set X, which we have used in 3.11. Consider the tensor algebra $\mathcal{T}(F)$, where F is the free vector space generated by X. Since $\mathcal{T}(F)$ is associative, we may view it as a Lie algebra under the commutator and define $\mathfrak{F}(X)$ as the Lie subalgebra generated by $F = \mathcal{T}_1(F)$. Given a set map from X to some Lie algebra $\mathfrak{g}$, we can compose with the inclusion $\mathfrak{g} \to \mathcal{U}(\mathfrak{g})$ and prolong the resulting set map first to a linear map $F \to \mathcal{U}(\mathfrak{g})$ and then to a homomorphism $\mathcal{T}(F) \to \mathcal{U}(\mathfrak{g})$ of unital associative algebras. By construction, the restriction of this homomorphism to F has values in the image of $i_{\mathfrak{g}} : \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$, so the same is true for the Lie subalgebra generated by F. Therefore, we obtain the required homomorphism $\mathfrak{F}(X) \to \mathfrak{g}$ by composing the restriction of the homomorphism to $\mathfrak{F}(X)$ with $(i_{\mathfrak{g}})^{-1} : \mathrm{im}(i_{\mathfrak{g}}) \to \mathfrak{g}$.

4.10. Induced modules and Verma modules. Consider a Lie algebra $\mathfrak{g}$ with universal enveloping algebra $\mathcal{U}(\mathfrak{g})$ and a Lie subalgebra $\mathfrak{k} \leq \mathfrak{g}$. Then the inclusion $\mathfrak{k} \to \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$ induces a homomorphism $\mathcal{U}(\mathfrak{k}) \to \mathcal{U}(\mathfrak{g})$ of unital algebras. By the PBW–theorem this homomorphism is injective, so we may view $\mathcal{U}(\mathfrak{k})$ as a subalgebra of $\mathcal{U}(\mathfrak{g})$. More precisely, $\mathcal{U}(\mathfrak{k})$ is simply the subalgebra generated by $\mathfrak{k} \subset \mathfrak{g} \subset \mathcal{U}(\mathfrak{g})$.

Now suppose that W is a representation of $\mathfrak{k}$ and thus of $\mathcal{U}(\mathfrak{k})$. Consider the tensor product $\mathcal{U}(\mathfrak{g}) \otimes W$ and the linear subspace $\mathcal{N}$ generated by all elements of the form $xy \otimes w - x \otimes (y \cdot w)$ for $x \in \mathcal{U}(\mathfrak{g})$, $y \in \mathcal{U}(\mathfrak{k})$ and $w \in W$. Then form the quotient $M_{\mathfrak{k}}(W) := (\mathcal{U}(\mathfrak{g}) \otimes W)/\mathcal{N}$. Since this is a special case of the tensor product of a left module with a right module over the associative algebra $\mathcal{U}(\mathfrak{k})$, it is often denoted as $\mathcal{U}(\mathfrak{g}) \otimes_{\mathcal{U}(\mathfrak{k})} W$.

Of course, the universal enveloping algebra $\mathcal{U}(\mathfrak{g})$ is naturally a (infinite dimensional) representation of $\mathfrak{g}$ by multiplication from the left, i.e. for $x \in \mathcal{U}(\mathfrak{g})$ and $X \in \mathfrak{g}$ we have $X \cdot x := Xx$. That this really is a representation follows immediately from the fact that $XY - YX = [X, Y]$ in $\mathcal{U}(\mathfrak{g})$. The subspace $\mathcal{N}$ is visibly invariant under this action (since left and right multiplications commute). Therefore, we get an induced action of $\mathfrak{g}$ on the quotient $M_{\mathfrak{k}}(W)$. The representation $M_{\mathfrak{k}}(W)$ is called an *induced representation* or an *induced module* for $\mathfrak{g}$.

Induced modules have a nice universal property which is usually referred to as *Frobenius reciprocity.* Suppose that V is any representation of $\mathfrak{g}$. Then we can simply restrict the action to $\mathfrak{k}$, thus making V into a $\mathfrak{k}$–representation, called the restriction of the $\mathfrak{g}$–representation V. Using this we have:

PROPOSITION. *Let $\mathfrak{g}$ be a Lie algebra $\mathfrak{k} \leq \mathfrak{g}$ a Lie subalgebra, W a representation of $\mathfrak{k}$ and $M_{\mathfrak{k}}(W) = \mathcal{U}(\mathfrak{g}) \otimes_{\mathcal{U}(\mathfrak{k})} W$ the induced $\mathfrak{g}$–module. Then for any representation V of $\mathfrak{g}$ there is a natural bijection*

$$\mathrm{Hom}_{\mathfrak{g}}(M_{\mathfrak{k}}(W), V) \cong \mathrm{Hom}_{\mathfrak{k}}(W, V).$$

PROOF. Suppose that $\Phi : M_{\mathfrak{k}}(W) \to V$ is a $\mathfrak{g}$–homomorphism and define $\varphi : W \to V$ by $\varphi(w) := \Phi(1 \otimes w)$, where we denote by $1 \otimes w$ also the class of this element of $\mathcal{U}(\mathfrak{g}) \otimes W$ in the quotient $M_{\mathfrak{k}}(W)$. For $Y \in \mathfrak{k}$ we have $\varphi(Y \cdot w) = \Phi(1 \otimes (Y \cdot w))$. But in the quotient the latter element represents the same class as $Y \otimes w$, and $\Phi(Y \otimes w) = Y \cdot \Phi(1 \otimes w) = Y \cdot \varphi(w)$, since Φ is a $\mathfrak{g}$ homomorphism. Hence φ is a $\mathfrak{k}$–homomorphism.

Conversely, given a $\mathfrak{k}$–homomorphism $\varphi : W \to V$, we define $\tilde{\Phi} : \mathcal{U}(\mathfrak{g}) \otimes W \to V$ by $\tilde{\Phi}(x \otimes w) := x \cdot \varphi(w)$, where we use the extension to $\mathcal{U}(\mathfrak{g})$ of the $\mathfrak{g}$–action on V. Since φ is a $\mathfrak{k}$–homomorphism it is also a $\mathcal{U}(\mathfrak{k})$–homomorphism. Hence we conclude that for $x \in \mathcal{U}(\mathfrak{g})$ and $y \in \mathcal{U}(\mathfrak{k})$ we get

$$\Phi(xy \otimes w) = x \cdot y \cdot \varphi(w) = x \cdot \varphi(y \cdot w) = \Phi(x \otimes (y \cdot w)).$$

Hence we see that Φ vanishes on $\mathcal{N}$ and thus factors to a linear map $M_{\mathfrak{k}}(W) \to V$, which by construction is a $\mathfrak{g}$–homomorphism. One immediately checks that the two constructions are inverse to each other. $\square$

Verma modules are a special class of induced modules. Suppose that $\mathfrak{g}$ is a complex semisimple Lie algebra and that we have chosen a Cartan subalgebra $\mathfrak{h} \leq \mathfrak{g}$ and a notion of positivity on $\mathfrak{h}_0^*$, so we have positive roots Δ^+. Then we define the *Borel subalgebra* $\mathfrak{b} \leq \mathfrak{g}$ as the direct sum $\mathfrak{h} \oplus \mathfrak{n}_+$, where $\mathfrak{n}_+$ is the direct sum of all root spaces corresponding to positive roots. Of course, $[\mathfrak{b}, \mathfrak{b}] = \mathfrak{n}_+$ and visibly $\mathfrak{n}_+$ is nilpotent, so $\mathfrak{b}$ is solvable. It is easy to see that $\mathfrak{b} \leq \mathfrak{g}$ is a maximal solvable subalgebra and it can be shown that any maximal solvable subalgebra of $\mathfrak{g}$ is conjugate to $\mathfrak{b}$ under an inner automorphism of $\mathfrak{g}$.

Since $\mathfrak{b}$ is solvable, we know that irreducible representations of $\mathfrak{b}$ are one–dimensional. More precisely, any such representation is given by a linear functional on $\mathfrak{b}/[\mathfrak{b}, \mathfrak{b}] = \mathfrak{h}$. Hence for any functional $\lambda \in \mathfrak{h}^*$ we get a one–dimensional irreducible representation $\mathbb{C}_\lambda$ of $\mathfrak{b}$ on $\mathbb{C}$, defined by $H \cdot z = \lambda(H)z$ for $H \in \mathfrak{h}$ and $X \cdot z = 0$ for $X \in \mathfrak{n}_+$. The induced module $M(\lambda) := M_{\mathfrak{b}}(\mathbb{C}_\lambda)$ is called the *Verma module* of $\mathfrak{g}$ with highest weight λ.

By construction, $M(\lambda)$ is generated as a $\mathcal{U}(\mathfrak{g})$–module, and thus as a $\mathfrak{g}$–module, by the class of the element $1 \otimes 1$. Moreover, Frobenius reciprocity takes a particularly nice and simple form for Verma modules: For a $\mathfrak{g}$–representation V a linear map $\mathbb{C}_\lambda \to V$ is of course determined by the image of $1 \in \mathbb{C}_\lambda$. Mapping 1 to $v_0 \in V$ defines a $\mathfrak{b}$–homomorphism if and only if each $H \in \mathfrak{h}$ acts on v_0 by multiplication by $\lambda(H)$, while each X in a positive root space acts trivially. But this means that $\mathrm{Hom}_{\mathfrak{b}}(\mathbb{C}_\lambda, V)$ and thus by the proposition $\mathrm{Hom}_{\mathfrak{g}}(M(\lambda), V)$ is in bijective correspondence with the space of highest weight vectors in V of weight λ. In particular, if V is irreducible and contains a highest weight vector of weight λ, then we obtain a homomorphism $M(\lambda) \to V$ which must be surjective since V is irreducible. Hence, any such irreducible representation can be realized as a quotient of a Verma module.

