Math 429 - Representation Theory II

Lie groups and algebras

New concepts will be written in **bold**, and new formulas will be $\boxed{\text{boxed}}$.

Material which you have already encountered in Math 211 and 314 will be marked as such.

Details in the proofs that we purposely leave out of the notes, so that you may work out for yourselves, will be colored in blue. Ask your instructors (in person / on the forum) for help.

Table of contents:

Lecture 1

1.1

A Lie group is a group that happens to also be a smooth manifold, with the two structures being compatible as in Definition 1. The main motivational example is the general linear group GL_n of invertible $n \times n$ matrices, which can be regarded as a subset of n^2 dimensional space. Before we give the precise definition, recall that a **group** G is a set endowed with

- an identity element $e \in G$,
- an involution $G \to G,\ g \mapsto g^{-1}$,
- an operation $G \times G \to G,\ (g, h) \mapsto gh$ that satisfies associativity.

The structures above must be compatible in the usual ways, that you recall from Math 211.

Meanwhile, recall that a topological space G is called a **manifold** (of dimension N) if it can be covered by open subsets (called **charts**) homeomorphic to open balls in $\mathbb{R}^N$, such that the overlaps between charts correspond to C^∞ (infinitely differentiable, a.k.a. "smooth") functions

$$\Big(\text{an open subset of } \mathbb{R}^N\Big) \to \Big(\text{an open subset of } \mathbb{R}^N\Big)$$

The charts allow us to apply much of the usual machinery of calculus to manifolds. For example, we have a notion of tangent space at any point $g \in G$

$$T_g G = \Big\{\text{vectors tangent to } G \text{ at } g\Big\} \tag{1}$$

At first glance, the definition above makes sense only when G lies inside $\mathbb{R}^N$ for some N, so that we can make sense of $T_g G$ in some "ambient" tangent space. But for an abstract manifold G, there is an equivalent definition of tangent spaces via **derivations**, which only takes as input the notion of smooth functions $G \to \mathbb{R}$ (as you would imagine, a function is called smooth if it corresponds to a smooth function (open subset of $\mathbb{R}^N$) $\to \mathbb{R}$ for every chart of G). Explicitly, a $\mathbb{R}$-linear map

$$v : \Big(\text{smooth functions on } G\Big) \to \mathbb{R} \tag{2}$$

is called a derivation at $g \in G$ if it satisfies the Leibniz rule

$$v(\alpha\beta) = v(\alpha) \cdot \beta(g) + \alpha(g) \cdot v(\beta) \tag{3}$$

for all smooth functions α, β on G. Then we define

$$T_g G = \Big\{\text{derivations at } g\Big\} \tag{4}$$

which is made into an $\mathbb{R}$-vector space by addition and scalar multiplication. As you would expect, a **smooth map** (i.e. a continuous function of the underlying topological spaces, which corresponds to a C^∞ function on every chart) $F : G \to G'$ between manifolds G, G' induces linear transformations

$$F_* : T_g G \to T_{g'} G' \tag{5}$$

for all $g \in G$ and $g' = F(g)$ (define (5) in terms of derivations). F as above is called

- an **immersion** if F_* is injective at every point $g \in G$;
- an **embedding** if it is both an immersion and a homeomorphism onto its image. In this case, we will often write

$$G \hookrightarrow G'$$

and call G a **submanifold** of G'.

Definition 1. *A* ***Lie group*** *is a set which has both a structure of group and of a smooth manifold, such that the inverse $G \to G$ and the operation $G \times G \to G$ are smooth functions.*

1.2

Beside the **general linear group** GL_n, here are some other important examples of Lie groups (please check for yourselves that these are indeed Lie groups):

- the **special linear group** $\boxed{SL_n = \left\{A \in \mathrm{Mat}_{n\times n} \middle| \det(A) = 1\right\}}$
- the **orthogonal group** $\boxed{O_n = \left\{A \in \mathrm{Mat}_{n\times n} \middle| A^T A = I_n\right\}}$
- the **special orthogonal group** $\boxed{SO_n = O_n \cap SL_n}$
- the **symplectic group** $\boxed{Sp_{2n} = \left\{A \in \mathrm{Mat}_{2n\times 2n} \middle| A^T \Omega A = \Omega\right\}}$, where

$$\Omega = \left(\begin{array}{c|c} 0 & I_n \\ \hline -I_n & 0 \end{array}\right)$$

- $\mathbb{R}^n$ with component-wise addition, and $(\mathbb{R}^*)^n$ with component-wise multiplication.

Notice that we've been intentionally vague about the coefficient field of our matrices in the examples above. This is because while manifolds are typically defined with respect to $\mathbb{R}$ in mind (so the corresponding matrix groups are $GL_n(\mathbb{R})$, $O_n(\mathbb{R})$, $Sp_{2n}(\mathbb{R})$ etc), one has an analogous theory with respect to $\mathbb{C}$. Here, one can do one of two things:

- Simply observe that $\mathbb{C}^N \cong \mathbb{R}^{2N}$, and so the matrix groups $GL_n(\mathbb{C})$, $O_n(\mathbb{C})$, $Sp_{2n}(\mathbb{C})$ etc are manifolds in the sense of Subsection 1.1; we will refer to them as **real manifolds** whenever there is a chance for confusion.
- Define **complex manifolds** as in Subsection 1.1, but requiring the charts to be homeomorphic to open subsets in $\mathbb{C}^N$, with the overlaps corresponding to holomorphic functions $\mathbb{C}^N \to \mathbb{C}^N$.

Any complex manifold (such as $GL_n(\mathbb{C})$, $O_n(\mathbb{C})$, $Sp_{2n}(\mathbb{C})$) is also a real manifold, but not vice versa. To see some more examples of this in the context of Lie groups, the **unitary group**

$$\boxed{U(n) = \left\{A \in \mathrm{Mat}_{n\times n}(\mathbb{C}) \middle| \bar{A}^T A = I_n\right\}} \tag{6}$$

and the **special unitary group**

$$\boxed{SU(n) = U(n) \cap SL_n(\mathbb{C})}$$

are real manifolds and not complex manifolds, despite the fact that they are defined as subsets of the $\mathbb{C}$ vector space of all complex matrices. For example, $U(1) = S^1 \subset \mathbb{C}^*$ with the operation given by rotation, and there is no reasonable sense in which a circle can be a complex manifold.

The world of real manifolds is richer in Lie groups than the world of complex manifolds. For instance, recall that the orthogonal group can be thought of as the set of linear transformations which preserves the Euclidean inner product. One can consider the generalized orthogonal groups, defined as the set of linear transformations which preserve a bilinear form of signature $(k, n-k)$:

$$O_{k,n-k} = \left\{ A \in \mathrm{Mat}_{n\times n} \middle| A^T \left(\begin{array}{c|c} I_k & 0 \\ \hline 0 & -I_{n-k} \end{array} \right) A = \left(\begin{array}{c|c} I_k & 0 \\ \hline 0 & -I_{n-k} \end{array} \right) \right\}$$

These groups are all non-isomorphic (except when $k \leftrightarrow n-k$) over the reals, but they are all isomorphic to O_n over the complex numbers, because -1 has a square root. The flip side of this is that the complex versions of Lie groups are in general better behaved than their real versions.

Remark. *We will use the terms "real Lie group" and "complex Lie group" to differentiate between Lie groups in the context of real manifolds and complex manifolds, respectively. Similarly, we will use the terms "smooth function" and "holomorphic function" to differentiate between the types of allowable functions in the two situations. We will write GL_n, SL_n, O_n etc for the corresponding groups with either real or complex coefficients.*

1.3

A **vector field** on a smooth manifold G will refer to a choice $\boldsymbol{v} = \{v_g \in T_g G\}_{g\in G}$ of tangent vectors that varies in a smooth fashion over the charts of G (we will not bother to make this precise, but if G is a submanifold of $\mathbb{R}^N$, the notion of smoothly varying vector field is precisely what you would intuitively expect). A vector field can be construed as a derivation

$$\boldsymbol{v} : \Big(\text{smooth functions on } G\Big) \to \Big(\text{smooth functions on } G\Big) \tag{7}$$

which satisfies the Leibniz rule in the following form

$$\boldsymbol{v}(\alpha\beta) = \boldsymbol{v}(\alpha) \cdot \beta + \alpha \cdot \boldsymbol{v}(\beta) \tag{8}$$

Example 1. *On $\mathbb{R}^N$, any tangent vector is a linear combination of the usual derivatives*

$$\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_N}$$

(followed by evaluation at the point where the tangent vector lives). A collection of tangent vectors

$$\left\{ \alpha_1(x_1, \dots, x_N) \frac{\partial}{\partial x_1} + \dots + \alpha_N(x_1, \dots, x_N) \frac{\partial}{\partial x_N} \right\}_{x_1,\dots,x_N \in \mathbb{R}}$$

determines a vector field iff the coefficients $\alpha_1, \dots, \alpha_N$ all depend smoothly on $x_1, \dots, x_N$.

If G is a Lie group, then we have left multiplication maps for each $g \in G$

$$G \xrightarrow{h \mapsto gh} G$$

which are smooth. Therefore, they induce linear isomorphisms between tangent spaces

$$T_h G \xrightarrow{\text{left multiplication by } g} T_{gh} G$$

Thus, we have a notion of **left invariant** vector field on G, namely a vector field which is preserved by the multiplication maps above. Since a left invariant vector field is completely determined by its value at the identity, we conclude that

$$\boxed{T_e G \cong \Big\{\text{left invariant vector fields on } G\Big\}} \tag{9}$$

1.4

It is customary to write $\text{Lie}(G) = T_e G$. Although a priori just a vector space, $\text{Lie}(G)$ can be endowed with an extra structure called a **Lie bracket**. The key result is the following.

Proposition 1. *If $\boldsymbol{v}$ and $\boldsymbol{w}$ are two vector fields (i.e. derivations (7)), then so is*

$$\boxed{[\boldsymbol{v}, \boldsymbol{w}]\,(\alpha) = \boldsymbol{v}(\boldsymbol{w}(\alpha)) - \boldsymbol{w}(\boldsymbol{v}(\alpha))} \tag{10}$$

Moreover, if $\boldsymbol{v}$ and $\boldsymbol{w}$ are left invariant vector fields on G, then so is $[\boldsymbol{v}, \boldsymbol{w}]$.

Proof. Explicitly, for any smooth functions α and β, we have

$$[\boldsymbol{v}, \boldsymbol{w}](\alpha\beta) = \boldsymbol{v}(\boldsymbol{w}(\alpha\beta)) - \boldsymbol{w}(\boldsymbol{v}(\alpha\beta)) = \boldsymbol{v}(\boldsymbol{w}(\alpha) \cdot \beta + \alpha \cdot \boldsymbol{w}(\beta)) - \boldsymbol{w}(\boldsymbol{v}(\alpha) \cdot \beta + \alpha \cdot \boldsymbol{v}(\beta)) =$$
$$= \boldsymbol{v}(\boldsymbol{w}(\alpha)) \cdot \beta + \boldsymbol{w}(\alpha) \cdot \boldsymbol{v}(\beta) + \boldsymbol{v}(\alpha) \cdot \boldsymbol{w}(\beta) + \alpha \cdot \boldsymbol{v}(\boldsymbol{w}(\beta)) - \boldsymbol{w}(\boldsymbol{v}(\alpha)) \cdot \beta - \boldsymbol{v}(\alpha) \cdot \boldsymbol{w}(\beta) - \boldsymbol{w}(\alpha) \cdot \boldsymbol{v}(\beta) -$$
$$- \alpha \cdot \boldsymbol{w}(\boldsymbol{v}(\beta)) = (\boldsymbol{v}(\boldsymbol{w}(\alpha)) - \boldsymbol{w}(\boldsymbol{v}(\alpha))) \cdot \beta + \alpha \cdot (\boldsymbol{v}(\boldsymbol{w}(\beta)) - \boldsymbol{w}(\boldsymbol{v}(\beta)))$$

Finally, if $\boldsymbol{v}$ and $\boldsymbol{w}$ are left invariant (i.e. preserved by the automorphisms of left multiplication), then so are the compositions $\boldsymbol{v} \circ \boldsymbol{w}$ and $\boldsymbol{w} \circ \boldsymbol{v}$ of these vector fields, hence so is $[\boldsymbol{v}, \boldsymbol{w}]$. □

The commutator of vector fields is actually a very explicit operation. In local coordinates $x_1, \dots, x_N$, we may write vector fields as

$$\boldsymbol{v} = \alpha_1 \frac{\partial}{\partial x_1} + \dots + \alpha_n \frac{\partial}{\partial x_n} \quad \text{and} \quad \boldsymbol{w} = \beta_1 \frac{\partial}{\partial x_1} + \dots + \beta_n \frac{\partial}{\partial x_n}$$

Then we have

$$[\boldsymbol{v}, \boldsymbol{w}] = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\alpha_i \frac{\partial \beta_j}{\partial x_i} - \beta_i \frac{\partial \alpha_j}{\partial x_i} \right) \frac{\partial}{\partial x_j} \tag{11}$$

It is easy to see that the operation (10) satisfies the following properties.