To proceed further, we have to describe the structure of $M(\lambda)$ in a bit more detail using the PBW–theorem. Let $\mathfrak{n}_-$ be the direct sum of all negative root spaces, and choose a basis of $\mathfrak{g}$ which is the union of a basis $\{X_1, \dots, X_\ell\}$ of $\mathfrak{n}_-$, a basis $\{H_1, \dots, H_r\}$ of $\mathfrak{h}$, and a basis $\{Z_1, \dots, Z_\ell\}$ of $\mathfrak{n}_+$. By the PBW–theorem the monomials $X_1^{i_1} \cdots Z_\ell^{j_\ell}$ form a basis for $\mathcal{U}(\mathfrak{g})$. If x is such a monomial which contains any Z's then the class of $x \otimes 1$ in $M(\lambda)$ visibly vanishes. If there are no Z's, then we can move the $H's$ to the other side, which causes multiplication by a factor only. This easily implies that the classes of the elements $X_1^{i_1} \cdots X_\ell^{i_\ell} \otimes 1$ form a basis of $M(\lambda)$, so $M(\lambda)$ is isomorphic to $\mathcal{U}(\mathfrak{n}_-) \otimes \mathbb{C}_\lambda$ as a vector space. Moreover, denoting by α_i the positive root such that $X_i \in \mathfrak{g}_{-\alpha_i}$, one immediately verifies that the class of $X_1^{i_1} \cdots X_\ell^{i_\ell} \otimes 1$ is a weight vector of weight $\lambda - i_1\alpha_1 - \cdots - i_\ell\alpha_\ell$. In particular, the λ–weight space of $M(\lambda)$ is one–dimensional.

Now assume that $N \subset M(\lambda)$ is a $\mathfrak{g}$–subrepresentation. Since $\mathfrak{h}$ by construction acts diagonalizably on $M(\lambda)$ we see from 4.2 that any element of N can be decomposed into a finite sum of weight vectors of different weights and each of the components lies in N. In particular, if some of the elements of N has a non–zero component of weight λ, then N contains the (one–dimensional) λ–weight space of $M(\lambda)$. Since $M(\lambda)$ is generated by this weight space, this implies $N = M(\lambda)$. Assume that $\{N_i : i \in I\}$ is an arbitrary family of proper $\mathfrak{g}$–subrepresentations of $M(\lambda)$ and consider the subspace $N \subset M(\lambda)$ spanned by all the N_i. Then any element of N is a linear combination of elements of N_i's, so N is a subrepresentation. On the other hand, splitting such an element into weight components we can never obtain a non–zero component of weight λ, and thus N is a proper submodule of $M(\lambda)$. This immediately implies that there is a unique maximal proper submodule $\mathcal{N} \subset M(\lambda)$, which may be defined as the subspace spanned by all proper subrepresentations of $M(\lambda)$.

Since $\mathcal{N} \neq M(\lambda)$ the quotient $L(\lambda) := M(\lambda)/\mathcal{N}$ is a nontrivial space, and of course it carries as representation of $\mathfrak{g}$. The element $1 \otimes 1$ visibly descends to a highest weight vector of weight λ in $L(\lambda)$. Now we can formulate the general result on existence of irreducible representations:

THEOREM. *For any linear functional $\lambda \in \mathfrak{h}^*$ the representation $L(\lambda)$ is irreducible and admits a highest weight vector of weight λ. Any irreducible representation V of $\mathfrak{g}$ which admits a highest weight vector of weight λ is isomorphic to $L(\lambda)$.*

If the weight λ is dominant and algebraically integral, then the representation $L(\lambda)$ is finite dimensional.

SKETCH OF PROOF. We have already observed that $L(\lambda)$ admits a highest weight vector of weight λ. If $W \subset L(\lambda)$ is a non–zero invariant subspace, then the preimage of W in $M(\lambda)$ is a $\mathfrak{g}$–submodule which strictly contains $\mathcal{N}$. But since $\mathcal{N}$ is the maximal proper submodule of $M(\lambda)$ we conclude that this preimage is all of $M(\lambda)$ and thus $W = L(\lambda)$. Hence $L(\lambda)$ is irreducible.

If V is an irreducible representation of $\mathfrak{g}$ which admits a highest weight vector of weight λ, then this highest weight vector gives rise to a non–zero homomorphism $f : M(\lambda) \to V$. The image of $\mathcal{N}$ under this homomorphism is a subrepresentation of V, which cannot be all of V since it does not contain any weight components of weight λ. Therefore, $\mathcal{N}$ is contained in the kernel, so f factors to a non–trivial homomorphism $L(\lambda) \to V$ which must be an isomorphism by irreducibility.

The proof that $L(\lambda)$ is finite dimensional for dominant integral λ is more involved. One shows that the set of weights of $L(\lambda)$ (including multiplicities) is invariant under the Weyl group. This is done using the representation theory of $\mathfrak{sl}(2, \mathbb{C})$ applied to the Lie subalgebras $\mathfrak{s}_\alpha$ for $\alpha \in \Delta$. Since the weight spaces of $M(\lambda)$ are immediately seen to be finite dimensional and there are only finitely many dominant weights in $M(\lambda)$, this implies that $L(\lambda)$ is finite dimensional. □

CHAPTER 5

Tools for dealing with finite dimensional representations

From Theorem 2.9 we know that any finite dimensional representation of a complex semisimple Lie algebra can be decomposed into a direct sum of irreducible representations. By the theorem of the highest weight in 4.3 the set of finite dimensional irreducible representations is in bijective correspondence with the set of dominant algebraically integral weights. Hence there remain two obvious problems: On the one hand, one has to ask how a decomposition into irreducible pieces can actually be carried out, or at least look for ways to determine which irreducible pieces show up in this decomposition. On the other hand, neither of the two methods for proving the existence of irreducible representations described in chapter 4 gives detailed information on these representations. So there is the natural task to get more information on the irreducible representations. For example, one may ask for the dimension of irreducible representations or the multiplicity of weights. We shall see that these two basic problems are closely related. In this chapter, we will discuss various tools that can be used to deal with these problems. In may cases, we will omit or only sketch proofs.

Decomposing representations

5.1. The isotypical decomposition. We start with the question of decomposing a finite dimensional representation of a complex semisimple Lie algebra $\mathfrak{g}$ into irreducible pieces. The first observation to make here is that one cannot expect to really get down to irreducible pieces in a natural way. Namely, suppose that W is irreducible and consider the direct sum $\tilde{W} = W \oplus W$. Take a highest weight vector $w_0 \in W$ and two complex numbers a, b with $(a, b) \neq (0, 0)$. Then of course (aw_0, bw_0) is a highest weight vector in $\tilde{W}$ and thus generates an irreducible subrepresentation isomorphic to W. Otherwise put, any choice of a basis in the two–dimensional highest weight space of $\tilde{W}$ leads to an identification of $\tilde{W}$ with the direct sum of two copies of W and there is no canonical choice available.

Hence what we should try to do is decomposing a given representation V of $\mathfrak{g}$ into pieces corresponding to non–isomorphic irreducible representations. Doing this is fairly straightforward: Fix a dominant integral weight λ and let us denote by Γ_λ the irreducible representation with highest weight λ. Of course, any linear combination of highest weight vectors in V with weight λ is again a highest weight vector with weight λ, so there is the subspace V^0_λ of highest weight vectors of weight λ. The dimension $\dim(V^0_\lambda)$ is called the *multiplicity* $m_\lambda(V)$ of Γ_λ in V. Now we define the λ*–isotypical component* $V_{(\lambda)} \subset V$ to be the $\mathfrak{g}$–subrepresentation generated by V^0_λ.

PROPOSITION. *Let V be a finite dimensional representation of a complex semisimple Lie algebra $\mathfrak{g}$. Then we have:*
(1) Choosing a basis in V^0_λ induces an isomorphism between $V_{(\lambda)}$ and the direct sum of $m_\lambda(V)$ many copies of Γ_λ.
(2) $V = \oplus_{\lambda: m_\lambda(V)>0} V_{(\lambda)}$, so V decomposes into the direct sum of its isotypical components.

(3) If W is another finite dimensional representation of $\mathfrak{g}$, the there is a natural bijection $\mathrm{Hom}_{\mathfrak{g}}(V,W) \cong \oplus_{\lambda:m_\lambda(V)>0} L(V^0_\lambda, W^0_\lambda)$.

PROOF. (1) Put $\ell = m_\lambda(V)$ and choose a basis $\{v_1,\dots,v_\ell\}$ of V^0_λ. Then from Theorem 4.2 we know that each v_i generates an irreducible subrepresentation W_i of $V_{(\lambda)}$ with highest weight λ, so $W_i \cong \Gamma_\lambda$. Moreover, by construction these subspaces span all of $V_{(\lambda)}$, since their span contains V^0_λ. Finally, it is clear that v_i does not lie in the span of the subspaces W_j for $j \neq i$. But this immediately implies that the intersection of W_i with this sum is a proper subrepresentation of W_i and thus zero by irreducibility. Hence we have $V_{(\lambda)} = W_1 \oplus \dots \oplus W_\ell$.
(2) By construction, the subspace spanned by all the isotypical components $V_{(\lambda)}$ is a subrepresentation of V, so by complete reducibility there is an invariant complement. If this would be nonzero, then it would contain at least one highest weight vector, but that vector would lie in one of the $V_{(\lambda)}$. Thus, the isotypical components span V and it remains to show that their sum is direct.

Let λ_0 be the highest weight such that $m_{\lambda_0}(V) > 0$. By construction, the isotypical components $V_{(\mu)}$ for $\mu \neq \lambda_0$ contain only weight vectors of weights strictly smaller than λ_0. Thus the interesction of $V_{(\lambda_0)}$ with the sum of the other isotypical components cannot contain any highest weight vectors, and hence must be zero. This means that $V_{(\lambda_0)}$ splits off as a direct summand, and inductively the claim follows.
(3) If $\varphi : V \to W$ is a homomorphism and $v \in V$ is a highest weight vector of weight λ, then $\varphi(v)$ is a highest weight vector of weight λ. Thus for each λ, we have $\varphi(V^0_\lambda) \subset W^0_\lambda$ so restriction defines a map $\mathrm{Hom}_{\mathfrak{g}}(V,W) \to \oplus L(V^0_\lambda, W^0_\lambda)$, which is obviously linear. Moreover, the restriction of φ to $V^0_{(\lambda)}$ determines the restriction of φ to $V_{(\lambda)}$, so (2) implies that this map is injective.

Conversely, assume that for each λ such that $m_\lambda(V) > 0$ we have given a linear map $\varphi_\lambda : V^0_\lambda \to W^0_\lambda$. Then consider $V \oplus W$. For $v \in V^0_\lambda$ we see that $(v, \varphi_\lambda(v))$ is a highest weight vector, and we denote by $\tilde{V}(\lambda)$ the $\mathfrak{g}$–subrepresentation generated by these Vectors. Finally, we define $\tilde{V}$ to be the subspace spanned by all $\tilde{V}(\lambda)$. The restriction of the first projection to $\tilde{V}$ defines a surjective homomorphism $\tilde{V} \to V$. As in (1) and (2) one sees that $\tilde{V}(\lambda)$ is a direct sum of $m_\lambda(V)$ copies of Γ_λ and the $\tilde{V}$ is the direct sum of the $\tilde{V}(\lambda)$. This easily implies that the first projection induces an isomorphism $\tilde{V} \to V$. Composing the inverse of this homomorphism with the restriction of the second projection to $\tilde{V}$, we obtain a homomorphism φ which induces the maps φ_λ on the spaces of highest weight vectors. □

REMARK. (1) The first two parts of this proposition show that a finite dimensional representation V is determined up to isomorphism by the multiplicities $m_\lambda(V)$.
(2) The last part of the proposition can be viewed as a generalization of Schur's lemma from 2.4. Indeed, for $V = \Gamma_\lambda$ we get $V^0_\mu = \{0\}$ for $\mu \neq \lambda$ while V^0_λ is one–dimensional. Hence part (3) in this case implies that $\mathrm{Hom}_{\mathfrak{g}}(V,V) \cong L(V^0_\lambda, V^0_\lambda)$ and on the right hand side we exactly have the multiples of the identity.

5.2. The Casimir element. We next discuss a tool that can be used to explicitly split a finite dimensional representation into isotypical components. The idea for this construction is fairly simple: We know from 4.9 that on a representation V of $\mathfrak{g}$ we automatically have an action of the universal enveloping algebra $\mathcal{U}(\mathfrak{g})$. Now suppose that we have an element $x \in \mathcal{U}(\mathfrak{g})$ which lies in the center of $\mathcal{U}(\mathfrak{g})$, i.e. which has the property that $xy = yx$ for all $y \in \mathcal{U}(\mathfrak{g})$. Then in particular $Xx = xX$ for all $X \in \mathfrak{g}$, which shows that the action of x on any representation V of $\mathfrak{g}$ commutes with the action of each element of $\mathfrak{g}$. In particular, by Schur's lemma x has to act as a multiple of the identity on any irreducible representation of $\mathfrak{g}$. One way to use this (which is important for studying infinite dimensional representations) is to analyze the action of the whole center of $\mathcal{U}(\mathfrak{g})$, which leads to the notion of infinitesimal character or central character. For many purposes it is sufficient to look at a specific element of this center, which is called the *Casimir element*, and we will only discuss this.

The idea for constructing the Casimir element is rather simple: In 4.9 we have constructed $\mathcal{U}(\mathfrak{g})$ as a quotient of the tensor algebra $\mathcal{T}(\mathfrak{g})$. In particular, we have a canonical map $\mathfrak{g} \otimes \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$. By definition, the action of $\mathfrak{g}$ on $\mathfrak{g} \otimes \mathfrak{g}$ is given by $X \cdot (Y \otimes Z) = [X, Y] \otimes Z + Y \otimes [X, Z]$. Thus the class of $X \cdot (Y \otimes Z)$ in $\mathcal{U}(\mathfrak{g})$ can be written as $(XY - YX)Z + Y(XZ - ZX) = XYZ - YZX$. In particular, if we find an $\mathfrak{g}$–invariant element in $\mathfrak{g} \otimes \mathfrak{g}$, then its class in $\mathcal{U}(\mathfrak{g})$ will commute with any $X \in \mathfrak{g}$. Since $\mathcal{U}(\mathfrak{g})$ is generated by $\mathfrak{g}$ this implies that the class will lie in the center.

Finding an invariant element in $\mathfrak{g} \otimes \mathfrak{g}$ is however very easy. The Killing form B identifies $\mathfrak{g}$ with $\mathfrak{g}^*$ as a representation of $\mathfrak{g}$ and thus $\mathfrak{g} \otimes \mathfrak{g}$ with $\mathfrak{g}^* \otimes \mathfrak{g} \cong L(\mathfrak{g}, \mathfrak{g})$. But of course, the identity map in $L(\mathfrak{g}, \mathfrak{g})$ is $\mathfrak{g}$–invariant, so we get an element with the required property, and this is the definition of the Casimir element $\Omega \in \mathcal{U}(\mathfrak{g})$. We can easily describe Ω explicitly: Taking a basis $\{X_i\}$ for $\mathfrak{g}$ and the dual basis $\{Y_i\}$ with respect to the Killing form, we have $\Omega = \sum_i Y_i X_i$. As an alternative to the considerations above, one may of course verify directly that $X\Omega = \Omega X$ for all $X \in \mathfrak{g}$ similarly as in the proof of Theorem 2.9.

To describe the properties of Ω we need one more ingredient. We define the *lowest form* $\delta \in \mathfrak{h}_0^*$ of the Lie algebra $\mathfrak{g}$ as half of the sum of all positive roots. Consider a simple root $\alpha \in \Delta^0$ and the corresponding simple reflection s_α on $\mathfrak{h}_0^*$. By definition, $s_\alpha(\alpha) = -\alpha$, and we have observed in 3.9 that for a positive root $\beta \neq \alpha$ we have $s_\alpha(\beta) \in \Delta^+$, so s_α only permutes the other positive roots. Thus we conclude that

$$s_\alpha(\delta) = \frac{1}{2} \sum_{\beta \in \Delta^+} s_\alpha(\beta) = -\alpha + \delta.$$

By definition of the simple reflection, this implies that $\frac{2\langle \delta, \alpha \rangle}{\langle \alpha, \alpha \rangle} = 1$ for all $\alpha \in \Delta^0$. Hence we may alternatively characterize δ as the sum of all fundamental weights, so this is the smallest integral element in the interior of the dominant Weyl chamber.

PROPOSITION. *Let $\mathfrak{g}$ be a complex semisimple Lie algebra with lowest form $\delta \in \mathfrak{h}_0^*$, and let $\langle\ ,\ \rangle$ be the inner product on $\mathfrak{h}_0^*$ induced by the Killing form.*

Then for a finite dimensional representation V of $\mathfrak{g}$ and a dominant integral weight $\lambda \in \mathfrak{h}_0^*$ the Casimir element $\Omega \in \mathcal{U}(\mathfrak{g})$ acts on the isotypical component $V_{(\lambda)} \subset V$ by mutlipication by $\langle \lambda, \lambda \rangle + 2\langle \lambda, \delta \rangle \in \mathbb{R}$.

PROOF. This is proved by choosing a special basis for $\mathfrak{g}$. We know from 3.5 that the decomposition $\mathfrak{g} = \mathfrak{h} \oplus \oplus_{\alpha \in \Delta^+} (\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha})$ is orthogonal with respect to the Killing form B. For $\alpha \in \Delta^+$ choose $E_\alpha \in \mathfrak{g}_\alpha$ and $F_\alpha \in \mathfrak{g}_{-\alpha}$ such that $B(E_\alpha, F_\alpha) = 1$. Further, we know that the restriction of B to $\mathfrak{h}_0^*$ is positive definite, so we may choose an orthonormal basis $\{H_i : i = 1, \dots, n\}$ of $\mathfrak{h}_0^*$, which then is a complex basis of $\mathfrak{h}$. From above, we conclude that we may write the Casimir element as

$$\Omega = \sum_{i=1}^{n} H_i H_i + \sum_{\alpha \in \Delta^+} (E_\alpha F_\alpha + F_\alpha E_\alpha).$$

In $\mathcal{U}(\mathfrak{g})$ we further have $E_\alpha F_\alpha = F_\alpha E_\alpha + [E_\alpha, F_\alpha]$, and $[E_\alpha, F_\alpha] = H_\alpha$, the element characterized by $B(H, H_\alpha) = \alpha(H)$ for all $H \in \mathfrak{h}$. Summing the H_α over all positive roots, we obtain $2H_\delta$ and thus

$$\Omega = 2H_\delta + \sum_{i=1}^{n} H_i H_i + 2 \sum_{\alpha\in\Delta^+} F_\alpha E_\alpha.$$

On a highest weight vector, the last sum by definition acts trivially, while elements of $\mathfrak{h}$ act by some scalar, so Ω acts by multiplication by a scalar on each highest weight vector. Since Ω commutes with the action of any $X \in \mathfrak{g}$ it acts by a scalar on each isotypical component. If $v \in V^0_\lambda$, then the action of H_δ on v is given by multiplication by $\lambda(H_\delta) = \langle\lambda, \delta\rangle$. The remaining sum acts on v by multiplication by $\sum_i \lambda(H_i)^2$, and since the H_i are orthonormal with respect to B this equals $\langle\lambda, \lambda\rangle$. □

Suppose we have given a representation V and we know the dominant integral weights λ such that $m_\lambda(V) > 0$. Then from the proposition we can compute the eigenvalues of Ω on the isotypical components. If these are all different, say $a_1, \dots, a_\ell$ then the projection onto the a_i–eigenspace is given by $\prod_{j\neq i} \frac{1}{a_i-a_j}(\Omega - a_j \,\mathrm{id})$, so we obtain explicit projections onto the isotypical components.

Example. The following example on the one hand shows the use of the Casimir element. On the other hand, it illustrates that for a complete understanding of a representation the main additional input needed is a better understanding of irreducible representations.