- anti-symmetry: $[\boldsymbol{v}, \boldsymbol{w}] = -[\boldsymbol{w}, \boldsymbol{v}]$, and

- the Jacobi identity $[\boldsymbol{v}_1, [\boldsymbol{v}_2, \boldsymbol{v}_3]] + [\boldsymbol{v}_2, [\boldsymbol{v}_3, \boldsymbol{v}_1]] + [\boldsymbol{v}_3, [\boldsymbol{v}_1, \boldsymbol{v}_2]] = 0$.

We conclude that the vector space $\text{Lie}(G)$ of (9) is endowed with an operation

$$\text{Lie}(G) \times \text{Lie}(G) \xrightarrow{[\cdot,\cdot]} \text{Lie}(G)$$

satisfying anti-symmetry and the Jacobi identity. If we regard $\text{Lie}(G)$ as the tangent space at $e \in G$, this operation is as follows: take any two tangent vectors, extend them uniquely to left-invariant vector fields on G, then take the commutator of the vector fields in question as per (10), and then restrict the corresponding vector field back to $e \in G$.

1.5

The following definition is an abstract version of the discussion in the previous Subsection, which actually makes sense over any ground field $\mathbb{K}$.

Definition 2. *A **Lie algebra** $\mathfrak{g}$ is a $\mathbb{K}$-vector space endowed with a **Lie bracket***

$$\mathfrak{g} \times \mathfrak{g} \xrightarrow{[\cdot,\cdot]} \mathfrak{g} \tag{12}$$

which is $\mathbb{K}$-bilinear in both arguments and satisfies

- *anti-symmetry: $[x, y] = -[y, x]$, and*
- *the Jacobi identity*
$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0 \tag{13}$$

***All Lie algebras studied in this course are finite-dimensional, and** char $\mathbb{K} = 0$.*

Thus, the discussion in the last paragraph of the previous Subsection shows that $\text{Lie}(G)$ has a Lie algebra structure, for all Lie groups G.

Example 2. *Let us work out the Lie algebra structure on*

$$\mathfrak{gl}_n = \text{Lie}(GL_n)$$

Because GL_n is an open subset of the vector space of all $n \times n$ matrices, it tangent spaces are all naturally identified with the vector space in question, so we have an identification

$$\mathfrak{gl}_n = \text{Mat}_{n\times n}$$

Show that the left-invariant vector field corresponding to $X \in \text{Mat}_{n\times n}$ is

$$(gX)_{g\in GL_n}$$

Given $X, Y \in \text{Mat}_{n\times n}$, calculate the commutator of the corresponding vector fields by formula (11)

$$[(gX)_{g\in GL_n}, (gY)_{g\in GL_n}] = \sum_{1\leq i,j,k,\ell\leq n} g_{ik} \left(X_{k\ell}Y_{\ell j} - Y_{k\ell}X_{\ell j}\right) \frac{\partial}{\partial E_{ij}}$$

where we write matrices $g = \sum_{1\leq i,j\leq n} g_{ij}E_{ij}$ etc, in terms of their matrix coefficients in the standard basis. Restricting the above equality to the identity element $g_{ij} = \delta_{ij}$ shows that the Lie bracket on $\mathfrak{gl}_n$ is given by

$$\mathfrak{gl}_n \times \mathfrak{gl}_n \xrightarrow{[\cdot,\cdot]} \mathfrak{gl}_n, \qquad \boxed{[X, Y] = XY - YX}$$

1.6

Let us now give examples of Lie algebras, beyond $\mathrm{Lie}(GL_n) = \mathfrak{gl}_n$. Firstly, note that it's quite easy to determine the Lie algebras of matrix groups (i.e. subgroups of GL_n cut out by polynomial equations). For instance you will show in the exercise session that

$$\mathrm{Lie}(SL_n) = \mathfrak{sl}_n = \left\{X \in \mathrm{Mat}_{n\times n} \middle| \mathrm{tr}\ X = 0\right\}$$
$$\mathrm{Lie}(O_n) = \mathfrak{o}_n = \left\{X \in \mathrm{Mat}_{n\times n} \middle| X^T + X = 0\right\} = \mathfrak{so}_n = \mathrm{Lie}(SO_n)$$
$$\mathrm{Lie}(Sp_{2n}) = \mathfrak{sp}_{2n} = \left\{X \in \mathrm{Mat}_{2n\times 2n} \middle| X^T\Omega + \Omega X = 0\right\}$$
$$\mathrm{Lie}(U(n)) = \mathfrak{u}(n) = \left\{X \in \mathrm{Mat}_{n\times n}(\mathbb{C}) \middle| \bar{X}^T + X = 0\right\}$$
$$\mathrm{Lie}(SU(n)) = \mathfrak{su}(n) = \left\{X \in \mathrm{Mat}_{n\times n}(\mathbb{C}) \middle| \bar{X}^T + X = 0 \text{ and } \mathrm{tr}\ X = 0\right\}$$

Secondly, there are many more Lie algebras out there than Lie groups. For one thing, Lie algebras can be defined over any field (including characteristic p) and they may be infinite-dimensional, neither of which situation is compatible with being a tangent space of a Lie group. For example, the **Virasoro algebra** is an important infinite-dimensional Lie algebra

$$\mathrm{Vir} = \mathbb{C}c \oplus \bigoplus_{n\in\mathbb{Z}} \mathbb{C}L_n, \qquad [c, L_n] = 0, \qquad [L_m, L_n] = (m-n)L_{m+n} + \delta_{m,-n}\frac{c(m^3-m)}{12}$$

which is the fundamental object in conformal field theory. We will not be studying Lie algebras of infinite dimension or positive characteristic in this course, but they are very rich subjects.

Remark. *When we use the notation $\mathfrak{gl}_n, \mathfrak{sl}_n, \mathfrak{o}_n, \mathfrak{sp}_{2n}$ as above, we are implicitly making a statement that does not depend on the ground field. When we wish to emphasize the fact that we are considering these Lie algebras over a specific ground field $\mathbb{K}$ (such as $\mathbb{R}$ or $\mathbb{C}$), we will denote them by*

$$\mathfrak{gl}_{n,\mathbb{K}}, \mathfrak{sl}_{n,\mathbb{K}}, \mathfrak{o}_{n,\mathbb{K}}, \mathfrak{sp}_{2n,\mathbb{K}}$$

etc. From a certain point onward in our course (and we will mention this explicitly), we will focus solely on the field of complex numbers and so we will write $\mathfrak{gl}_n$ instead of $\mathfrak{gl}_{n,\mathbb{C}}$ etc thereafter.

Lecture 2

2.1

Many of the usual constructions for groups apply to Lie groups, but we must be careful to make sure the manifold structure is preserved. For example, a **Lie group homomorphism**

$$F : G \to G' \tag{14}$$

is required to be both a group homomorphism and a smooth function. Analogously, a linear transformation

$$f : \mathfrak{g} \to \mathfrak{g}'$$

is called a **Lie algebra homomorphism** if it preserves the Lie bracket:

$$\boxed{f([x,y]) = [f(x), f(y)]} \tag{15}$$

$\forall x, y \in \mathfrak{g}$, where the LHS involves the Lie bracket in $\mathfrak{g}$ and the RHS involves the Lie bracket in $\mathfrak{g}'$.

Proposition 2. *If $F : G \to G'$ is a Lie group homomorphism, then the induced derivative*

$$f : \mathfrak{g} \to \mathfrak{g}'$$

is a Lie algebra homomorphism, where $\mathfrak{g} = \mathrm{Lie}(G)$ and $\mathfrak{g}' = \mathrm{Lie}(G')$.

Proof. The Lie group homomorphism F induces a derivative on vector fields, which is easily seen to take left-invariant vector fields to left-invariant vector fields. The fact that this derivative satisfies (15) is automatic. □

2.2

We will now review the basic representation theory of Lie groups, generalizing the treatments that you have already encountered it in Math 211 or 314. We say that a Lie group G **acts** on a manifold X, denoted by

$$\boxed{G \curvearrowright X} \tag{16}$$

if there exists a smooth function

$$G \times X \to X, \qquad (g, x) \mapsto \Phi_g(x) = g \cdot x$$

that simultaneously satisfies the usual properties from group theory

$$\Phi_{gg'} = \Phi_g \circ \Phi_{g'} \qquad \text{and} \qquad \Phi_e = \mathrm{Id}_X \tag{17}$$

and is a smooth/holomorphic function of real/complex manifolds.

Example 3. *Because the groups SL_n, O_n, SO_n etc are subgroups of GL_n, they naturally act on an n-dimensional vector space. More interesting actions can be obtained by observing that various subsets of n-dimensional space are preserved by the aforementioned actions, for instance*

$$O_n(\mathbb{R}) \curvearrowright S^{n-1} = \left\{ x_1^2 + \cdots + x_n^2 = 1 \right\} \subset \mathbb{R}^n$$
$$U(n) \curvearrowright S^{2n-1} = \left\{ |z_1|^2 + \cdots + |z_n|^2 = 1 \right\} \subset \mathbb{C}^n$$

As in the usual case of group theory, we have the usual actions of G on itself

$$\begin{aligned} &\text{left action } g \cdot h = gh \\ &\text{right action } g \cdot h = hg^{-1} \\ &\text{adjoint action } g \cdot h = ghg^{-1} \end{aligned}$$

The **orbits** of an action $G \curvearrowright X$ are the sets

$$Gx = \left\{ g \cdot x \middle| g \in G \right\}$$

as x runs over X. While the left and right actions $G \curvearrowright G$ have a single orbit (in other words, they are transitive), the orbits of the adjoint action are just the conjugacy classes of G.

2.3

If a Lie group G acts on a vector space V in such a way that all the action maps

$$\Phi_g : V \to V$$

are linear transformations, then we say that V is a **representation of** G. This can be rephrased in terms of the Lie group (with operation given by composition)

$$GL(V) = \Big\{ \text{invertible linear transformations } V \to V \Big\} \tag{18}$$

in that to give a representation $G \curvearrowright V$ is the same as to give a Lie group homomorphism

$$\boxed{G \to GL(V)}, \quad g \mapsto \Big(\Phi_g : V \to V \Big) \tag{19}$$

Taking the differential of (19) at the identity gives us

$$\boxed{\mathfrak{g} \to \mathfrak{gl}(V)}, \quad x \mapsto \Big(\phi_x : V \to V \Big) \tag{20}$$

where we write $\mathfrak{g} = \text{Lie}(G)$ and define

$$\mathfrak{gl}(V) = \text{End}(V) := \Big\{ \text{linear transformations } V \to V \Big\} \tag{21}$$

Note that (21) is a vector space with respect to addition and a Lie algebra with respect to commutator, and it coincides with $\text{Lie}(GL(V))$. The following statement is an immediate consequence of (20) being a Lie algebra homomorphism, which follows from Proposition 2.

Proposition 3. *The assignment* (20) *is a* ***Lie algebra representation****, i.e. an assignment*

$$\Big\{ \phi_x : V \to V \Big\}_{x \in \mathfrak{g}}$$

of linear transformations (which depend linearly on x) such that

$$\phi_{[x,y]} = \phi_x \circ \phi_y - \phi_y \circ \phi_x \tag{22}$$

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

2.4

The adjoint action $\mathrm{Ad}_g(h) = ghg^{-1}$

$$G \curvearrowright G, \quad \mathrm{Ad}_g : G \to G \tag{23}$$

does not constitute a representation because G is not a vector space, but its derivative

$$\boxed{G \curvearrowright \mathfrak{g}, \quad \mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}} \tag{24}$$

is a representation, according to the following.

Proposition 4. *Formula* (24) *is a Lie group representation, and its derivative*

$$\boxed{\mathfrak{g} \curvearrowright \mathfrak{g}, \quad \mathrm{ad}_x : \mathfrak{g} \to \mathfrak{g}} \tag{25}$$

is a Lie algebra representation. Explicitly,

$$\boxed{\mathrm{ad}_x(y) = [x, y]} \tag{26}$$

for all $x, y \in \mathfrak{g}$. Both (24) *and* (25) *are called the* ***adjoint representation****.*

Proof. The fact that (24) is a Lie group representation is immediate, as $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ is a linear function (as are all derivatives of smooth functions) which inherits the multiplicativity property (17) from $\mathrm{Ad}_g : G \to G$. Therefore, its derivative (25) is a Lie algebra representation by Proposition 2. Formula (26) is best proved by an alternative description of the Lie bracket involving one-parameter subgroups (to be covered in the exercise session), but let us give an explicit computation when $\mathfrak{g} = \mathfrak{gl}_n$. We have

$$\mathrm{Ad}_g(h) = ghg^{-1}$$

for all $g, h \in GL_n$. Letting $h = 1 + tY$ for a "small" value of t and any $n \times n$ matrix Y (this is reasonable, since we are identifying the tangent space at $e \in GL_n$ with the ambient space of all matrices), we see that the adjoint representation is given by

$$\mathrm{Ad}_g(Y) = gYg^{-1}$$

for all $g \in GL_n$, $Y \in \mathfrak{gl}_n$. To differentiate the formula above, let $g(t) = 1 + tX$ for "small" t. Then

$$g(t)^{-1} = 1 - tX + \frac{t^2X^2}{2} - \dots$$

and so

$$\mathrm{ad}_X(Y) = \lim_{t \to 0} \frac{\mathrm{Ad}_{g(t)}(Y) - Y}{t} = \lim_{t \to 0} \frac{(1 + tX)Y(1 - tX + \dots) - Y}{t} = XY - YX$$

which coincides with the Lie bracket of $\mathfrak{gl}_n$. $\square$

Example 4. *For any vector space V, we have a tautological representation*

$$GL(V) \curvearrowright V$$

We may extend this action naturally to any tensor product of symmetric and exterior powers of V

$$GL(V) \curvearrowright \cdots \otimes S^k V \otimes \wedge^\ell V \otimes \cdots$$

*You have seen at the very end of Math 314 that these representation come into play in **Schur-Weyl duality**. This is a statement that for any $n \in \mathbb{N}$, we have a decomposition*

$$\underbrace{V \otimes \cdots \otimes V}_{n \text{ factors}} = \bigoplus_{\text{partition } \lambda} L(\lambda) \otimes S_\lambda$$

of representations of $GL(V) \times S_n$ (the symmetric group permutes the factors in the LHS), where in the RHS we write S_λ for the irreducible Specht modules of S_n, and $L(\lambda)$ for the irreducible representations of $GL(V)$. We will characterize the latter in more detail in Lectures 13 and 14.