Consider $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C})$, let $V = \mathbb{C}^n$ be the standard representation with standard basis $\{v_1, \dots, v_n\}$ and consider the third tensor power $\otimes^3 V$. From 4.7 we know that each v_i is a weight vector of weight e_i. Using this, we can immediately determine all weights of $\otimes^3 V$: Any element $v_{i_1} \otimes v_{i_2} \otimes v_{i_3}$ is a weight vector of weight $e_{i_1} + e_{i_2} + e_{i_3}$. If the three vectors are all different, then there are six vectors with the given weight (corresponding to the permutations of the factors). If two of them are equal, then there are only three vectors (determined by the position of the "other" basis vector) while for all three equal, there is only one possible weight vector. Hence we see that the weights of $\otimes^3 V$ are $3e_i$ with multiplicity one, $2e_i + e_j$ for $j \neq i$ with multiplicity 3 and $e_i + e_j + e_k$ with three different indices with multiplicity 6. The highest weight if of course $3e_1$, so we see that $m_{3e_1}(\otimes^3 V) = 1$, and this corresponds to the irreducible subrepresentation $S^3 V \subset \otimes^3 V$. Looking at the standard basis of $S^3 V$ we see that it contains each of the above weights with multiplicity one.

Taking an invariant complement $V' \subset \otimes^3 V$ (for example the kernel of the symmetrization map) we see that the weights of V' are $2e_i + e_j$ with multiplicity 2 and $e_i + e_j + e_k$ with multiplicity 5. The highest among these weights is $2e_1 + e_2$, which in terms of the fundamental weights $\omega_1, \dots, \omega_{n-1}$ equals $\omega_1 + \omega_2$, see 4.7. Hence we conclude that $m_{2e_1+e_2}(\otimes^3 V) = m_{2e_1+e_2}(V') = 2$, so we get two copies of the irreducible representation $\Gamma_{2e_1+e_2}$. Recall that the Weyl group of $\mathfrak{g}$ is $\mathfrak{S}_n$ which acts by permutations of the e_i. Symmetry of the weights under the Weyl group thus implies that $V_{(2e_1+e_2)}$ contains each of the weights $2e_i + e_j$ for $j \neq i$ with multiplicity two. Hence the only possible weights in an invariant complement V'' to $V'_{(2e_1+e_2)}$ in V' are $e_i + e_j + e_k$ for i, j, k all different. The multiplicity of this weight in V'' is 5 minus twice the multiplicity of $e_i + e_j + e_k$ in $\Gamma_{2e_1+e_2}$, so this may be 1, 3, or 5. The only dominant weight among these is $e_1 + e_2 + e_3$, so we must have $m_{e_1+e_2+e_3}(\otimes^3 V) = 1$, 3, or 5, and to determine the multiplicity it suffices to determined the multiplicity of the weight $e_1 + e_2 + e_3$ in $\Gamma_{2e_1+e_2}$. Having done this, we obtain $m_{e_1+e_2+e_3}(V)$ many copies of $\Gamma_{e_1+e_2+e_3} \cong \Lambda^3 V$ and by symmetry of the weights under the Weyl group, these exhaust all of V''.

Thus we conclude that the dominant integral weights λ such that $m_\lambda(\otimes^3 V) > 0$ are exactly $\lambda_0 = 3e_1$, $2e_1 + e_2 = \lambda_0 - \alpha_1$ and $e_1 + e_2 + e_3 = \lambda_0 - 2\alpha_1 - \alpha_2$. To compute the eigenvalues of the Casimir element on the corresponding isotypical components, we observe that $\langle\lambda, \lambda\rangle + 2\langle\lambda, \delta\rangle = \langle\lambda + \delta, \lambda + \delta\rangle - \langle\delta, \delta\rangle$. Hence replacing λ by $\lambda - \mu$, the eigenvalue changes by $-2\langle\mu, \lambda + \delta\rangle + \langle\mu, \mu\rangle$. Of course, replacing the Killing form by a multiple just changes Ω by some multiple, so we can use the inner products as before to decide whether we get the same eigenvalue on two isotypical components. In terms of the fundamental weights, we have $\lambda_0 = 3\omega_1$ and $\delta = \omega_1 + \cdots + \omega_{n-1}$, so the inner products of $\lambda_0 + \delta$ with a root is easy to compute. In particular, $\langle\alpha_1, \lambda_0 + \delta\rangle = 4$ and $\langle\alpha_1, \alpha_1\rangle = 2$, so the eigenvalue on $(\otimes^3 V)_{2e_1+e_2}$ differs from the one on S^3V by -6. On the other hand, $\langle 2\alpha_1 + \alpha_2, \lambda_0 + \delta\rangle = 9$ and $\langle 2\alpha_1 + \alpha_2, 2\alpha_1 + \alpha_2\rangle = 6$, and thus the eigenvalue on $(\otimes^3 V)_{e_1+e_2+e_3}$ differs from the one on S^3V by -12. Hence we conclude that the eigenvalues on the three isotypical components are pairwise different, so the splitting of $\otimes^3 V$ into isotypical components coincides with the splitting into eigenspaces for Ω.

Formulae for multiplicities, characters, and dimensions

5.3. Formulae for multiplicities of weights. Let us turn to the question of a better understanding of irreducible representations. As we have seen in the last example, it is very useful to have a general way to determine the dimensions of the weight spaces of an irreducible representation of given highest weight. There are several different formulae for these multiplicities available, each of which has its advantages and disadvantages.

We first discuss the *Freudenthal multiplicity formula*. This expresses the multiplicity of a weight μ in the irreducible representation Γ_λ (i.e. the dimension of the weight space $(\Gamma_\lambda)_\mu$) in terms of the multiplicities of higher weights. Since we know that the weights and their multiplicities are invariant under the Weyl group, it suffices to determine the multiplicities of the dominant weights. Moreover, we know in advance that the highest weight λ of Γ_λ has multiplicity one, so we can compute the multiplicities in an iterative way. To state the formula, let us denote by $n_\mu(\Gamma_\lambda)$ the multiplicity of μ in Γ_λ, let $\langle\ ,\ \rangle$ be the inner product on $\mathfrak{h}_0^*$ induced by the Killing form and $\|\ \|$ the corresponding norm, i.e. $\|\mu\|^2 = \langle\mu, \mu\rangle$. Then the Freudenthal multiplicity formula reads as

$$(2\langle\lambda - \mu, \mu + \delta\rangle + \|\lambda - \mu\|^2)n_\mu(\Gamma_\lambda) = 2\sum_{\alpha\in\Delta^+}\sum_{k\geq 1}\langle\mu + k\alpha, \alpha\rangle n_{\mu+k\alpha}(\Gamma_\lambda),$$

where δ is the lowest form of the Lie algebra $\mathfrak{g}$, see 5.2. If $\mu \neq \lambda$ is a dominant weight of Γ_λ we know that $\lambda - \mu$ is a linear combination of simple roots with non–negative integral coefficients. This immediately implies that $\langle\lambda - \mu, \mu + \delta\rangle > 0$, so we see that the numerical factor in the left hand side of the Freudenthal multiplicity formula is positive. Hence knowing the multiplicities of the higher weights $\mu + k\alpha$ we can compute $n_\mu(\Gamma_\lambda)$ from the formula.

Before we sketch the proof, let us apply the formula to sort out the remaining open point in Example 5.2: For $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C})$ we want to determine $n_\mu(\Gamma_\lambda)$ with $\mu = e_1 + e_2 + e_3$ and $\lambda = 2e_1 + e_2$. Denoting the simple roots by $\alpha_1, \dots, \alpha_{n-1}$ we have $\alpha_i = e_i - e_{i+1}$, and hence $\mu = \lambda - (\alpha_1 + \alpha_2)$. Hence we conclude that the only expressions of the form $\mu + k\alpha$ with $\alpha \in \Delta$ and $k \geq 1$ are $\mu + 1\alpha_1 = \lambda - \alpha_2$, $\mu + 1\alpha_2 = \lambda - \alpha_1$ and $\mu + 1(\alpha_1 + \alpha_2) = \lambda$. Of course, we know that $n_\lambda(\Gamma_\lambda) = 1$. On the other hand, $\lambda - \alpha_1 = e_1 + 2e_2$ and $\lambda - \alpha_2 = 2e_1 + e_3$, so these weights both lie in the Weyl orbit of λ and thus have multiplicity one, too. (Alternatively, we could also determine their multiplicity using the

Freudenthal formula). In terms of the fundamental weights ω_i we have $\mu = \omega_3$, so this has zero inner product with α_1, α_2 and $\alpha_1+\alpha_2$. Thus for each of the three roots we have $\langle \mu+\alpha,\alpha\rangle = \langle\alpha,\alpha\rangle = 2$, and thus the right hand side of the Freudenthal formula gives $2(2+2+2) = 12$. For the left hand side we get $(2\langle\alpha_1+\alpha_2,\mu+\delta\rangle + \|\alpha_1+\alpha_2\|^2)n_\mu(\Gamma_\lambda) = 6n_\mu(\Gamma_\lambda)$, so we conclude that $n_\mu(\Gamma_\lambda) = 2$. Therefore, we conclude from the discussion in Example 5.2 that $\otimes^3 V \cong S^3V \oplus (\Gamma_{2e_1+e_2} \oplus \Gamma_{2e_1+e_2}) \oplus \Lambda^3 V$.

Next we briefly sketch the proof of the Freudenthal formula for a dominant weight μ, details for this case and the general proof can be found in [**Fulton-Harris**, §25.1]: Consider the Casimir element Ω introduced in 5.2. From Proposition 5.2 we know that Ω acts on $V = \Gamma_\lambda$ by multiplication by $\|\lambda\|^2 + 2\langle\lambda,\delta\rangle > 0$. In particular, $\Omega(V_\mu) = V_\mu$ and the trace of the restriction $\Omega|_{V_\mu}$ is given by $(\|\lambda\|^2 + 2\langle\lambda,\delta\rangle)n_\mu(\Gamma_\lambda)$. Freudenthal's formula is then obtained by computing this trace in a different way. From the proof of Proposition 5.2 we know that $\Omega = 2H_\delta + \sum_{i=1}^n H_iH_i + 2\sum_{\alpha\in\Delta^+} F_\alpha E_\alpha$. Visibly, any of the three parts leaves V_μ invariant, the first summand acts by multiplication by $2\langle\mu,\delta\rangle$ while the middle sum acts by multiplication by $\|\mu\|^2$. Bringing these two parts to the other side, we see that

$$(\|\lambda\|^2 + 2\langle\lambda-\mu,\delta\rangle - \|\mu\|^2)n_\mu(\Gamma_\lambda) = \operatorname{tr}\left((\sum_{\alpha\in\Delta^+} F_\alpha E_\alpha)|_{V_\mu}\right).$$

One immediately verifies that the left hand side of this equation equals the left hand side of the Freudenthal formula. To compute the contribution of the right hand side, we fix $\alpha \in \Delta^+$, consider the subalgebra $\mathfrak{s}_\alpha$ and the subspace $\oplus_{n\in\mathbb{Z}} V_{\mu+n\alpha}$ which is visibly invariant under $\mathfrak{s}_\alpha$. From the representation theory of $\mathfrak{sl}(2,\mathbb{C})$ we know that this subspace forms an unbroken string of the form $V_\nu \oplus V_{\nu-\alpha} \oplus \dots \oplus V_{\nu-N\alpha}$ for some integer N, and we define k by $\nu = \mu + k\alpha$. The space V_ν spans a subrepresentation isomorphic to $\dim(V_\nu) = n_{\mu+k\alpha}(\Gamma_\lambda)$ many copies of the irreducible representation with highest weight N. The intersection of V_μ with this subrepresentation also has dimension $n_{\mu+k\alpha}(\Gamma_\lambda)$. From Proposition 3.4 we know that for elements of a standard basis we have $F \cdot E \cdot v_k = k(N-k+1)v_k$. If we split off the irreducible copies generated by V_ν, then the highest weight is $\nu - \alpha$ with multiplicity $n_{\mu+(k-1)\alpha}(\Gamma_\lambda) - n_{\mu+k\alpha}(\Gamma_\lambda)$, and this corresponds to $N' = N-2$ and $k' = k-1$. Here the standard generators contribute an eigenvalue $k'(N'-k'+1)$. Collecting the contribution with a factor $n_{\mu+k\alpha}(\Gamma_\lambda)$, we obtain (still for standard generators) $k(N-k+1) - (k-1)(N-k) = N$. But the standard generators correspond to $B(E,F) = \frac{2}{\|\alpha\|^2}$ whence we see that the contribution for $F_\alpha E_\alpha$ equals $\frac{\|\alpha\|^2}{2} N n_{\mu+k\alpha}(\Gamma_\lambda)$. But the highest weight N may be computed explicitly as $N = \frac{2\langle\beta,\alpha\rangle}{\|\alpha\|^2}$ and $\beta = \mu + k\alpha$. Working down taking off irreducible components step by step, this directly leads to the Freudenthal formula.

There is another formula for the multiplicities of weights due to Kostant, which has a different flavor. The advantage of this formula is that it is not of recursive character but directly computes the multiplicity of a weight. The disadvantage is that it needs summation over the whole Weyl group and involves a counting function, whose values are sometimes difficult to determine: Define the *Kostant partition function* $\mathcal{P} : \mathfrak{h}_0^* \to \mathbb{N}$ as follows. Put $\mathcal{P}(0) = 1$ and for $\mu \neq 0$ let $\mathcal{P}(\mu)$ be the number of tuples (a_α) of non–negative integers such that $\mu = \sum_{\alpha\in\Delta^+} a_\alpha \alpha$. Hence $\mathcal{P}(\mu)$ is the number of different ways to write μ as a sum of positive roots. In particular, $\mathcal{P}(\mu) = 0$ unless μ is positive. Denoting as before by δ the lowest form of $\mathfrak{g}$ the *Kostant multiplicity formula* then states that

$$n_\mu(\Gamma_\lambda) = \sum_{w\in W} \operatorname{sgn}(w)\mathcal{P}(w(\lambda+\delta)-(\mu+\delta)).$$

To use this formula efficiently, one usually has to show first that only few summands give a nonzero contribution. For example, in the case $\mathfrak{g} = \mathfrak{sl}(n,\mathbb{C})$, $\lambda = 2e_1 + e_2$ and $\mu = e_1+e_2+e_3 = \lambda-(\alpha_1+\alpha_2)$ that we have considered above, one can use elementary combinatorial arguments to show that only the summand with $w = \mathrm{id}$ can give a nonzero contribution. But then $\mathcal{P}(\alpha_1+\alpha_2) = 2$, corresponding to the decompositions $\alpha = \alpha$ and $\alpha = \alpha_1 + \alpha_2$, so we recover the multiplicity result from before.

5.4. The Weyl character formula and some consequences. For any finite dimensional representation V of a semisimple Lie algebra $\mathfrak{g}$, one can encode the multiplicities of all weight spaces into one function called the character of V. To do this, one considers the group ring $\mathbb{Z}[\mathfrak{h}_0^*]$, which we may simply think of the set of all formal finite linear combinations of the form $\sum a_\mu e^\mu$ for elements e^μ which satisfy the product rule $e^\mu e^\nu = e^{\mu+\nu}$. Such linear combinations may be added in an obvious way and multiplied according to $(\sum a_\mu e^\mu)(\sum b_\nu e^\nu) = \sum_\lambda(\sum_{\mu+\nu=\lambda} a_\mu b_\nu)e^\lambda$. Alternatively, one may think about the elements of $\mathbb{Z}[\mathfrak{h}_0^*]$ as functions $\mathfrak{h}_0^* \to \mathbb{Z}$ which map all but finitely many elements of $\mathfrak{h}_0^*$ to zero and then the above operations are just the pointwise addition of functions and the *convolution* $(fg)(\nu) = \sum_{\lambda+\mu=\nu} f(\lambda)g(\mu)$.

Given a finite dimensional representation V of $\mathfrak{g}$ one then defines the *character* $\chi(V) \in \mathbb{Z}[\mathfrak{h}_0^*]$ of V by $\chi(V) = \sum_{\lambda\in \mathrm{wt}(V)} \dim(V_\lambda)e^\lambda$. The advantage of this point of view is that it is very nicely compatible with the basic operations on representations. From the definition one immediately concludes that $\chi(V\oplus W) = \chi(V)+\chi(W)$ and $\chi(V\otimes W) = \chi(V)\chi(W)$.

For a weight $\mu\in\mathfrak{h}_0^*$, we define $A_\mu \in \mathbb{Z}[\mathfrak{h}_0^*]$ by $A_\mu := \sum_{w\in W}\operatorname{sgn}(w)e^{w(\mu)}$. The simplest way to state the *Weyl character formula* is now that for any dominant integral weight λ, the character of the irreducible representation Γ_λ with highest weight λ satisfies

$$A_\delta\chi(\Gamma_\lambda) = A_{\lambda+\delta},$$

where as before δ denotes the lowest form of $\mathfrak{g}$.

In this form it is not clear how much this formula can actually tell us about the character $\chi(\Gamma_\lambda)$. However, it turns out that one may enlarge the commutative ring $\mathbb{Z}[\mathfrak{h}_0^*]$ in such a way that A_δ becomes invertible, so in that larger commutative ring one has the expression $\chi(\Gamma_\lambda) = \frac{A_{\lambda+\delta}}{A_\delta}$. The price for this however is that $\frac{1}{A_\delta}$ is an infinite series, so handling the formula is fairly complicated and usually involves quite a bit of combinatorics.

Weyl's original proof of the character formula is done on the level of Lie groups. It uses that any complex semisimple Lie algebra possesses a unique (up to isomorphism) real form such that any connected Lie group with that Lie algebra is compact. The problem is then reduced to this compact group and the proof is done using analysis, see [**Fulton-Harris**, §26.2] for a sketch of this proof. There are also purely algebraic proofs. In particular, one may deduce the formula from the Freudenthal multiplicity formula, see [**Fulton-Harris**, §25.2]. The Weyl character formula in turn easily implies (and actually is equivalent to) the Kostant multiplicity formula. The key to this is that one may obtain an explicit formula for $\frac{1}{A_\delta}$ which has the form $e^{-\delta}\sum\mathcal{P}(\mu)e^{-\mu}$, where the sum goes over all μ such that $\mathcal{P}(\mu)\neq 0$.

A very useful consequence of the character formula is the *Weyl dimension formula* which computes the dimension of Γ_λ. Clearly, for $\chi(\Gamma_\lambda) = \sum a_\mu e^\mu$ we have $\dim(\Gamma_\lambda) =$

$\sum a_\mu$. Unfortunately, trying to express this sum from the character formula naively one ends up with $\frac{0}{0}$, but it is not too difficult to deduce the formula

$$\dim(\Gamma_\lambda) = \prod_{\alpha\in\Delta^+} \frac{\langle \lambda+\delta,\alpha\rangle}{\langle \delta,\alpha\rangle},$$

where the inner product is induced by (a multiple of) the Killing form.