2.5

Many of the basic notions from Math 314 apply to Lie groups as they did to finite groups. Given representations $G \curvearrowright V$ and $G \curvearrowright W$ (determined by collections $\{\Phi_g : V \to V\}_{g\in G}$ and $\{\Psi_g : W \to W\}_{g\in G}$, respectively) a G-**intertwiner** is a linear transformation

$$f : V \longrightarrow W$$

such that the following diagram commutes

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \Phi_g \downarrow & & \downarrow \Psi_g \\ V & \xrightarrow{f} & W \end{array}$$

for all $g \in G$. If we write $\Phi_g(v) = g \cdot v$ and $\Psi_g(w) = g \cdot w$ for all $v \in V$ and $w \in W$, then the property of being a G-intertwiner is equivalent to $f(g \cdot v) = g \cdot f(v)$ for all $g \in G, v \in V$. If a G-intertwiner is moreover bijective, then we call it an **isomorphism**. Recall that a subset of a vector space is called a subspace if and only if it is preserved under addition of vectors and scalar multiplication. If we have a representation $G \curvearrowright V$, then a subspace $W \subseteq V$ is called a **subrepresentation** if

$$\Phi_g(W) \subseteq W$$

for all $g \in G$. Moreover, in this case there is an induced **quotient representation**

$$G \curvearrowright V/W$$

Given representations V and W of G, we can make their direct sums, tensor products and duals into G-representations via

$$G \curvearrowright V \oplus W, \quad g \cdot (v, w) = (g(v), g(w)) \tag{27}$$

$$G \curvearrowright V \otimes W, \quad g \cdot (v \otimes w) = g(v) \otimes g(w) \tag{28}$$

$$G \curvearrowright V^\vee, \quad (g \cdot \lambda)(v) = \lambda(g^{-1} \cdot v) \tag{29}$$

The natural analogues of all notions above (intertwiners, isomorphisms, sub-and-quotient representations) apply equally well to Lie algebras as to Lie groups. The only difference lies in formulas (27), (28), (29), which must be modified in the case of Lie algebra representations to

$$\mathfrak{g} \curvearrowright V \oplus W, \quad x \cdot (v, w) = (x(v), x(w)) \tag{30}$$
$$\mathfrak{g} \curvearrowright V \otimes W, \quad x \cdot (v \otimes w) = x(v) \otimes w + v \otimes x(w) \tag{31}$$
$$\mathfrak{g} \curvearrowright V^\vee, \qquad (x \cdot \psi)(v) = -\psi(x(v)) \tag{32}$$

2.6

We will state the following basic facts for representations of Lie groups G, but they apply equally well for representations of Lie algebras $\mathfrak{g}$.

Definition 3. *A representation $G \curvearrowright V$ is called **irreducible** if it does not have any proper subrepresentations (i.e. no subrepresentations other than 0 or V).*

One of the main tools in representation theory is the following result, known as **Schur's lemma**.

Lemma 1. *Suppose we have a G-intertwiner $f : V \to W$ between two representations of G, which is not identically 0. If V is irreducible, then f is injective. If W is irreducible, then f is surjective.*

As an immediate corollary of Lemma 1, any non-zero intertwiner between two irreducible representations must be an isomorphism. All of the above is the same for Lie groups as it was for finite groups, but some things do not generalize so easily. An example of this is Maschke's theorem, which says that any complex finite-dimensional representation of a finite group $G \curvearrowright V$ has the property that any subrepresentation $W \subset V$ has a complement

$$V \cong W \oplus W' \tag{33}$$

such that W' is also a subrepresentation of V (an important consequence of this is that finite-dimensional complex representations of finite groups are completely reducible, i.e. isomorphic to direct sums of irreducible representations). This result completely fails for Lie groups in general. For example, consider the action of $\mathbb{C}$ (Lie group with respect to addition) on $V = \mathbb{C}^2$ via

$$x \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \tag{34}$$

The subspace $W = \{b = 0\}$ is a one-dimensional subrepresentation of V, but because it is the unique such subrepresentation, it is impossible to find a decomposition (33). However, compact Lie groups and unitary representations will give us a setting in which we can salvage some of these results. We will study these (and the corresponding Lie algebras) in Lecture 4.

Lecture 3

3.1

Let G be a Lie group. A subgroup $H \subset G$ is called a

- Lie subgroup if $H \hookrightarrow G$ is an immersion
- closed Lie subgroup if $H \hookrightarrow G$ is an embedding

The terminology for closed Lie subgroups is motivated by the (non-obvious) fact that they are also closed as topological spaces. Most Lie subgroups of interest will turn out to be closed, for example the stabilizers of Lie group actions $G \curvearrowright M$

$$\mathrm{Stab}_G(m) = \left\{ g \in G \middle| g \cdot m = m \right\}$$

are all closed Lie subgroups of G. However, there exist examples of non-closed Lie subgroups, e.g. the image of the group homomorphism $\mathbb{R} \to \mathbb{R}^2/\mathbb{Z}^2, t \mapsto (t, ct)$ for some $c \in \mathbb{R}\backslash\mathbb{Q}$.

Theorem 1. *(a) If $H \subseteq G$ is a normal closed Lie subgroup, then*

$$\boxed{G/H}$$

has an induced structure of a Lie group.

(b) If $f : G \to G'$ is a Lie group homomorphism, then $\mathrm{Ker}\ f$ *is a normal closed Lie subgroup, and we obtain an induced Lie group homomorphism*

$$G/\mathrm{Ker}\ f \hookrightarrow G'$$

which is an immersion. $\mathrm{Im}\ f$ *is a Lie subgroup of G' on general grounds; if it is moreover a closed Lie subgroup of G' then we have the following analogue of the first isomorphism theorem*

$$\boxed{G/\mathrm{Ker}\ f \cong \mathrm{Im}\ f}$$

(c) The center $Z(G)$ is a closed Lie subgroup.

3.2

Given a Lie algebra $\mathfrak{g}$ in the generality of Definition 2, a subspace

- $\mathfrak{h} \subset \mathfrak{g}$ is called a (Lie) **subalgebra** if it is closed under the Lie bracket, i.e. $[\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{h}$.
- $\mathfrak{h} \subset \mathfrak{g}$ is called an **ideal** if $[\mathfrak{g}, \mathfrak{h}] \subseteq \mathfrak{h}$. The kernel of any Lie algebra homomorphism is an ideal.

Check for yourself that if $\mathfrak{h} \subset \mathfrak{g}$ is an ideal, then $\mathfrak{g}/\mathfrak{h}$ inherits a Lie algebra structure. The following result is an analogue of the correspondence (or lattice) theorem that you saw for groups in Math 211.

Theorem 2. *If $\mathfrak{h} \subset \mathfrak{g}$ is an ideal, then there is a one-to-one correspondence*

$$\Big(\text{Lie subalgebras } \mathfrak{h} \subseteq \mathfrak{a} \subseteq \mathfrak{g}\Big) \leftrightarrow \Big(\text{Lie subalgebras } \bar{\mathfrak{a}} \subseteq \bar{\mathfrak{g}}\Big) \tag{35}$$

given by $\bar{\mathfrak{a}} = \pi(\mathfrak{a})$ and $\mathfrak{a} = \pi^{-1}(\bar{\mathfrak{a}})$, where

$$\pi : \mathfrak{g} \to \bar{\mathfrak{g}} := \mathfrak{g}/\mathfrak{h}$$

is the natural projection. In (35)*, $\mathfrak{a}$ is an ideal if and only if $\bar{\mathfrak{a}}$ is an ideal.*

One may also represent quotients in terms of **short exact sequences** of Lie algebras

$$0 \to \mathfrak{h} \xrightarrow{\phi} \mathfrak{g} \xrightarrow{\psi} \bar{\mathfrak{g}} \to 0 \tag{36}$$

with the implication being that $\mathrm{Im}\ \phi$ is an ideal in $\mathfrak{g}$, and $\psi : \mathfrak{g}/\mathrm{Im}\ \phi \to \bar{\mathfrak{g}}$ is an isomorphism.

Definition 4. *Given Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$ over the same ground field, their **direct sum***

$$\boxed{\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2} \tag{37}$$

has Lie bracket defined by

$$\Big[(x_1, x_2), (y_1, y_2)\Big] = \Big([x_1, y_1], [x_2, y_2]\Big)$$

In other words, the subalgebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$ in (37) *are ideals of $\mathfrak{g}$.*

For example, show that we have an isomorphism of Lie algebras

$$\mathfrak{gl}_{n,\mathbb{K}} \cong \mathbb{K} \oplus \mathfrak{sl}_{n,\mathbb{K}} \tag{38}$$

over any ground field $\mathbb{K}$ of characteristic that does not divide n, where the right-hand side is the direct sum of the trivial one-dimensional Lie algebra and the special linear Lie algebra.

Definition 5. *Given an element x in a Lie algebra $\mathfrak{g}$, its **centralizer** is*

$$\boxed{\mathfrak{z}_x(\mathfrak{g}) = \Big\{y \in \mathfrak{g} \Big| [x, y] = 0\Big\}} \tag{39}$$

*The intersection of all the centralizers is called the **center** of $\mathfrak{g}$*

$$\boxed{\mathfrak{z}(\mathfrak{g}) = \Big\{y \in \mathfrak{g} \Big| [x, y] = 0,\ \forall x \in \mathfrak{g}\Big\}} \tag{40}$$

Theorem 3. *Let G be a real/complex Lie group with real/complex Lie algebra $\mathfrak{g}$.*

(a) If $H \subseteq G$ is a Lie subgroup, not necessarily closed, then $\mathfrak{h} = \mathrm{Lie}(H)$ is a Lie subalgebra of $\mathfrak{g}$ (this correspondence $H \rightsquigarrow \mathfrak{h}$ is invertible if we restrict to connected H).

(b) If $H \subseteq G$ is a normal closed Lie subgroup, then $\mathfrak{h} = \mathrm{Lie}(H)$ is an ideal of $\mathfrak{g}$, and

$$\boxed{\mathrm{Lie}(G/H) \cong \mathfrak{g}/\mathfrak{h}}$$

(c) $\mathrm{Lie}(Z(G)) = \mathfrak{z}(\mathfrak{g})$ if G is connected.

3.3

The above Theorem establishes a correspondence between subgroups of a Lie group G and subalgebras of $\mathfrak{g} = \text{Lie}(G)$. The following result generalizes this fact.

Theorem 4. *(a) For any real/complex Lie groups G and G' (with the former being connected), there exists an injective function*

$$\Big(\textit{Lie group homomorphisms } G \to G'\Big) \rightsquigarrow \Big(\textit{Lie algebra homomorphisms } \mathfrak{g} \to \mathfrak{g}'\Big) \qquad (41)$$

where $\mathfrak{g} = \text{Lie}(G)$, $\mathfrak{g}' = \text{Lie}(G')$, given by the derivative at the identity.

(b) If furthermore G is simply connected, then (41) *is a bijection.*

It is clear why Theorem 4 requires G to be connected, because $\mathfrak{g}$ only "knows" about the connected component of the identity in G. The simply-connected assumption is necessary to rule out examples like

$$S^1 = \Big\{e^{2\pi i x}\Big| x \in \mathbb{R}\Big\} \quad \text{with} \quad \text{Lie}(S^1) = \mathbb{R}$$

(the operation on S^1 is multiplication, while the Lie bracket on $\mathbb{R}$ is trivial), in which case

$$\text{Hom}_{\text{Lie group}}(S^1, S^1) \cong \mathbb{Z} \quad \text{but} \quad \text{Hom}_{\text{Lie algebra}}(\mathbb{R}, \mathbb{R}) = \mathbb{R}$$

Proof. of Theorem 4: (a) The correspondence (41) is given by the derivative, and we have already seen that it takes a Lie group homomorphism $F : G \to G'$ to a Lie algebra homomorphism $f = F_* : \mathfrak{g} \to \mathfrak{g}'$. By Exercise Sheet 2, Problem 3, we have

$$F(\exp(x)) = \exp(f(x))$$

for any x in a neighborhood of $0 \in \mathfrak{g}$. This means that knowledge of f determines F completely in a neighborhood of $e \in G$. However, any connected Lie group is generated by any neighborhood of the identity (it is not hard to show that the subgroup generated by any open subset must be open; if $H \subset G$ is the subset generated by an open neighborhood U of the identity, then $G \backslash H$ is also open, because for any $g \in G \backslash H$ we must have $gU \cap H = \emptyset$; since G is connected, this implies that $H = G$) so we conclude that f completely determines F, i.e. the assignment (41) is injective.