Let us use the dimension formula to prove irreducibility of the tracefree symmetric powers of the standard representation $\mathbb{C}^{2n}$ of $\mathfrak{g} = \mathfrak{so}(2n,\mathbb{C})$. So we have to compute $\dim(\Gamma_\lambda)$ for $\lambda = ke_1$. From the description of the fundamental weights in 4.8 one immediately reads off that $\delta = \omega_1 + \dots + \omega_n = (n-1)e_1 + (n-2)e_2 + \dots + e_{n-1}$. From 3.8 we know that $\Delta^+ = \{e_i \pm e_j : i < j\}$. Clearly, a positive root α contributes non–trivially to the product in the dimension formula only if $\langle \lambda,\alpha\rangle \neq 0$, so we only have to consider the roots $e_1 \pm e_j$ for $j = 2,\dots,n$. For the roots $e_1 - e_j$ the contribution to the dimension formula is

$$\prod_{j=2}^{n} \frac{(n+k-1)-(n-j)}{(n-1)-(n-j)} = \prod_{j=2}^{n} \frac{k+j-1}{j-1} = \frac{(n+k-1)!}{k!(n-1)!}.$$

For the roots $e_1 + e_j$, we obtain the contribution

$$\prod_{j=2}^{n} \frac{(n+k-1)+(n-j)}{(n-1)+(n-j)} = \prod_{j=2}^{n} \frac{2n+k-j-1}{2n-j-1} = \frac{(2n+k-3)!(n-2)!}{(n+k-2)!(2n-3)!}.$$

Multiplying up and canceling we get

$$\dim(\Gamma_{ke_1}) = \frac{(2n+k-3)!(n+k-1)}{k!(2n-3)!(n-1)} = \frac{n+k-1}{n-1}\binom{2n+k-3}{2n-3}$$

We have to show that this coincides with $\dim(S^k\mathbb{C}^{2n}) - \dim(S^{k-2}\mathbb{C}^{2n})$. Inserting the definitions, one easily sees that

$$\binom{2n+k-1}{2n-1} - \binom{2n+k-3}{2n-1} = \frac{(2n+k-1)(2n+k-2)-k(k-1)}{(2n-1)(2n-2)}\binom{2n+k-3}{2n-3},$$

and one immediately verifies that this coincides with the above expression.

5.5. Decomposing tensor products. As an example for further results for dealing with finite dimensional representations we discuss the problem of determining multiplicities of irreducible factors and decomposing tensor products of irreducible representations. We start with two dominant integral weights λ and μ for a complex semisimple Lie algebra $\mathfrak{g}$, consider the irreducible representations Γ_λ and Γ_μ with these highest weights and their tensor product $\Gamma_\lambda \otimes \Gamma_\mu$. In particular, we are interested in finding the dominant integral weights ν for which the multiplicity $m_\nu(\Gamma_\lambda \otimes \Gamma_\mu)$ is positive and we want to compute this multiplicities.

Suppose that we have solved this problem and that V and W are two representations such that for each weight ν we know the multiplicities $m_\nu(V)$ and $m_\nu(W)$. Then $V \cong \oplus_\lambda \Gamma_\lambda^{m_\lambda(V)}$ and $W \cong \oplus_\mu \Gamma_\mu^{m_\mu(W)}$, and compatibility of the tensor product with direct sums immediately implies that $V \otimes W \cong \oplus_{\lambda,\mu}(\Gamma_\lambda \otimes \Gamma_\mu)^{m_\lambda(V)m_\mu(W)}$. In particular, this implies $m_\nu(V\otimes W) = \sum_{\lambda,\mu} m_\lambda(V)m_\mu(W)m_\nu(\Gamma_\lambda\otimes\Gamma_\mu)$. Hence we can determine the multiplicities of a tensor product (which determine it up to isomorphism) provided that we know the multiplicities of the factors.

We start with an elementary result:

Proposition. *If ν is a dominant integral weight such that $m_\nu(\Gamma_\lambda \otimes \Gamma_\mu) > 0$, then there is a weight μ' of Γ_μ such that $\nu = \lambda + \mu'$. Moreover, if this is the case, then $m_\nu(\Gamma_\lambda \otimes \Gamma_\mu) \leq n_{\lambda-\nu}(\Gamma_\mu)$.*

Proof. (1) Take bases of Γ_λ and Γ_μ consisting of weight vectors and consider the induced basis of $\Gamma_\lambda \otimes \Gamma_\mu$. Then any weight vector in the tensor product of weight ν may be written as a finite sum $\sum_{i=1}^n v_i \otimes w_i$ for weight vectors $v_i \in \Gamma_\lambda$ and $w_i \in \Gamma_\mu$ whose weights add up to ν. By adding up the left components which have the same w_i, we may assume that the w_i are linearly independent and then we renumber the factors in the sum in such a way that the weights λ_i of v_i satisfy $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Now suppose that an expression $\sum v_i \otimes w_i$ of this form is a highest weight vector. Then any element E_α in a positive root space annihilates the sum, which implies that we can write $(E_\alpha \cdot v_1) \otimes w_1$ as a linear combination of elements of the form $v_i \otimes (E_\alpha \cdot w_i)$ for $i \geq 1$ and of the form $(E_\alpha \cdot v_i) \otimes w_i$ for $i \geq 2$. Now let $\varphi : \Gamma_\mu \to \mathbb{C}$ be the linear functional which maps w_1 to 1 and vanishes on all other elements of the chosen basis. Applying $(\mathrm{id} \otimes \varphi) : \Gamma_\lambda \otimes \Gamma_\mu \to \Gamma_\lambda$ to the above linear combination, all the terms $(E_\alpha \cdot v_i) \otimes w_i$ for $i \geq 2$ are killed, and we conclude that $E_\alpha \cdot v_1$ is a linear combination of $v_1, \dots, v_n$. But by construction $E_\alpha \cdot v_1$ is a weight vector of weight $\lambda_1 + \alpha > \lambda_1 \geq \cdots \geq \lambda_n$, so this is only possible if $E_\alpha \cdot v_1 = 0$. Since this holds for all positive roots, we conclude that v_1 is a highest weight vector, so $\lambda_1 = \lambda$ and thus ν is the sum of λ and the weight of w_1.

Let $v_0 \in \Gamma_\lambda$ be a highest weight vector. Then we know that v_0 is unique up to complex multiples, so we see from above that any highest weight vector of weight ν can be written as a sum of $v_0 \otimes w$ with $0 \neq w \in (\Gamma_\mu)_{\lambda-\nu}$ and elements of the form $v_i \otimes w_i$ where each v_i is a weight vector of weight $< \lambda$. Of course, any non–trivial linear combination of highest weight vectors is again a highest weight vector. Now if there were more than $n_{\lambda-\nu}(\Gamma_\mu) = \dim((\Gamma_\mu)_{\lambda-\nu})$ many linearly independent highest weight vectors of weight ν, then the second components in their "leading terms" $v_0 \otimes w_j$ are linearly dependent, so we could find a nontrivial linear combination in which the leading terms cancel, which is a contradiction. □

This result is particularly useful if one of the two factors in the tensor product is simple. For example, consider the case $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C})$ and the standard representation $V = \mathbb{C}^n = \Gamma_{e_1}$. Then the weights of V are simply $e_1, \dots, e_n$ and each weight space has dimension one. Thus we conclude that $m_\nu(\Gamma_\lambda \otimes V) > 0$ is only possible for $\nu = \lambda + e_i$ for some $i = 1, \dots, n$ and in any case $m_\nu(\Gamma_\lambda \otimes V) \leq 1$. In particular, $\Gamma_\lambda \otimes V$ always decomposes into a direct sum of pairwise non–isomorphic irreducible representations. It is also easy to see in this case that the splitting into irreducibles coincides with the splitting into eigenspaces for the Casimir element introduced in 5.3. Indeed, the eigenvalue on the isotypical component with highest weight $\lambda + e_i$ is by Proposition 5.3 given as

$$\|\lambda + e_i\|^2 + 2\langle \lambda + e_i, \delta\rangle = \langle \lambda, \lambda + 2\delta\rangle + \|e_i\|^2 + 2\langle e_i, \lambda + \delta\rangle.$$

If for $i < j$ we would get the same eigenvalue for $\lambda + e_i$ and $\lambda + e_j$, then since all e_ℓ have the same length, this would mean that $0 = 2\langle e_i - e_j, \lambda + \delta\rangle$. But this is impossible, since $e_i - e_j \in \Delta^+$ and $\lambda + \delta$ lies in the interior of the positive Weyl chamber, so the inner product has to be positive.

Let us also note that the above result gives us yet another way to sort out the problem on the isotypical decomposition of $\otimes^3 V$. Interpreting $\otimes^3 V$ as $(\otimes^2 V) \otimes V = (S^2 V \otimes V) \oplus (\Lambda^2 V \otimes V)$, we see from above that for any dominant integral weight ν we have $m_\nu(S^2 V \otimes V) \leq 1$ and $m_\nu(\Lambda^2 V \otimes V) \leq 1$, and thus $m_\nu(\otimes^3 V) \leq 2$. Thus, $\otimes^3 V$ can contain at most 2 copies of $\Lambda^3 V$, but we had noted already that there have to be 1, 3, or 5 copies, so 1 is the only possibility.

Finally we just remark that there are several explicit formulae for $m_\nu(\Gamma_\lambda \otimes \Gamma_\mu)$. These usually are either of recursive nature or involve summation over the Weyl group and can be derived from the Weyl character formula. A particularly useful example of such a formula, which is valid for arbitrary complex semisimple Lie algebras is called *Klymik's formula* or *Racah's formula*:

$$m_\nu(\Gamma_\lambda \otimes \Gamma_\mu) = \sum_{w \in W} \operatorname{sgn}(w) n_{\nu+\delta-w(\lambda+\delta)}(\Gamma_\mu).$$

To use this formula efficiently, the first step usually again is to show that only few Weyl group elements lead to a non–trivial contribution to the sum.

Young symmetrizers and Weyl's construction

We conclude our discussion by outlining an explicit description of the irreducible representations of $\mathfrak{gl}(n, \mathbb{C})$ and $\mathfrak{sl}(n, \mathbb{C})$ based on the representation theory of permutation groups. This also leads to a complete description of the isotypical decomposition of the tensor powers of the standard representation, as well as constructions for general finite dimensional representations. Much of this can be generalized to the other classical simple Lie algebras. We start by discussing some general facts on representations of finite groups.

5.6. Representations of finite groups. We have already met the concept of a representation of a finite group in 1.2. The first fundamental observation is that finite dimensional representations of finite groups are always completely reducible: Suppose that we have given a representation of G on V, i.e. a homomorphism $G \to GL(V)$. Then we can construct a G–invariant inner product on V by averaging over any inner product: Choose any inner product $(\ ,\ )$ on V and define

$$\langle v, w \rangle := \sum_{g \in G} (g \cdot v, g \cdot w).$$

By construction, this is invariant, i.e. $\langle v, w \rangle = \langle g \cdot v, g \cdot w \rangle$ for each $g \in G$. In particular, if $W \subset V$ is an invariant subspace, then $W^\perp \subset V$ is an invariant complement to W, so complete reducibility follows.

Next, we observe that any finite group G admits a canonical finite dimensional representation, called the *left regular representation*. Consider the space $\mathbb{C}^G$ of complex valued functions on G. Of course, the dimension of this space equals the number $|G|$ of elements of G. From 1.2 we see that the left action of G on itself by multiplication gives rise to a representation of G on $\mathbb{C}^G$, defined by $(g \cdot \varphi)(g') = \varphi(g^{-1} g')$. (Of course, an analog of this representation exists also for Lie groups, but it is infinite dimensional and thus much more complicated to deal with, in particular for non–compact groups.) Let us denote by $e_g \in \mathbb{C}^G$ for $g \in G$ the function given by $e_g(g) = 1$ and $e_g(h) = 0$ for all $h \neq g$. Then any $\varphi \in \mathbb{C}^G$ can be written as $\sum_{g \in G} a_g e_g$ for some numbers $a_g \in \mathbb{C}$. In this picture, the left regular representation is characterized by $g \cdot e_h = e_{gh}$.

Similarly as in 5.4 we can now make $\mathbb{C}^G$ into an associative algebra by defining $e_g * e_h := e_{gh}$ for all $g, h \in G$ and extending this bilinearly. In the picture of functions, this corresponds to the convolution $(\varphi * \psi)(g) = \sum_h \varphi(h)\psi(h^{-1}g)$. This associative algebra is called the complex *group algebra* of G and denoted by $\mathbb{C}[G]$. Note the $g \mapsto e_g$ defines an injective homomorphism from G into the group of invertible elements of $\mathbb{C}[G]$, and under this homomorphism the left regular representation corresponds to multiplication from the left in $\mathbb{C}[G]$. If A is any associative algebra and $f : G \to A$ is a homomorphism from G to the group of invertible elements of A, then this uniquely extends to a homomorphism $\tilde{f} : \mathbb{C}[G] \to A$ by $\tilde{f}(\sum a_g e_g) = \sum a_g f(g)$. Hence we may think of $\mathbb{C}[G]$ as an analog of the universal enveloping algebra (or the other way round). In particular, complex representations of G uniquely extend to $\mathbb{C}[G]$.

Using the theory of characters, one shows that (up to isomorphism) there are only finitely many irreducible representations $V_1, \dots, V_n$ of G (and n equals the number of conjugacy classes of G). Now for any i we obtain from above an algebra homomorphism $\mathbb{C}[G] \to L(V_i, V_i)$. Under this homomorphism, the regular representation corresponds to action on the values of maps, so as a G–representation we have to consider the first copy of V_i in $L(V_i, V_i)$ as the direct sum of $\dim(V_i)$ copies of the trivial representation, and thus $L(V_i, V_i)$ as $V_i^{\dim(V_i)}$. Again using characters, one shows that these homomorphisms induce an isomorphism $\mathbb{C}[G] \cong \oplus_{i=1}^n V_i^{\dim(V_i)}$ of G–representations, so this is the analog of the isotypical decomposition for the regular representation. For details about this see chapters 2 and 3 of [**Fulton-Harris**].

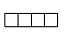
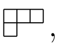
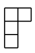
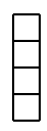
Thus one can find any irreducible representation of G inside the group algebra $A := \mathbb{C}[G]$. Now there is a simple way to construct G–invariant subspaces of A. Namely, take any element $c \in A$ and consider the subspace $Ac := \{ac : a \in A\}$. Of course, this is invariant under left multiplications of elements of A, and thus in particular a G–invariant subspace of the left regular representation.

5.7. From permutation groups to $\mathfrak{gl}(V)$. As above, let G be a finite group with group algebra $A = \mathbb{C}[G]$, and suppose that V is a finite dimensional representation of G. Then we define a right action $V \times G \to V$ of G on V by $v \cdot g := g^{-1} \cdot v$. This evidently satisfies $v \cdot (gh) = (v \cdot g) \cdot h$, and we extend it by linearity to a map $V \times A \to V$, which then satisfies $v \cdot (ab) = (v \cdot a) \cdot b$. Given an element $c \in A$, we can then form $Vc := \{v \cdot c : v \in V\}$, which obviously is a linear subspace of V.

Next, consider the vector space $\mathcal{B} := \operatorname{Hom}_G(V, V)$ of G–equivariant maps $\varphi : V \to V$. Of course, for $\varphi \in \mathcal{B}$ we have $\varphi(v \cdot g) = \varphi(v) \cdot g$ for all $g \in G$, and thus $\varphi(v \cdot a) = \varphi(v) \cdot a$ for all $a \in A$. In particular, $\varphi(Vc) \subset Vc$ for all $\varphi \in \mathcal{B}$ and $c \in A$. Composition makes $\mathcal{B}$ into an associative algebra, this has a canonical representation on V, and for each $c \in A$, the subspace Vc is invariant.

Let W be another representation of A (viewed as usual as a left module). Then we define $V \otimes_A W$ to be the quotient of the tensor product $V \otimes W$ by the linear subspace generated by all elements of the form $(v \cdot a) \otimes w - v \otimes (a \cdot w)$ with $v \in V$, $w \in W$ and $a \in A$. For any $\varphi \in \mathcal{B}$ we can consider the map $\varphi \otimes \mathrm{id}_W : V \otimes W \to V \otimes W$. One immediately verifies that this preserves the subspace defined above, so it descends to a linear map $V \otimes_A W \to V \otimes_A W$ hence making this space into a representation of $\mathcal{B}$. The fundamental fact for our purposes is that if W is an irreducible A–module then $V \otimes_A W$ is an irreducible $\mathcal{B}$–module. Moreover, for $c \in A$ consider $W = Ac$ and the map $V \otimes Ac \to Vc$ defined by $v \otimes ac \mapsto v \cdot (ac) = (v \cdot a) \cdot c$. Visibly, this descends to a linear map $V \otimes_A Ac \to Vc$, which is easily seen to be an isomorphism of $\mathcal{B}$–modules. Proofs of these two facts can be found in [**Fulton-Harris**, Lemma 6.22].

To pass to Lie algebras, we consider the case of a permutation group $G = \mathfrak{S}_k$, an arbitrary vector space V and the action of G on $\otimes^k V$ given by permutation of the factors. Using the maps f_σ defined by $f_\sigma(v_1 \otimes \dots \otimes v_k) = v_{\sigma_1} \otimes \dots \otimes v_{\sigma_k}$ as in 4.5 this reads as $x \cdot \sigma := f_{\sigma^{-1}}(x)$ for $x \in \otimes^k V$. The obvious action of $L(V, V)$ on $\otimes^k V$ is clearly by G–equivariant maps, and one verifies that any G–equivariant endomorphism of $\otimes^k V$ is of this form. Hence we have $L(V, V) \cong \operatorname{Hom}_G(\otimes^k V, \otimes^k V)$. In particular, from above we see that if $c \in A = \mathbb{C}[\mathfrak{S}_k]$ is an element such that Ac is an irreducible representation of $\mathfrak{S}_k$, then $(\otimes^k V)c \cong (\otimes^k V) \otimes_A Ac$ is an irreducible representation of $L(V,V)$ and thus of $\mathfrak{gl}(V)$.

5.8. Irreducible representations of $\mathfrak{S}_k$. It turns out that the irreducible representations of $\mathfrak{S}_k$ are indexed by all partitions of k. (This is not so surprising, since from above we know that irreducible representations are parametrized by conjugacy classes which are described by partitions via the length of cycles in a decomposition of a permutation into a product of cycles.) A *partition* γ of k is an ℓ–tuple $(\gamma_1, \dots, \gamma_\ell)$ of integers such that $\gamma_1 \geq \gamma_2 \geq \dots \geq \gamma_\ell > 0$ and such that $\gamma_1 + \dots + \gamma_\ell = k$. For example, $k = 3$ admits the partitions (3), $(2,1)$ and $(1,1,1)$, and $k = 4$ has the partitions (4), $(3,1)$, $(2,2)$, $(2,1,1)$, and $(1,1,1,1)$. To a partition $\gamma = (\gamma_1, \dots, \gamma_\ell)$ one associates a *Young diagram* as an array of ℓ lines of boxes such that the ith line consists of γ_i boxes. In this way, the partitions of $k = 4$ listed above are represented by the Young diagrams , , , , and .

To each partition γ of k one next associates an element $c_\gamma \in A := \mathbb{C}[\mathfrak{S}_k]$ as follows: Take the Young diagram of γ and number the boxes by the numbers $\{1, \dots, k\}$ row by row. Let $P \subset \mathfrak{S}_k$ be the set of those permutations of $\{1, \dots, k\}$ which leave each row of the diagram invariant, and $Q \subset \mathfrak{S}_k$ as those which leave each column invariant. By construction, P and Q are subgroups of $\mathfrak{S}_k$. Then define $a_\gamma, b_\gamma, c_\gamma \in A$ by $a_\gamma := \sum_{\sigma \in P} e_\sigma \in A$, $b_\gamma := \sum_{\sigma \in Q} \operatorname{sgn}(\sigma) e_\sigma$, and $c_\gamma := a_\gamma b_\gamma$. The element $c_\gamma \in A$ is called the *Young symmetrizer* corresponding to the partition γ (respectively to its Young diagram). Consider the partition (k), so a Young diagram with one row only. Then by definition $P = \mathfrak{S}_k$ and $Q = \{\mathrm{id}\}$, so $b_\gamma = e_{\mathrm{id}}$, the unit element of A and $a_\gamma = c_\gamma$ acts on $\otimes^k V$ as $k!$ times the symmetrization. Similarly, for the partition $(1, \dots, 1)$ we obtain $k!$ times the alternation as the action on $\otimes^k V$. More generally, for $\gamma = (\gamma_1, \dots, \gamma_\ell)$ we get $(\otimes^k V)a_\gamma = S^{\gamma_1} V \otimes \dots \otimes S^{\gamma_\ell} V$ while for the action of b_γ the result is isomorphic to a tensor product of γ_1 exterior powers of V whose degrees are determined by the numbers of boxes in the columns of the Young diagram.