(b) Let us show that any Lie algebra homomorphism $f : \mathfrak{g} \to \mathfrak{g}'$ can be lifted to a Lie group homomorphism $F : G \to G'$. To this end, we invoke the Baker-Campbell-Hausdorff formula from Exercise Sheet 2, Problem 5:

$$\exp(x) \exp(y) = \exp\left(x + y + \frac{[x, y]}{2} + \dots\right)$$

Then the function

$$F : G \to G', \qquad F(\exp(x)) = \exp(f(x))$$

gives a Lie group homomorphism in a neighborhood of the identity $e \in U \subset G$, because

$$F(\exp(x) \exp(y)) = F\left(\exp\left(x + y + \frac{[x, y]}{2} + \dots\right)\right) = \exp\left(f(x) + f(y) + f\left(\frac{[x, y]}{2}\right) + \dots\right) =$$

$$= \exp\left(f(x) + f(y) + \frac{[f(x), f(y)]}{2} + \dots\right) = \exp(f(x))\exp(f(y)) = F(\exp(x))F(\exp(y))$$

As we have seen in part (a), the group G is generated by U. Thus, for any $g \in G$ we can choose

$$e = g_0, g_1, \dots, g_{k-1}, g_k = g \tag{42}$$

where each g_{i-1} is close enough to g_i so that $g_i g_{i-1}^{-1} \in U$. This means that we can define

$$F(g) = F(g_k g_{k-1}^{-1})F(g_{k-1} g_{k-2}^{-1}) \dots F(g_2 g_1^{-1})F(g_1 g_0^{-1}) \tag{43}$$

To show that this is well-defined, the key observation is that the value of $F(g)$ above is independent of the choice of (42). To see this, consider any

$$e = g_0, g_1, \dots, g_{k-1}, g_k = g = g'_{k'}, g'_{k'-1}, \dots, g'_1, g'_0 = e$$

then string a path through the g_i's and a path through the g'_i's, and then take a homotopy between the two paths (which exists because G is simply connected). We leave the details to you. The independence of (43) on the choice of (42) also proves $F(gh) = F(g)F(h)$, because one can construct a sequence (42) from e to gh by stringing together an analogous sequence for g and one for h. □

3.4

The following result is sometimes known as Lie's third theorem.

Theorem 5. *Any finite-dimensional real/complex Lie algebra $\mathfrak{g}$ has the property that*

$$\mathfrak{g} \cong \mathrm{Lie}(G)$$

for a unique (up to isomorphism) simply connected real/complex Lie group G.

Proof. This is a quite difficult result, unless one accepts **Ado's theorem**: any finite-dimensional Lie algebra (over any field) has a **faithful** finite-dimensional representation, i.e. one such that (20) is injective. Therefore, we may assume that the Lie algebra in Theorem 5 is of the form

$$\mathfrak{g} \subset \mathfrak{gl}_{n,\mathbb{K}}$$

for $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Let $G' \subset GL_n(\mathbb{K})$ to be the closure of the subgroup generated by $\{\exp(X)\}_{X \in \mathfrak{g}}$ with respect to the usual matrix exponential. To obtain a simply connected Lie group, we define

$$G = \Big\{\text{paths } \gamma : [0,1] \to G, \gamma(0) = e\Big\} \Big/ \text{homotopy}$$

made into a Lie group with respect to pointwise multiplication of paths. The covering map

$$G \to G', \qquad \gamma \mapsto \gamma(1)$$

is a Lie group homomorphism. It is well-known that G and G' have the same tangent space at the identity, which is $\mathfrak{g}$ by construction. The uniqueness is a special case of Theorem 4. □

3.5

A particularly important special case of the results in the previous Subsection arises in the study of Lie group representations

$$G \curvearrowright V \quad \Leftrightarrow \quad \text{homomorphisms } G \to GL(V) \tag{44}$$

Theorem 4 implies that if G is connected (for example $SL_n(\mathbb{R})$, $SO_n(\mathbb{R})$, $Sp_{2n}(\mathbb{R})$, $U(n)$, $SU(n)$, $GL_n(\mathbb{C})$, $SL_n(\mathbb{C})$, $SO_n(\mathbb{C})$, $Sp_{2n}(\mathbb{C})$), then such a representation is completely determined by the induced representation of $\mathfrak{g} = \text{Lie}(G)$

$$\mathfrak{g} \curvearrowright V \quad \Leftrightarrow \quad \text{homomorphisms } \mathfrak{g} \to \mathfrak{gl}(V) \tag{45}$$

If moreover G is simply connected (for example $SU(n), SL_n(\mathbb{C}), Sp_{2n}(\mathbb{C})$), then any representation (45) can be uniquely lifted to a representation (44). This is very convenient, as it reduces the study of Lie group representations (which is a more complicated problem that interweaves algebra and geometry) to the purely algebraic problem of studying Lie algebra representations. So for instance, the discussion on finite-dimensional representations of $U(n)$ from the next Lecture will also imply that analogous discussion on finite-dimensional representations of $\mathfrak{u}(n)$. Conversely, the classification of finite-dimensional representations of $\mathfrak{sl}_{2,\mathbb{C}}$ in Lecture 5 will determine an analogous classification of finite-dimensional representations of $SL_2(\mathbb{C})$.

Lecture 4

4.1

You have already seen a version of what follows in Math 314, but we include it for review. Recall that a Hilbert space is a $\mathbb{C}$-vector space V endowed with an inner product

$$V \otimes V \xrightarrow{\langle \cdot, \cdot \rangle} \mathbb{C}$$

which is linear in the first argument, satisfies $\langle v_1, v_2 \rangle = \overline{\langle v_2, v_1 \rangle}$ for all $v_1, v_2 \in V$, and $\langle v, v \rangle \in \mathbb{R}_{>0}$ if $v \neq 0$. A linear transformation $f : V \to V$ is called **unitary** if it preserves the inner product

$$\Big\langle f(v_1), f(v_2) \Big\rangle = \Big\langle v_1, v_2 \Big\rangle$$

If $V = \mathbb{C}^n$ with the inner product

$$\Big\langle (z_1, \dots, z_n), (w_1, \dots, w_n) \Big\rangle = z_1 \overline{w}_1 + \dots + z_n \overline{w}_n \tag{46}$$

then a unitary linear transformation is simply given by a unitary matrix A as in (6), i.e. $f(v) = Av$.

Definition 6. *If V is a Hilbert space, then a representation $G \curvearrowright V$ is called **unitary** if all the action maps $\Phi_g : V \to V$ are unitary linear transformations.*

Unitary representations satisfy Maschke's theorem. Indeed, given a subrepresentation $W \subseteq V$ of a finite-dimensional representation $G \curvearrowright V$ over the complex numbers, one defines

$$W' = \Big\{ v \in V \Big| \langle v, W \rangle = 0 \Big\}$$

The fact that all the Φ_g are unitary (i.e. preserve the inner product) implies that all the Φ_g preserve W' (i.e. W' is a subrepresentation), while the fact that $V = W \oplus W'$ follows from the bilinearity and positive definiteness of the inner product. After repeated applications of Maschke's theorem, one concludes that finite-dimensional unitary representations are **completely reducible**, i.e.

$$V \cong W_1 \oplus \dots \oplus W_k$$

where $W_1, \dots, W_k$ are irreducible representations.

4.2

A complex representation V of a Lie group G is called **unitarizable** if it admits an inner product with respect to which it is a unitary representation. Unitarizable representations also have desirable properties like Maschke's theorem, and the following discussion provides a large source of examples.

Definition 7. *A real Lie group is called **compact** if it is compact as a topological space.*

Since most Lie groups we will encounter are matrix groups, compactness is equivalent (by the Heine-Borel theorem) with being closed and bounded. Thus, we see that $SL_n, SO_n, Sp_{2n}, \dots$ for $n \geq 2$ are not compact, since one can easily cook up a matrix in each of these sets that has top-left entry of arbitrarily large absolute value. However, the unitary group $U(n)$ is compact because

- it is closed, as the equation $\bar{A}^T A = I_n$ is polynomial in the real and imaginary parts of the entries $\{a_{ij}\}_{1\leq i,j\leq n}$ of a matrix A,
- it is bounded, as $\sum_{1\leq i,j\leq n} |a_{ij}|^2 = \mathrm{tr}(\bar{A}^T A) = n$

Proposition 5. *Any complex representation V of a compact Lie group G is unitarizable.*

Proof. Since $V \cong \mathbb{C}^n$ for some n, we can consider the standard inner product

$$V \otimes V \xrightarrow{\langle \cdot,\cdot \rangle} \mathbb{C}$$

given by formula (46). We may define another inner product by "averaging" the one above over the compact Lie group G

$$\Big\langle v_1, v_2 \Big\rangle^{\text{avg}} = \int_G \Big\langle \Phi_g(v_1), \Phi_g(v_2) \Big\rangle dg \tag{47}$$

where dg is known as a **Haar measure** on G (i.e. a measure which is invariant under the right action, i.e. $dg = d(gh^{-1})$ for all $h \in G$; its existence is a difficult theorem). The compactness of G means that the formula above is well-defined, and the fact that it determines an inner product is a straightforward check of the axioms, which we recommend you do. To show that V with the inner product (47) is a unitary representation, choose any $h \in G$ and note that

$$\Big\langle \Phi_h(v_1), \Phi_h(v_2) \Big\rangle^{\text{avg}} = \int_G \Big\langle \Phi_{gh}(v_1), \Phi_{gh}(v_2) \Big\rangle dg = \int_G \Big\langle \Phi_g(v_1), \Phi_g(v_2) \Big\rangle d(gh^{-1})$$

The RHS of the above is equal to the RHS of (47) precisely because dg is a Haar measure. □

4.3

As we have seen in the previous Subsection, the representation theory of compact Lie groups is simpler than that of arbitrary Lie groups. However, the two can be related by a procedure known as "Weyl's unitary trick". The following is an important result, which we will not prove.

Theorem 6. *Any Lie group has a maximal compact subgroup, and any two such maximal compact subgroups are conjugates of each other.*

Thus, we will often speak of "the" maximal compact subgroup, at least up to conjugation. For example, in the following table we list maximal compact subgroups of important matrix groups

$$\begin{aligned} GL_n(\mathbb{C}) &\rightsquigarrow U(n) \\ SL_n(\mathbb{C}) &\rightsquigarrow SU(n) \\ GL_n(\mathbb{R}) &\rightsquigarrow O_n(\mathbb{R}) \\ SL_n(\mathbb{R}) &\rightsquigarrow SO_n(\mathbb{R}) \end{aligned}$$

The maximal compact subgroup of $Sp_{2n}(\mathbb{R})$ is also isomorphic to $U(n)$, while the maximal compact subgroup of $Sp_{2n}(\mathbb{C})$ is its intersection with $U(2n)$ inside square matrices of size $2n$. Induction and restriction (which you learned in the context of finite groups in Math 314) allow one to relate the representations of a Lie group with those of its maximal compact subgroup. When it comes to Lie algebras, compact Lie groups are special for the following reason.

Theorem 7. *If G is a connected compact Lie group, then the exponential map $\mathfrak{g} \to G$ is surjective.*

Proof. An important result called the **Peter-Weyl theorem** implies that any compact Lie group G embeds into some $U(n)$. Since the exponential map $\mathfrak{g} \to G$ is then restricted from the exponential map

$$\mathfrak{u}(n) \to U(n)$$

then it suffices to show that the latter is surjective. This is a well-known consequence of the fact that any unitary matrix is diagonalizable

$$g = P \cdot \text{diag}(e^{i\theta_1}, \dots, e^{i\theta_n}) \cdot P^{-1}$$

with P unitary and $\theta_1, \dots, \theta_n \in \mathbb{R}$. Therefore, the logarithm of g

$$x = P \cdot \text{diag}(i\theta_1, \dots, i\theta_n) \cdot P^{-1}$$

is well-defined and lies in $\mathfrak{u}(n)$. □

4.4

Until now, we have presented results that were completely parallel between the real and complex settings. We will now show the rich interplay between these two settings.

Definition 8. *If $\mathfrak{k}$ is a real Lie algebra, we call*

$$\mathfrak{k}_{\mathbb{C}} = \mathfrak{k} \otimes_{\mathbb{R}} \mathbb{C} = \mathfrak{k} \oplus \mathfrak{k}i$$

(with the same Lie bracket, but extended to complex coefficients) the ***complexification*** *of $\mathfrak{k}$.*

Conversely, if a complex Lie algebra $\mathfrak{g}$ is isomorphic to $\mathfrak{k}_{\mathbb{C}}$ for a real Lie algebra $\mathfrak{k}$, then we call $\mathfrak{k}$ a **real form** of $\mathfrak{g}$. Complexification is a powerful tool, but it is not always obvious, as shown by the following example

$$(\mathfrak{sl}_{n,\mathbb{R}})_{\mathbb{C}} \cong \mathfrak{sl}_{n,\mathbb{C}} \cong \mathfrak{su}(n)_{\mathbb{C}} \tag{48}$$

(in other words, $\mathfrak{sl}_n(\mathbb{C})$ has two interesting real forms). While the first isomorphism is obvious (just extend the coefficients of matrices from real to complex), we invite you to prove the second one.

Definition 9. *Let K be a compact real Lie group. A* ***complexification*** *of K is a complex Lie group G together with an embedding of smooth manifolds*

$$K \hookrightarrow G$$

which is universal with respect to Lie group homomorphisms from K to complex Lie groups.

Theorem 8. *Any compact real Lie group K admits a complexification G, whose maximal compact subgroup is K itself. The Lie algebra $\mathfrak{k} = \text{Lie}(K)$ is a real form of $\mathfrak{g} = \text{Lie}(G)$.*

In (48), we showed that $\mathfrak{sl}_{n,\mathbb{C}}$ is a complexification of $\mathfrak{su}(n)$. Let us now show that

$$SL_n(\mathbb{C}) \quad \text{is a complexification of} \quad SU(n) \tag{49}$$

Any compact subgroup of $SL_n(\mathbb{C})$ preserves an inner product on $\mathbb{C}^n$ by the argument in the proof of Proposition 5, so any such compact subgroup must be contained inside a conjugate of the unitary group; thus, we conclude that $SU(n)$ is a maximal compact subgroup of $SL_n(\mathbb{C})$. To see that (49) satisfies the universality property of Definition 9, a good way is to realize $SU(n)$ and $SL_n(\mathbb{C})$ as one and the same matrix group, but the former with real coefficients and the latter with complex coefficients. The solution is to consider

$$\left\{A, B \in \mathrm{Mat}_{n\times n} \middle| AA^T + BB^T = I_n, AB^T = BA^T, \det(A + Bi) = 1\right\} \tag{50}$$

with multiplication given by

$$(A, B)(A', B') = (AA' - BB', AB' + BA')$$

Check that the above multiplication defines a group. When A, B are real matrices, we see that $A+Bi$ is a unitary matrix, and so we recognize the above group as $SU(n)$. When A, B are complex matrices, one needs to check that any matrix $g \in SL_n(\mathbb{C})$ can be uniquely written as $A+Bi$, where A and B are complex matrices which satisfy the conditions in (50) (set $A = \frac{g+g^{T,-1}}{2}, B = \frac{g-g^{T,-1}}{2i}$).

Remark: as a partial converse to Theorem 8, we have the following (you are not expected to know what "semisimple" means yet).

Theorem 9. *Suppose a complex Lie algebra $\mathfrak{g}$ is semisimple, in the sense of Definition 13. Then it has a real form $\mathfrak{k}$ which is the Lie algebra of a compact Lie group K. Moreover, if G is the connected complex Lie group with Lie algebra $\mathfrak{g}$, one can choose K to be a maximal compact subgroup of G.*

4.5

We have already seen that compact Lie groups such as $SU(n)$ have completely reducible complex finite-dimensional representations

$$V \cong V_1 \oplus \cdots \oplus V_k$$

where V_i are all irreducible. Since $SU(n)$ is simply connected, then Theorem 4 (see also the discussion in Subsection 3.5) tells us that representations of $\mathfrak{su}(n)$ are also completely reducible.

Proposition 6. *If $\mathfrak{k}$ is a real Lie algebra with complexification $\mathfrak{g} = \mathfrak{k}_{\mathbb{C}}$, then any complex representation of $\mathfrak{k}$ can be uniquely made into a representation of $\mathfrak{g}$. Moreover, any $\mathfrak{k}$-intertwiner can be uniquely upgraded to an $\mathfrak{g}$-intertwiner.*

The Proposition above is straightforward. As a consequence of Proposition 6 and formula (48), we conclude that the representation theory of $\mathfrak{sl}_{n,\mathbb{C}}$ inherits the properties of the representation theory of $\mathfrak{su}(n)$. In particular, any complex finite-dimensional representation of $\mathfrak{sl}_{n,\mathbb{C}}$ is completely reducible. Of course, there are more explicit ways to see the complete reducibility of $\mathfrak{sl}_{n,\mathbb{C}}$ representations, and we will now approach this from the viewpoint of semisimple Lie algebras.

Lecture 5

5.1

Before we develop the general theory of Lie algebras, let us focus on the simplest example:

$$\mathfrak{sl}_2 = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| a, b, c, d \in \mathbb{C}, a + d = 0 \right\} \tag{51}$$

In keeping with standard practice in representation theory, we use the notation $\mathfrak{sl}_2$ even though the subsequent discussion requires our working over the field of complex numbers (the most important thing being that we restrict attention to complex representations hereafter). Thus, $\mathfrak{sl}_2$ refers to a three-dimensional complex Lie algebra, with basis given by

$$\mathfrak{sl}_2 = \mathbb{C}E \oplus \mathbb{C}H \oplus \mathbb{C}F$$

where

$$E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{52}$$

The Lie bracket can be easily computed from commutators of matrices

$$[H, E] = 2E, \qquad [H, F] = -2F, \qquad [E, F] = H \tag{53}$$

It is easy to construct a one-dimensional representation $L(0) = \mathbb{C}$ of $\mathfrak{sl}_2$ (just have all of E, F, H act as 0) and a two-dimensional representation $L(1) = \mathbb{C}^2$ of $\mathfrak{sl}_2$ (the natural action on $\mathbb{C}^2$ given by the 2×2 matrices (52)). Convince yourselves that they are both irreducible representations.

5.2

The action $\mathfrak{sl}_2 \curvearrowright \mathbb{C}^2$ generalizes to the symmetric powers of $\mathbb{C}^2$

$$S^n\mathbb{C}^2 = \left\{ \text{linear combinations of } w_1 \otimes \cdots \otimes w_n \middle| w_1, \ldots, w_n \in \mathbb{C}^2 \right\} \Big/ \Big(\cdots w \otimes w' \cdots - \cdots w' \otimes w \cdots \Big)$$

If we let e_1, e_2 be the standard basis of $\mathbb{C}^2$, then we obtain a basis of the n-th symmetric power

$$S^n\mathbb{C}^2 = \bigoplus_{i=0}^{n} \mathbb{C}v_i, \qquad \text{where} \qquad v_i = e_1^{\otimes i} \otimes e_2^{\otimes n-i}$$

and so $S^n\mathbb{C}^2$ has dimension $n + 1$. Moreover, the action $\mathfrak{sl}_2 \curvearrowright \mathbb{C}^2$ extends to an action

$$\boxed{\mathfrak{sl}_2 \curvearrowright L(n) = S^n\mathbb{C}^2}$$

given by the formula $x \cdot (w_1 \otimes \cdots \otimes w_n) = \sum_{i=1}^n w_1 \otimes \cdots \otimes x(w_i) \otimes \cdots \otimes w_n$ and explicitly by

$$E \cdot v_i = (n - i)v_{i+1} \tag{54}$$

$$F \cdot v_i = iv_{i-1} \tag{55}$$

$$H \cdot v_i = (2i - n)v_i \tag{56}$$

It is easy to check from the formulas above that $L(n)$ is irreducible and that H acts by a diagonalizable operator. We will soon see that the latter is actually a general feature. Recall that we only consider complex representations in this Lecture.

Proposition 7. *Any $n+1$ dimensional irreducible representation of $\mathfrak{sl}_2$ is isomorphic to $L(n)$.*

Proof. Consider an irreducible representation $\mathfrak{sl}_2 \curvearrowright V$ and let us consider the eigenspaces of H

$$V_\ell = \left\{ v \in V \middle| H \cdot v = \ell v \right\}$$

It is an easy consequence of (53) that

$$E \cdot V_\ell \subseteq V_{\ell+2}$$

$$F \cdot V_\ell \subseteq V_{\ell-2}$$

for all ℓ. This implies that the direct sum of V_ℓ is a subrepresentation of V, which must be non-zero because H has at least one non-zero eigenvector (this requires our working over the complex numbers). Because V is irreducible, this implies that

$$V = \bigoplus_{\ell \in \mathbb{C}} V_\ell$$

Because V is finite-dimensional, only finitely many of the V_ℓ's are nonzero. Let us consider a maximal such ℓ, i.e. such that $V_\ell \neq 0$ but $V_{\ell+2} = 0$. For any $0 \neq v \in V_\ell$, show using (53) that

$$v, Fv, F^2 v, \dots, F^n v \tag{57}$$

form a subrepresentation of V, where $n+1$ is the smallest positive integer such that $F^{n+1}v = 0$. The vectors (57) are linearly independent, because they lie in different eigenspaces of H. Because V is irreducible, we conclude that

$$V = \bigoplus_{i=0}^{n} \mathbb{C} F^i v$$

It is easy to check that the assignment $F^i v \mapsto \frac{v_{n-i}}{(n-i)!}$ gives an isomorphism $V \cong L(n)$.

□

5.3

Having fully characterized the irreducible representations of $\mathfrak{sl}_2$, let us now characterize general finite-dimensional representations. The last paragraph of Subsection 4.5 shows that a finite-dimensional representation of $\mathfrak{sl}_2 \curvearrowright V$ splits up as a direct sum of irreducible representations. By Proposition 7, we have

$$V \cong L(n_1) \oplus \dots \oplus L(n_k) \tag{58}$$

for certain $n_1, \dots, n_k \in \mathbb{N}$. This property is called the **complete reducibility** of finite-dimensional representations of $\mathfrak{sl}_2$.

Proposition 8. *Any finite-dimensional representation $\mathfrak{sl}_2 \curvearrowright V$ has a decomposition*

$$V = \bigoplus_{\ell \in \mathbb{Z}} V_\ell \tag{59}$$

where $V_\ell = \{v \in V | Hv = \ell v\}$. Those ℓ's such that $V_\ell \neq 0$ in the above formula are called the ***weights*** *of V, and the corresponding V_ℓ are called the* ***weight spaces****.*

Note how different the situation would have been if we replaced $\mathfrak{sl}_2$ by the one-dimensional Lie algebra $\mathbb{C}H \cong \mathfrak{gl}_1$. Since a representation of the latter Lie algebra boils down to an abritrary linear transformation H on a vector space, there would be no restriction on the eigenvalues of such an H. In stark contrast, the presence of E and F in the Lie algebra $\mathfrak{sl}_2$ forces H to act as a diagonalizable matrix with integer eigenvalues.

Proof. of Proposition 8: By (58), it suffices to prove the result for $V = L(n)$. In this case, we saw in (56) that the irreducible representation $L(n)$ has a weight decomposition (59) with weights

$$n, n-2, \dots, 2-n, -n \tag{60}$$

□

Because the weights of irreducible representations are symmetric around the origin, and because every finite-dimensional representations of $\mathfrak{sl}_2$ is a direct sum of the form (58), we have the following consequence (which plays an important part in the hard Lefschetz theorem in algebraic geometry).

Corollary 1. *If V is a finite-dimensional representation of $\mathfrak{sl}_2$, then for all $k \in \mathbb{N}$, the linear transformations E^k and F^k give isomorphisms between the k-th and $-k$-th weight subspaces of V.*

Corollary 2. *Assume that a representation $\mathfrak{sl}_2 \curvearrowright V$ has finite-dimensional weight subspaces, and the subrepresentation generated by every vector is finite-dimensional. Then V is finite-dimensional.*

Proof. If V were not finite-dimensional, then we can inductively construct a subrepresentation

$$L(n_1) \oplus \cdots \oplus L(n_k) \hookrightarrow V \tag{61}$$

for arbitrarily large $k \in \mathbb{N}$ (with $n_1, \dots, n_k \in \mathbb{N}$ being arbitrary). Indeed, consider a subrepresentation as in (61) and let v be a vector not inside it. By the hypothesis, v generates a finite-dimensional $\mathfrak{sl}_2$ representation W. Letting $\bar{W}$ be the sum of W and $L(n_1) \oplus \cdots \oplus L(n_k)$ means that $\bar{W}$ is a finite-dimensional representation of $\mathfrak{sl}_2$, hence completely reducible. We may then find natural numbers $n_{k+1}, \dots, n_{k+k'}$ such that

$$L(n_1) \oplus \cdots \oplus L(n_{k+k'}) \cong \bar{W} \hookrightarrow V$$

However, every one of the irreducible representations in (61) contributes a dimension of 1 to either the 0-th or the 1-st weight subspace of V, by (60). Since the latter are finite-dimensional, we obtain a contradiction. □

5.4

We will now define the **Casimir operator** for $\mathfrak{sl}_2$, which will be introduced in all its glory in Subsection 8.4. Consider any representation

$$\mathfrak{sl}_2 \curvearrowright V$$

and consider the linear transformation

$$C = EF + FE + \frac{H^2}{2} : V \to V \tag{62}$$

(above, we slightly abuse notation by writing E, F, H for the linear transformations on V induced by the same-named elements of $\mathfrak{sl}_2$). We stress the fact that C is not an element of $\mathfrak{sl}_2$, but it can be understood as an element in the universal enveloping algebra of $\mathfrak{sl}_2$, as in the following Lecture. In the meantime, let us observe that the defining relations in the Lie algebra $\mathfrak{sl}_2$ imply that

$$C = 2FE + H + \frac{H^2}{2} \tag{63}$$

More importantly, we have the following property.