To formulate the description of the irreducible representations of $\mathfrak{S}_k$, we need one more definition. Consider a partition $\gamma = (\gamma_1, \dots, \gamma_\ell)$ of k and the corresponding Young diagram. For a box in the Young diagram, define the *Hook length* of the box to be the number of boxes which lie below the given one in the same column, plus the number of boxes which lie right of the given one in the same row plus one. Otherwise put, given a box, draw a hook with edge in the box and legs pointing down and right, and count how many boxes the hook meets. Then define $d(\gamma)$ to be the quotient of $k!$ by the product of the hook lengths of all boxes. For example for the two trivial partitions $k = (k)$ and $k = (1, \dots, 1)$ there is just one row respectively just one column in the diagram, and hence the hook lengths of the boxes are just $k, k-1, \dots, 1$, and hence in both cases we get $d(\gamma) = 1$. As another example, consider the partition $(3,1)$ of $k = 4$. Then the Hook lengths of the four boxes are just 4, 2, 1, and 1, so $d(\gamma) = 3$ in this case.

Theorem. *Let γ be a partition of k, and let $c_\gamma \in A := \mathbb{C}[\mathfrak{S}_k]$ be the corresponding Young symmetrizer. Then c_γ is a multiple of a projection, i.e. $(c_\gamma)^2$ is a nonzero multiple of c_γ, and $V_\gamma := Ac_\gamma$ is an irreducible representation of $\mathfrak{S}_k$ of dimension $d(\gamma)$. The map $\gamma \mapsto V_\gamma$ defines a bijection from the set of partitions of k to the set of isomorphism classes of irreducible representations of $\mathfrak{S}_k$.*

The proof of this result can be found in [**Fulton-Harris**, §4.2–4.3]. It is not difficult and elementary apart from some basic results on character theory which are needed to show that the representations V_γ exhaust all irreducible representations of $\mathfrak{S}_k$.

5.9. Schur functors and Weyl's construction. Let V be a complex vector space and consider the kth tensor power $\otimes^k V$. From 5.7 we have a right action of $\mathfrak{S}_k$ on $\otimes^k V$ given by $(v_1 \otimes \cdots \otimes v_k) \cdot \sigma := v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(k)}$ which commutes with the action of $\mathfrak{gl}(V)$. This action extends to an action of the group algebra $A = \mathbb{C}[\mathfrak{S}_k]$, which still commutes with the action of $\mathfrak{gl}(V)$. Given a partition γ of k with corresponding Young symmetrizer $c_\gamma \in A$ we can form $\mathbb{S}_\gamma(V) := (\otimes^k V)c_\gamma \cong (\otimes^k V) \otimes_A Ac_\gamma$. From 5.8 we know that Ac_γ is an irreducible representation of $\mathfrak{S}_k$, so we conclude from 5.7 that $\mathbb{S}_\gamma(V)$ is an irreducible representation of $\mathfrak{gl}(V)$. If $f : V \to W$ is a linear map, then we have the induced map $\otimes^k f : \otimes^k V \to \otimes^k W$, and of course this equivariant for the actions of A on both components, so it restricts to a linear maps $\mathbb{S}_\gamma(f) : \mathbb{S}_\gamma(V) \to \mathbb{S}_\gamma(W)$. From the construction it follows that $\mathbb{S}_\gamma(g \circ f) = \mathbb{S}_\gamma(g) \circ \mathbb{S}_\gamma(f)$. Therefore, $\mathbb{S}_\gamma$ defines a functor on the category of finite dimensional vector spaces, called the *Schur functor* associated to the partition γ. This construction of representations is also referred to as *Weyl's construction.*

Note that for the obvious partitions (k) and $(1, \dots, 1)$ we recover the symmetric power $S^k V$ respectively the exterior power $\Lambda^k V$. The second example shows that it may well happen that $\mathbb{S}_\gamma(V) = \{0\}$. Indeed, if $\gamma = (\gamma_1, \dots, \gamma_\ell)$, then from the construction of c_γ it follows immediately that applying the Schur functor involves alternating over ℓ components, so the result must be zero if $\ell > \dim(V)$. It turns out that this is the only restriction, so $\mathbb{S}_\gamma(V) \neq \{0\}$ if $\ell \leq \dim(V)$.

This machinery also immediately leads to a complete description of $\otimes^k V$ as a representation of $\mathfrak{gl}(V)$. The description of the decomposition of the left regular representation of $\mathfrak{S}_k$ from 5.6 can be stated as $A \cong \oplus_\gamma (Ac_\gamma)^{d(\gamma)}$, where the sum is over all partitions γ and $d(\gamma) = \dim(Ac_\gamma)$ is the number introduced in 5.8. It is obvious from the definition in 5.7 that $\otimes^k V \cong (\otimes^k V) \otimes_A A$, so this decomposes as $\oplus_\gamma (\otimes^k V) \otimes_A (Ac_\gamma)^{d(\gamma)}$, and we obtain

$$\otimes^k V \cong \oplus_\gamma \mathbb{S}_\gamma(V)^{d(\gamma)}$$

as a representation of $\mathfrak{gl}(V)$.

Taking $V = \mathbb{C}^n$, we may restrict the representation $\mathbb{S}_\gamma(V)$ to $\mathfrak{g} := \mathfrak{sl}(n, \mathbb{C}) \subset \mathfrak{gl}(n, \mathbb{C})$. Since the element $t\,\mathrm{id} \in \mathfrak{gl}(n, \mathbb{C})$ acts on $\otimes^k \mathbb{C}^n$ by multiplication by t^k, it acts in the same way on $\mathbb{S}_\gamma(\mathbb{C}^n)$. Hence any $\mathfrak{g}$–invariant subspace of $\mathbb{S}_\gamma(\mathbb{C}^n)$ is automatically $\mathfrak{gl}(n, \mathbb{C})$–invariant, which implies that $\mathbb{S}_\gamma(\mathbb{C}^n)$ is an irreducible representation of $\mathfrak{g}$. From above we know that we obtain a nontrivial result only if $\gamma = (\gamma_1, \dots, \gamma_\ell)$ with $\ell \leq n$. One verifies that the highest weight of the representation $\mathbb{S}_\gamma(\mathbb{C}^n)$ is then given by $\gamma_1 e_1 + \cdots + \gamma_\ell e_\ell$. In particular, $\mathbb{S}_\gamma(\mathbb{C}^n) \cong \mathbb{S}_{\tilde{\gamma}}(\mathbb{C}^n)$ if and only if there is some integer j such that $\tilde{\gamma}_i = \gamma_i + j$ for all $i = 1, \dots, n$. To get a unique description of the irreducible representations, one may therefore simply restrict to partitions $\gamma = (\gamma_1, \dots, \gamma_\ell)$ with $\ell < n$. Also, the decomposition of $\otimes^k \mathbb{C}^n$ from above is exactly the decomposition into isotypical components, if one adds up isomorphic pieces.

There is a lot of further information on these representations available, for example combinatorial formulae for the dimension of $\mathbb{S}_\gamma(\mathbb{C}^n)$ which simplify the Weyl dimension formula from 5.4, and for the isotypical decomposition of a tensor product of two such representations. Information on these and further results in this direction can be found at various places in [**Fulton-Harris**].

There is a simple analog of Weyl's construction for the other classical Lie algebras. In all cases, there is a bilinear form b, which is invariant under the action of the given Lie algebra $\mathfrak{g}$. Given the standard representation V of $\mathfrak{g}$ and considering $\otimes^k V$, any choice of indices i and j with $1 \leq i < j \leq k$ gives rise to a $\mathfrak{g}$–homomorphism $\otimes^k V \to \otimes^{k-2} V$. Now one has to take the intersection of the kernels of these maps and intersect it with $\mathbb{S}_\gamma(V)$ to obtain an irreducible representation of $\mathfrak{g}$. It turns out that for $\mathfrak{g} = \mathfrak{sp}(2n, \mathbb{C})$, all irreducible representations can be obtained in this way, and one can also describe the isotypical decomposition of $\otimes^k V$ explicitly, see [**Fulton-Harris**, §17.3]. For the orthogonal algebras $\mathfrak{so}(n, \mathbb{C})$, it is clear that one cannot obtain spin representations. It turns out however, that all vector representations can be obtained by this construction. This means that in the odd case, one gets all irreducible representations whose highest weight has the last component in the expansion as a linear combination of fundamental weights is even. In the even case, the corresponding condition is that the sum of the coefficients of the last two fundamental weights has to be even. Details about this can be found in [**Fulton-Harris**, §19.5].

Bibliography

[1] *Lectures on Lie Groups and Lie Algebras*, London Mathematical Society Student Texts 32, Cambridge University Press, 1995.

[Fulton-Harris] W. Fulton, J. Harris, *Representation Theory A First Course*, Graduate Texts in Mathematics 129, Springer, 1991.

[2] J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory* Grad. Texts in Math. Vol. 9, Springer 1972

[Knapp] A.W. Knapp, *Lie Groups Beyond an Introduction*, Progress in Mathematics 140, Birkhäuser, 1996.

[Kolař-Michor-Slovák] I. Kolař, P.W. Michor, J. Slovák, *Invariant Operations in Differential Geometry*, Springer, 1993

[Michor] P. Michor, *Topics in Differential Geometry*, lecture notes, available online at http://www.mat.univie.ac.at/∼michor/tgbook.ps

[3] H. Samelson, *Notes on Lie algebras* Van Nostrand Reinhold Mathematical Studies, No. 23 Van Nostrand Reinhold Co., New York- London-Melbourne 1969

Index