Proposition 9. *The operator C commutes with E, F, H. In particular, if V is irreducible, then C is a scalar multiple of the identity (by Schur's Lemma).*

Proof. The Proposition is easy, but very important, so we present the details. Formulas (53) and the Leibniz rule for commutators of products imply that

$$[E, C] = E[E, F] + [E, F]E + \frac{H[E, H]}{2} + \frac{[E, H]H}{2} = EH + HE - HE - EH = 0$$
$$[F, C] = [F, E]F + F[F, E] + \frac{H[F, H]}{2} + \frac{[F, H]H}{2} = -HF - FH + HF + FH = 0$$
$$[H, C] = [H, E]F + E[H, F] + [H, F]E + F[H, E] = 2EF - 2EF - 2FE + 2FE = 0$$

□

To compute the constant by which C acts on the irreducible representation $L(n)$, it suffices to compute its action on the highest weight vector v_n. Because $E \cdot v_n = 0$, then (63) implies that

$$C \cdot v_n = \frac{n(n+2)}{2} v_n$$

so we conclude that the constant in question is $\frac{n(n+2)}{2}$. For later purposes, it will be important to calculate the action of the terms EF and FE on the ℓ-th weight subspace of $L(n)$ for all $\ell \in \mathbb{Z}$, which is also quite easy from formulas (54)-(55)

$$EF \cdot v_i = i(n - i + 1) \cdot v_i \quad \Rightarrow \quad EF\Big|_{\text{weight } \ell} = \frac{(n + \ell)(n - \ell + 2)}{4} \tag{64}$$

$$FE \cdot v_i = (i + 1)(n - i) \cdot v_i \quad \Rightarrow \quad FE\Big|_{\text{weight } \ell} = \frac{(n + \ell + 2)(n - \ell)}{4} \tag{65}$$

$$\frac{H^2}{2} \cdot v_i = \frac{(2i - n)^2}{2} \cdot v_i \quad \Rightarrow \quad \frac{H^2}{2}\Big|_{\text{weight } \ell} = \frac{\ell^2}{2} \tag{66}$$

This gives us another proof of $C|_{\text{weight } \ell} = \frac{(n+\ell)(n-\ell+2)}{4} + \frac{(n+\ell+2)(n-\ell)}{4} + \frac{\ell^2}{2} = \frac{n(n+2)}{2}$, $\forall \ell$.

5.5

Let us finally present the character theory of $\mathfrak{sl}_2$, which will give us a very elegant way to determine the numbers $n_1, \ldots, n_k$ in (58). Specifically, for a finite-dimensional representation with weight decomposition (59), we define its **character** as

$$\chi_V = \sum_{\ell \in \mathbb{Z}} (\dim V_\ell) t^\ell$$

It is easy to check that the character satisfies the properties

$$\chi_{V \oplus V'} = \chi_V + \chi_{V'} \tag{67}$$

$$\chi_{V \otimes V'} = \chi_V \chi_{V'} \tag{68}$$

$$\chi_{V^\vee} = \overline{\chi_V} \tag{69}$$

with respect to direct sum, tensor product and dual representations (see (30), (31), (32)). Moreover, the explicit description of the weight spaces of $L(n)$ in (60) shows that

$$\chi_{L(n)} = \frac{t^{n+1} - t^{-n-1}}{t - t^{-1}} \tag{70}$$

By (67), the character of an arbitrary finite-dimensional representation V that decomposes as (58) would be

$$\chi_V = \sum_{i=1}^{k} \frac{t^{n_i+1} - t^{-n_i-1}}{t - t^{-1}}$$

and it is easy to see that one can extract the set $\{n_1, \ldots, n_k\}$ from χ_V. Thus, the decomposition (58) is unique up to reordering the summands, and is completely determined by χ_V. This allows us to prove the following formula for all $m \geq n$, known as the **Clebsch-Gordan** rule

$$L(m) \otimes L(n) \cong L(m+n) \oplus L(m+n-2) \oplus \cdots \oplus L(m-n+2) \oplus L(m-n) \tag{71}$$

simply by showing that the left and right-hand sides have the same character (use (67) and (68)).

Lecture 6

6.1

We will now systematically study Lie algebras $\mathfrak{g}$ and their representations

$$\mathfrak{g} \curvearrowright V \tag{72}$$

First of all, while the notion of Lie algebra representations may seem strange at first (for any $x \in \mathfrak{g}$ you get a linear transformation $\phi_x : V \to V$ such that

$$\phi_{[x,y]} = \phi_x \circ \phi_y - \phi_y \circ \phi_x$$

for all $x, y \in \mathfrak{g}$), it is actually a particular case of the familiar notion of an algebra representation.

Definition 10. *Let $\mathfrak{g}$ be a Lie algebra over a field $\mathbb{K}$. Make the vector space*

$$T\mathfrak{g} = \mathbb{K} \bigoplus \mathfrak{g} \bigoplus \mathfrak{g} \otimes \mathfrak{g} \bigoplus \mathfrak{g} \otimes \mathfrak{g} \otimes \mathfrak{g} \bigoplus \dots$$

into an algebra (i.e. a ring with $\mathbb{K}$ inside) via concatenation of tensors. Then the quotient

$$\boxed{U\mathfrak{g} = T\mathfrak{g} \Big/ \Big(x \otimes y - y \otimes x - [x,y] \Big)_{x,y \in \mathfrak{g}}} \tag{73}$$

is called the ***universal enveloping algebra*** *of $\mathfrak{g}$.*

Note that in the right-hand side of (73), we factor by the two-sided ideal consisting of linear combinations of tensors

$$t_1 \otimes x \otimes y \otimes t_2 - t_1 \otimes y \otimes x \otimes t_2 - t_1 \otimes [x,y] \otimes t_2 \tag{74}$$

for any tensors $t_1, t_2 \in T\mathfrak{g}$ and any $x, y \in \mathfrak{g}$. The effect this has on $U\mathfrak{g}$ is to ensure that $[x,y]$ is identified with the commutator $x \otimes y - y \otimes x$ in all formulas. While the construction of the universal enveloping algebra might seem a little dry, prove by yourself the fact that (72) is a Lie algebra representation if and only if

$$U\mathfrak{g} \curvearrowright V$$

is an algebra module, i.e. we have for all $z \in U\mathfrak{g}$ a linear transformation $\phi_z : V \to V$ such that

$$\phi_{zw} = \phi_z \circ \phi_w \quad \text{and} \quad \phi_1 = \mathrm{Id}_V$$

6.2

In general, quotients of non-commutative algebras such as (73) are quite badly behaved, e.g. they could have zero divisors. However, this is not true for universal enveloping algebras of Lie algebras due to an important structural result known as the Poincaré-Birkhoff-Witt (PBW) theorem. In a nutshell, this theorem starts from a $\mathbb{K}$-basis

$$x_1, \dots, x_n \tag{75}$$

of $\mathfrak{g}$ as a vector space, and claims that the symbols

$$x_1^{a_1} \dots x_n^{a_n} = \underbrace{x_1 \otimes \dots \otimes x_1}_{a_1 \text{ factors}} \otimes \underbrace{x_2 \otimes \dots \otimes x_2}_{a_2 \text{ factors}} \otimes \dots \otimes \underbrace{x_n \otimes \dots \otimes x_n}_{a_n \text{ factors}} \tag{76}$$

give rise to a basis of $U\mathfrak{g}$, as $a_1, \dots, a_n$ range over the non-negative integers.

Theorem 10. *We have a vector space isomorphism*

$$U\mathfrak{g} = \bigoplus_{a_1,\dots,a_n=0}^{\infty} \mathbb{K} \cdot x_1^{a_1} \dots x_n^{a_n} \tag{77}$$

Let us show that the symbols $x_1^{a_1} \dots x_n^{a_n}$ span $U\mathfrak{g}$ (the fact that they are linearly independent is more difficult, and we will not prove it). By definition, $U\mathfrak{g}$ is spanned by tensors of the form $x_{i_1} \otimes \dots \otimes x_{i_k}$ where $1 \leq i_1, \dots, i_k \leq n$. We are trying to prove the fact that any such tensor can be "ordered" so as to have $i_1 \leq \dots \leq i_k$. If at some point we have $i_s > i_{s+1}$, we apply the equality

$$\Big(\dots \otimes x_{i_s} \otimes x_{i_{s+1}} \otimes \dots \Big) = \Big(\dots \otimes x_{i_{s+1}} \otimes x_{i_s} \otimes \dots \Big) + \Big(\dots \otimes [x_{i_s}, x_{i_{s+1}}] \otimes \dots \Big)$$

One can re-express the Lie bracket $[x_{i_s}, x_{i_{s+1}}]$ as a linear combination of x_j's, and note that the right-most term in the above expression has $k-1$ tensor factors. Thus, after finitely many such steps, any tensor $x_{i_1} \otimes \dots \otimes x_{i_k}$ can be written as a linear combination of ordered tensors.

Note that the symbols (76) run over a basis of the symmetric algebra

$$S\mathfrak{g} = \mathbb{K} \bigoplus \mathfrak{g} \bigoplus S^2\mathfrak{g} \bigoplus S^3\mathfrak{g} \bigoplus \dots$$

Therefore, (77) is an isomorphism of vector spaces

$$U\mathfrak{g} \cong S\mathfrak{g} \tag{78}$$

However, we note that (78) is not an isomorphism of graded vector spaces. Indeed, while $S\mathfrak{g}$ and $T\mathfrak{g}$ are graded algebras (with $S^n\mathfrak{g}$ and $\mathfrak{g}^{\otimes n}$ in degree n), the quotient $U\mathfrak{g}$ is not graded because we set the degree 2 element $x \otimes y - y \otimes x$ equal to the degree 1 element $[x, y]$. However, $U\mathfrak{g}$ is a **filtered algebra**, i.e. there exists a sequence of subspaces

$$U_0\mathfrak{g} \subset U_1\mathfrak{g} \subset \dots \subset U_n\mathfrak{g} \subset \dots \subset U\mathfrak{g} \quad \text{such that } U\mathfrak{g} = \bigcup_{n=0}^{\infty} U_n\mathfrak{g}$$

such that $U_k\mathfrak{g} \cdot U_\ell\mathfrak{g} \subseteq U_{k+\ell}\mathfrak{g}$. Indeed, the natural choice is to define $U_k\mathfrak{g}$ as the image of $\oplus_{i=0}^k \mathfrak{g}^{\otimes i}$ in the quotient (73), which is a good idea because any commutation relation between $\leq k$ tensor factors will also involve $\leq k$ tensor factors.

Proposition 10. *For any $x \in U_k\mathfrak{g}$ and $y \in U_\ell\mathfrak{g}$, we have*

$$xy - yx \in U_{k+\ell-1}\mathfrak{g}$$

Therefore, the induced associated graded algebra

$$\text{gr } U\mathfrak{g} = \bigoplus_{n=0}^{\infty} U_n\mathfrak{g} \Big/ U_{n-1}\mathfrak{g}$$

is commutative.

Proof. Assume $x = x_1 \otimes \cdots \otimes x_k$ and $y = y_1 \otimes \cdots \otimes y_\ell$. Then the commutator $xy - yx$ equals

$$x_1 \otimes \cdots \otimes x_k \otimes y_1 \otimes \cdots \otimes y_\ell - y_1 \otimes \cdots \otimes y_\ell \otimes x_1 \otimes \cdots \otimes x_k =$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{\ell} x_1 \otimes \cdots \otimes x_{i-1} \otimes y_1 \otimes \cdots \otimes y_{j-1} \otimes \underbrace{\Big(x_i \otimes y_j - y_j \otimes x_i\Big)}_{=[x_i, y_j]} \otimes y_{j+1} \otimes \cdots \otimes y_\ell \otimes x_{i+1} \otimes \cdots \otimes x_k$$

which is clearly in $U_{k+\ell-1}\mathfrak{g}$.

□

With Proposition 10 in mind, we can upgrade the Poincaré-Birkhoff-Witt theorem to the existence of an isomorphism of graded algebras

$$\boxed{\operatorname{gr} U\mathfrak{g} \cong S\mathfrak{g}} \tag{79}$$

which sends $x_{i_1} \otimes \cdots \otimes x_{i_k}$ to $x_{i_1} \dots x_{i_k}$. In particular, this shows that the natural composition

$$\mathfrak{g} \hookrightarrow T\mathfrak{g} \twoheadrightarrow U\mathfrak{g} \tag{80}$$

is injective, which is not obvious by the mere fact that $U\mathfrak{g}$ is a quotient of $T\mathfrak{g}$ by a two-sided ideal.

Example 5. *When $\mathfrak{g} = \mathfrak{sl}_2$, the universal enveloping algebra is quite simple*

$$U\mathfrak{sl}_2 = \mathbb{K}\Big\langle E, F, H\Big\rangle \Big/ (HE = E(H+2), HF = F(H-2), EF - FE = H)$$

This allows us to construct numerous representations of $\mathfrak{sl}_2$, such as for any $\lambda \in \mathbb{K}$

$$M(\lambda) = \mathbb{K}[F] = \mathbb{K} \oplus \mathbb{K}F \oplus \mathbb{K}F^2 \oplus \dots \tag{81}$$

via $E \cdot 1 = 0$, $H \cdot 1 = \lambda$ and F acting by multiplication. The rest of the action is determined by the defining relations of $U\mathfrak{sl}_2$. The above infinite-dimensional representation has weights $\lambda, \lambda - 2, \lambda - 4, \dots,$ in stark contrast with finite-dimensional representations, which we have seen have weights which are all integers and symmetric around 0. In Lecture 13, we will see that (81) *is an example of a Verma module.*

6.3

We will now introduce the natural Lie algebra versions of solvable and nilpotent groups, respectively. Note that a Lie algebra $\mathfrak{g}$ is called **abelian** if its Lie bracket is identically 0. The definitions in the present Lecture apply to any ground field $\mathbb{K}$.

Definition 11. *A Lie algebra $\mathfrak{g}$ is called* ***solvable*** *if it has a chain of Lie subalgebras*

$$0 = \mathfrak{g}_0 \subset \mathfrak{g}_1 \subset \cdots \subset \mathfrak{g}_k = \mathfrak{g}$$

such that $\mathfrak{g}_{i-1}$ is an ideal in $\mathfrak{g}_i$ and $\mathfrak{g}_i / \mathfrak{g}_{i-1}$ is abelian for all $i \in \{1, \dots, k\}$.

As in group theory, Definition 11 may be re-expressed in terms of commutators. Given subspaces A, B of a Lie algebra $\mathfrak{g}$, we will write

$$[A, B] = \text{span}\Big\{[a, b]\Big| a \in A, b \in B\Big\}$$

In particular, the **derived (or commutator) Lie subalgebra** of $\mathfrak{g}$ is

$$\boxed{D\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]} \tag{82}$$

Then it is a well-known fact, which we invite you to prove, that $\mathfrak{g}$ is solvable if and only if the so-called **derived** series

$$\mathfrak{g} \supseteq D\mathfrak{g} \supseteq D(D\mathfrak{g}) \supseteq \dots \tag{83}$$

eventually terminates with the 0 subalgebra.

Definition 12. *A Lie algebra $\mathfrak{g}$ is called* ***nilpotent*** *if it has a chain of ideals*

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

such that $[\mathfrak{g}, \mathfrak{g}_i] \subset \mathfrak{g}_{i-1}$ for all $i \in \{1, \dots, k\}$.

It is a well-known fact, which we invite you to prove, that $\mathfrak{g}$ is nilpotent if and only if the series

$$\mathfrak{g} \supseteq [\mathfrak{g}, \mathfrak{g}] \supseteq [\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}]] \supseteq \dots$$

eventually terminates with the 0 subalgebra. Since a Lie algebra is abelian if and only if $[\mathfrak{g}, \mathfrak{g}] = 0$, we conclude that

$$\text{abelian} \subset \text{nilpotent} \subset \text{solvable}$$

holds in the world of Lie algebras as it did in the case of groups. The fundamental example of abelian, nilpotent and solvable Lie algebras are

$$\mathfrak{h} = \Big\{\text{diagonal matrices}\Big\} \tag{84}$$

$$\mathfrak{n} = \Big\{\text{strictly upper triangular matrices}\Big\} \tag{85}$$

$$\mathfrak{b} = \Big\{\text{upper triangular matrices}\Big\} \tag{86}$$

respectively, all regarded as Lie subalgebras of $\mathfrak{gl}_n$ with the usual Lie bracket commutator. It is a good idea to check the above statements. The following result is proved just like in basic group theory.

Proposition 11. *Any subalgebra or quotient of an abelian / nilpotent / solvable Lie algebra is abelian / nilpotent / solvable. Conversely, if we have an ideal*

$$\mathfrak{i} \subseteq \mathfrak{g}$$

such that $\mathfrak{i}$ and $\mathfrak{g}/\mathfrak{i}$ are solvable, then $\mathfrak{g}$ is solvable (in order to have this property for nilpotent Lie algebra, we would need to further assume that $\mathfrak{i}$ lies in the center of $\mathfrak{g}$).

6.4

For the remainder of this Lecture, we assume that the ground field $\mathbb{K}$ is algebraically closed. It is well-known that a commutative family of endomorphisms of a finite-dimensional vector space can be simultaneously triangularized. In particular, if an abelian Lie algebra $\mathfrak{g}$ acts on a representation V, there exists a basis of V so that the action is given by upper triangular matrices. It turns out that the same is true for solvable Lie algebras, as in the following important result called **Lie's theorem**.

Theorem 11. *If a solvable Lie algebra $\mathfrak{g}$ acts on a finite-dimensional representation V (over an algebraically closed field), then there exists a basis of V so that the action is given by upper triangular matrices.*

Proof. Let us write $\{\phi_x\}_{x\in\mathfrak{g}} \in \text{End}(V)$ for the operators that comprise the action. It suffices to show that all the ϕ_x have a joint eigenvector $v \in V$, because then one can obtain the desired result by induction on $\dim V$ (we can obtain a full flag of subspaces of V which is preserved by the ϕ_x by taking a full flag of subspaces of $V/\mathbb{K}v$ and appending $\mathbb{K}v$ to it). We will prove the aforementioned claim by induction on $\dim \mathfrak{g}$. Since $\mathfrak{g}$ is solvable, we have

$$D\mathfrak{g} \subsetneq \mathfrak{g}$$

so the quotient $\mathfrak{g}/D\mathfrak{g}$ is a non-zero abelian Lie algebra. Therefore, any codimension 1 subspace in $\mathfrak{g}/D\mathfrak{g}$ is an ideal, and by the correspondence theorem we conclude that there exists a codimension 1 ideal $\mathfrak{i} \subset \mathfrak{g}$. By Proposition 11, $\mathfrak{i}$ is a solvable Lie algebra, and so the induction hypothesis implies that there exists $0 \neq v \in V$ such that

$$xv = \lambda(x)v \tag{87}$$

for all $x \in \mathfrak{i}$, where λ is a linear functional on $\mathfrak{i}$. Pick $y \in \mathfrak{g}\backslash\mathfrak{i}$ and consider the subspace

$$W = \text{span}\Big\{v, yv, y^2v, \dots\Big\}$$

For any $x \in \mathfrak{i}$ and $k \geq 0$, you can prove the elementary identity

$$xy^kv = y^kxv + \sum_{i=1}^{k} \binom{k}{i} y^{k-i} \underbrace{[\dots[[x,y],y],\dots,y]}_{i\ y\text{'s}} v \tag{88}$$

Since all the commutators in the right-hand side lie in $\mathfrak{i}$, the right-hand side of the expression above lies in W; we conclude that the action of any $x \in \mathfrak{i}$ preserves W. Even more so, x acts upper triangularly in the basis $v, yv, \dots, y^{\dim W-1}v$ of W, with $\lambda(x)$ on the diagonal. Therefore,

$$\text{tr}\,(x|_W) = \lambda(x)\dim W$$

for all $x \in \mathfrak{i}$. However, y also preserves W, and $[\mathfrak{i}, y] \subseteq \mathfrak{i}$ on account of $\mathfrak{i}$ being an ideal of $\mathfrak{g}$. The fact that commutators have trace 0 implies that

$$\lambda([x,y]) = 0 \tag{89}$$

for all $x \in \mathfrak{i}$. Let us now consider the non-zero subspace

$$W' = \Big\{v \in V \Big| x \cdot v = \lambda(x)v,\ \forall x \in \mathfrak{i}\Big\}$$

for the same linear functional as in (87). For any $x \in \mathfrak{i}$, $w \in W'$ and note that

$$x \cdot (y \cdot w) = y \cdot (x \cdot w) + [x, y] \cdot w = \lambda(x) y \cdot w + \lambda([x, y]) \cdot w$$

Since the second term in the right-hand side vanishes by (89), we conclude that $y \cdot w \in W'$. Thus, the action of y preserves W', so it has an eigenvector $w' \in W'$. This w' will be an eigenvector for the whole of $\mathfrak{g} = \mathfrak{i} \oplus \mathbb{K} y$.

□

Corollary 3. *All irreducible finite-dimensional representations (over an algebraically closed field) of a solvable Lie algebra are one-dimensional.*

Lecture 7

7.1

In the last lecture, we saw that if V is a representation of a solvable Lie algebra $\mathfrak{g}$ over an algebraically closed field $\mathbb{K}$, then the action homomorphism

$$\mathfrak{g} \to \mathfrak{gl}(V)$$

lands in the Lie subalgebra $\mathfrak{b} \subset \mathfrak{gl}(V)$ of upper triangular matrices (with respect to some basis). The analogous result for nilpotent Lie algebras and strictly upper triangular matrices is false (for example, V being a one-dimensional representation of a one-dimensional Lie algebra), but we have the following replacement. In what follows, we require $\operatorname{char} \mathbb{K} = 0$, but $\mathbb{K}$ needn't be algebraically closed.

Theorem 12. *If a Lie algebra $\mathfrak{g}$ acts on a finite-dimensional representation V by nilpotent operators, then there exists a basis of V so that the action is given by strictly upper triangular matrices.*

Proof. It suffices to show that all the operators $\{\phi_x\}_{x\in\mathfrak{g}}$ that comprise the action on V annihilate a non-zero vector $v \in V$, because then one can obtain the desired result by induction on the $\dim V$ (we can obtain a full flag of subspaces of V which is preserved by the ϕ_x by taking a full flag of subspaces of $V/\mathbb{K}v$ and appending $\mathbb{K}v$ to it). We will prove that the action of all $x \in \mathfrak{g}$ annihilate a common non-zero vector by induction on $\dim \mathfrak{g}$. First of all, we may replace $\mathfrak{g}$ by the image of the action homomorphism

$$\mathfrak{g} \to \mathfrak{gl}(V)$$

which allows us to assume $\mathfrak{g} \subseteq \mathfrak{gl}(V)$. Then let us assume that

$$\mathfrak{g} \text{ has a codimension 1 ideal } \mathfrak{i} \tag{90}$$

By the induction hypothesis for $\mathfrak{i}$, the subspace

$$W = \left\{ v \in V \middle| x \cdot v = 0, \ \forall x \in \mathfrak{i} \right\}$$

is non-zero. Fix $y \in \mathfrak{g}\backslash\mathfrak{i}$. Because

$$x \cdot (y \cdot w) = y \cdot (x \cdot w) + \underbrace{[x, y]}_{\in \mathfrak{i}} \cdot w = 0$$

for all $x \in \mathfrak{i}$ and $w \in W$, we conclude that the action of y sends W to itself. However, the action of y is via a nilpotent operator. Since a nilpotent operator always annihilates a non-zero vector, we conclude that there exists $0 \neq w \in W$ such that $y \cdot w = 0$. As $x \cdot w = 0$ for all $x \in \mathfrak{i}$, we conclude that the action of all elements of $\mathfrak{g}$ annihilates w.

Let us now explain why there exists a codimension 1 ideal $\mathfrak{i} \subset \mathfrak{g}$, thus justifying the assumption (90). We choose $\mathfrak{i}$ to be maximal proper Lie subalgebra of $\mathfrak{g}$, and assume that $\operatorname{codim} \mathfrak{i} > 1$. Consider the representation

$$\mathfrak{i} \curvearrowright \mathfrak{g}/\mathfrak{i}, \qquad x \cdot (y \bmod \mathfrak{i}) = ([x, y] \bmod \mathfrak{i})$$

The action above is by nilpotent operators, being a block of the action $\operatorname{ad} : \mathfrak{i} \curvearrowright \mathfrak{g}$, which is by nilpotent operators due to the assumption $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ (there is a subtlety to prove here: if $X \in \mathfrak{gl}(V)$ is a nilpotent operator, show that $\operatorname{ad}_X : \mathfrak{gl}(V) \to \mathfrak{gl}(V)$ is nilpotent, i.e. $\operatorname{ad}_X^n = 0$ for some n). The inductive hypothesis in the representation above implies that there exists $y \in \mathfrak{g}\backslash\mathfrak{i}$ such that $[\mathfrak{i}, y] \subseteq \mathfrak{i}$. This implies that $\mathfrak{i} \oplus \mathbb{K}y$ is a larger proper Lie subalgebra than $\mathfrak{i}$, which provides a contradiction. ∎

Theorem 12 implies the following result, commonly known as **Engel's theorem**.

Corollary 4. *A Lie algebra* $\mathfrak{g}$ *is nilpotent if and only if* $\mathrm{ad}_x \in \mathrm{End}(\mathfrak{g})$ *is nilpotent for all* $x \in \mathfrak{g}$.

The "only if" implication is straightforward. For the "if" implication, Theorem 12 applied to the adjoint representation gives us a flag of subspaces

$$0 = V_0 \subset V_1 \subset \cdots \subset V_n = \mathfrak{g}$$

so that $\mathrm{ad}_x(V_i) \subseteq V_{i-1}$ for all i. Therefore, for any $x_1, \ldots, x_n \in \mathfrak{g}$, we have

$$0 = \mathrm{ad}_{x_1} \circ \mathrm{ad}_{x_2} \circ \cdots \circ \mathrm{ad}_{x_{n-1}}(x_n) = [x_1, [x_2, \ldots, [x_{n-1}, x_n] \ldots]]$$

which precisely means that $\underbrace{[\mathfrak{g}, [\mathfrak{g}, \ldots, [\mathfrak{g}, \mathfrak{g}] \ldots]]}_{n \text{ copies of } \mathfrak{g}} = 0$.

7.2

By analogy with the situation of groups, we say that a Lie algebra is **simple** if it has no proper ideals, and it is not abelian. The reason for the latter restriction is that it ensures that

$$[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g} \tag{91}$$

as the ideal $[\mathfrak{g}, \mathfrak{g}]$ cannot be 0 if $\mathfrak{g}$ is not abelian. As a consequence, we have:

Corollary 5. *Any one-dimensional representation of a simple Lie algebra is 0.*

The following notion is a more general version of simplicity.

Definition 13. *A Lie algebra is called* ***semisimple*** *if it has no solvable ideals other than 0.*

Note that any simple Lie algebra is also semisimple. The reason is that its only ideals are 0 and $\mathfrak{g}$, but $\mathfrak{g}$ could not be solvable, because otherwise $[\mathfrak{g}, \mathfrak{g}]$ would be a proper ideal. The following notion, that of radical of a Lie algebra, measures how far a general Lie algebra is from being semisimple.

Proposition 12. *In any Lie algebra* $\mathfrak{g}$*, the sum of two solvable ideals (i.e. the smallest ideal containing the two) is solvable. Therefore, there exists a maximal solvable ideal called the* ***radical***

$$\boxed{\mathrm{rad}(\mathfrak{g}) \subseteq \mathfrak{g}}$$

and $\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$ *is semisimple.*

Proof. Let $\mathfrak{i}$ and $\mathfrak{j}$ be solvable ideals of $\mathfrak{g}$. The natural isomorphism

$$(\mathfrak{i} + \mathfrak{j})/\mathfrak{i} \cong \mathfrak{j}/(\mathfrak{i} \cap \mathfrak{j})$$

(itself an analogue of the second isomorphism theorem for Lie algebras) realizes $\mathfrak{i}+\mathfrak{j}$ as an extension of the solvable Lie algebras $\mathfrak{i}$ and a quotient of $\mathfrak{j}$. By Proposition 11, $\mathfrak{i}+\mathfrak{j}$ is solvable. Therefore, the maximal solvable ideal $\mathrm{rad}(\mathfrak{g})$ must be defined as the sum of all solvable ideals. The semisimplicity of $\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$ follows from the maximality of $\mathrm{rad}(\mathfrak{g})$ and the correspondence Theorem 2 (complete this argument).

□

The definition of the radical entails the fact that any Lie algebra admits a short exact sequence

$$0 \to \mathrm{rad}(\mathfrak{g}) \xrightarrow{\iota} \mathfrak{g} \xrightarrow{\pi} \mathfrak{g}_{ss} \to 0$$

where $\mathfrak{g}_{ss}$ is semisimple. In fact, an important result called **Levi's theorem** states that the above short exact sequence actually splits if the ground field has characteristic 0, i.e.

$$\exists \psi : \mathfrak{g}_{ss} \to \mathfrak{g} \quad \text{s.t.} \quad \pi \circ \psi = \mathrm{Id}_{\mathfrak{g}_{ss}} \tag{92}$$

This means that $\mathfrak{g}_{ss}$ can be perceived as a subalgebra of $\mathfrak{g}$, although not as an ideal.

7.3

By generalizing Corollary 3, any irreducible representation (over an algebraically closed field)

$$\mathfrak{g} \curvearrowright V$$

has the property that $\mathrm{rad}(\mathfrak{g})$ acts by scalars. Proof: as in the proof of Theorem 11, the subspace

$$W = \left\{ v \in V \middle| x \cdot v = \lambda(x) v, \forall x \in \mathrm{rad}(\mathfrak{g}) \right\}$$

is non-zero for some linear functional λ on $\mathrm{rad}(\mathfrak{g})$. However, you can prove by analogy with Theorem 11 that any $y \in \mathfrak{g} \backslash \mathrm{rad}(\mathfrak{g})$ also sends W to W; because V is irreducible, then $V = W$. As the ideal

$$[\mathfrak{g}, \mathrm{rad}(\mathfrak{g})] \subseteq \mathfrak{g} \tag{93}$$

acts on irreducible representations by 0, it is customary to quotient out this ideal from $\mathfrak{g}$. This naturally leads to the following.

Definition 14. *A Lie algebra $\mathfrak{g}$ is called **reductive** if $\mathrm{rad}(\mathfrak{g}) = \mathfrak{z}(\mathfrak{g})$.*

The defining condition of a reductive Lie algebra is actually equivalent to $\mathrm{rad}(\mathfrak{g}) \subseteq \mathfrak{z}(\mathfrak{g}) \quad \Leftrightarrow$ $[\mathfrak{g}, \mathrm{rad}(\mathfrak{g})] = 0$, because the opposite inclusion $\mathfrak{z}(\mathfrak{g}) \subseteq \mathrm{rad}(\mathfrak{g})$ is true in any Lie algebra (prove this claim). With this in mind, Levi's theorem (92) implies that a reductive Lie algebra splits as

$$\mathfrak{g} \cong \mathfrak{z}(\mathfrak{g}) \oplus \mathfrak{g}_{ss}$$

In this case, $\mathfrak{g}_{ss}$ is actually an ideal of $\mathfrak{g}$, and the direct sum in the RHS is in the sense of Definition 4. Thus, we see that reductive Lie algebras are obtained from semisimple ones by adding a center. The main example is (38), in which the general linear Lie algebra (reductive) is the direct sum of the special linear Lie algebra (semisimple, as we will shortly see) and a one-dimensional center.

7.4

Semisimple and reductive Lie algebras can be described in terms of their bilinear forms, as per the following notion.

Definition 15. *If $\mathfrak{g}$ is a Lie algebra over a field $\mathbb{K}$, then a symmetric bilinear form*

$$\mathfrak{g} \times \mathfrak{g} \xrightarrow{(\cdot,\cdot)} \mathbb{K}$$

is called ***invariant*** *if*

$$([x, y], z) + ([x, z], y) = 0 \tag{94}$$

for all $x, y, z \in \mathfrak{g}$. We will write ***s.i.b.f.*** *for a symmetric invariant bilinear form.*

Condition (94) is nice because the orthogonal complement of an ideal with respect to a s.i.b.f. is also an ideal. The following is a great source of s.i.b.f.'s (we leave this claim as an exercise to you).

Proposition 13. *For any representation $\mathfrak{g} \curvearrowright V$, the assignment*

$$(x, y)_V = \mathrm{tr}_V (\phi_x \circ \phi_y) \tag{95}$$

is a s.i.b.f.

For example, if V is the usual n-dimensional representation of $\mathfrak{gl}_n$, then $(X, Y)_V = \mathrm{tr}(XY)$. For a general Lie algebra $\mathfrak{g}$, a special role is played by setting V to be the adjoint representation, in which case the s.i.b.f.

$$(x, y)_{\mathfrak{g}} = \mathrm{tr}_{\mathfrak{g}}(\mathrm{ad}_x \circ \mathrm{ad}_y) \tag{96}$$

is called the **Killing form**.

Theorem 13. *If the s.i.b.f.* (95) *is non-degenerate for some representation V, then $\mathfrak{g}$ is reductive.*

Proof. There exists a flag of subrepresentations

$$0 = V_0 \subset V_1 \subset \cdots \subset V_k = V$$

with V_i/V_{i-1} irreducible for all $i \in \{1, \dots, k\}$. The fact that the matrices ϕ_x are all block upper triangular with respect to the flag above implies that

$$(x, y)_V = \sum_{i=1}^{k} (x, y)_{V_i/V_{i-1}}$$

for all $x, y \in \mathfrak{g}$. If we take $x \in [\mathfrak{g}, \mathrm{rad}(\mathfrak{g})]$, then we have already seen in (93) that such x act by 0 in all irreducible representations, and thus lie in the kernel of $(\cdot, \cdot)_V$. Since the latter is assumed non-degenerate, this implies $x = 0$, as desired.

□

As a consequence of this Theorem, all the matrix Lie algebras that we encountered in this course ($\mathfrak{gl}_n, \mathfrak{sl}_n, \mathfrak{o}_n, \mathfrak{sp}_{2n}, \mathfrak{u}(n), \mathfrak{su}(n)$) are reductive, which you can prove by showing that the s.i.b.f. given by their usual matrix representation is non-degenerate.

7.5

We have just seen that the s.i.b.f.'s (95) give a criterion for a Lie algebra being reductive. We will now see that the Killing form is even more powerful, as evidenced by the following result, typically called **Cartan's criterion of solvability/semisimplicity**, which holds over any field of characteristic 0.

Theorem 14. *(a) A Lie algebra $\mathfrak{g}$ is solvable if and only if*

$$([x,y],z)_{\mathfrak{g}} = 0, \quad \forall x,y,z \in \mathfrak{g}$$

(b) A Lie algebra is semisimple if and only if its Killing form is non-degenerate.

We will prove Theorem 14 using certain tools that will be developed in next lecture. But let us apply it to $\mathfrak{gl}_n = \text{scalar matrices} \oplus \mathfrak{sl}_n$: clearly scalar matrices are in the kernel of the Killing form, which shows that $\mathfrak{gl}_n$ is not semisimple. On the other hand, if we let V be the usual n-dimensional representation of $\mathfrak{gl}_n$, we have

$$(E_{ij}, E_{i'j'})_V = \mathrm{tr}_V(E_{ij}E_{i'j'}) = \delta_{i'j}\delta_{ij'}$$

which is clearly non-degenerate. Thus, we conclude that $\mathfrak{gl}_n$ is reductive. The bilinear form above is also non-degenerate for $\mathfrak{sl}_n$, but because the latter Lie algebra is simple, we actually have the following result.

Lemma 2. *Any two non-degenerate s.i.b.f.'s on a simple Lie algebra (over an algebraically closed field) are proportional.*

Proof. It is easy to check that any non-degenerate s.i.b.f. on a Lie algebra $\mathfrak{g}$ gives an isomorphism

$$\boxed{\mathfrak{g} \cong \mathfrak{g}^*}$$

of representations of the Lie algebra $\mathfrak{g}$ (where the LHS is the adjoint representation and the RHS is the dual of the adjoint representation). The fact that $\mathfrak{g}$ is simple is equivalent to the adjoint representation being irreducible, and then the fact that any two isomorphisms $\mathfrak{g} \cong \mathfrak{g}^*$ differ by a constant multiple is a consequence of Schur's Lemma.

□

7.6

Let us give without proof certain connections between the Killing form and real forms of complex Lie algebras, as in Lecture 4.

Theorem 15. *(a) For any compact real Lie group G, its Lie algebra $\mathfrak{g}$ is reductive and its Killing form is negative-semidefinite (the kernel of the form is just $\mathfrak{z}(\mathfrak{g})$).*

(b) As a partial converse, if $\mathfrak{g}$ is a real Lie algebra with a negative-definite Killing form, then any connected Lie group with Lie algebra $\mathfrak{g}$ is compact.

The situation of real Lie algebras with positive-definite Killing form is much simpler and less interesting than the above Theorem: there are no such Lie algebras except 0. To see this, let $\mathfrak{g}$ be a real Lie algebra, and pick an orthonormal basis

$$\mathfrak{g} = \bigoplus_{i=1}^{n} \mathbb{R} \cdot x_i$$

for the Killing form. If we let γ_{ij}^k be the structure constants for the Lie bracket, i.e.

$$[x_i, x_j] = \mathrm{ad}_{x_i}(x_j) = \sum_k \gamma_{ij}^k x_k$$

then we have for all $i \in \{1, \ldots, n\}$

$$(x_i, x_i)_{\mathfrak{g}} = \mathrm{tr}_{\mathfrak{g}}(\mathrm{ad}_{x_i} \circ \mathrm{ad}_{x_i}) = \sum_{j,k=1}^{n} \gamma_{ij}^k \gamma_{ik}^j \tag{97}$$

However, the identity

$$\gamma_{ij}^k = ([x_i, x_j], x_k)_{\mathfrak{g}} = -([x_i, x_k], x_j)_{\mathfrak{g}} = -\gamma_{ik}^j \tag{98}$$

implies that the right-hand side of (97) is non-positive, which contradicts the positive-definiteness of the Killing form.

Lecture 8

8.1

Semisimple Lie algebras have two important features: complete reducibility of representations, and Jordan decompositions, both of which we will now state and prove. The ground field $\mathbb{K}$ must have characteristic 0 throughout the entire lecture.

Theorem 16. *If V is any finite-dimensional representation of a semisimple Lie algebra $\mathfrak{g}$, then for any subrepresentation $W \subset V$ there exists another subrepresentation $W' \subset V$ such that*

$$V = W \oplus W' \tag{99}$$

*After repeated applications of this result, we conclude that any representation of a semisimple Lie algebra is a direct sum of irreducible representations, which is called **complete reducibility**.*

The semisimplicity assumption is key, as we have already seen in (34) an example where complete reducibility fails. In fact, that example failed because the action was given in terms of matrices which are not diagonalizable, which leads us into a discussion of Jordan decompositions. Let us start with the following result, which we leave to you.

Proposition 14. *A linear transformation $f : V \to V$ is called **semisimple** if for any subspace $W \subset V$ preserved by f there exists another subspace preserved by f such that*

$$V = W \oplus W'$$

If V is a finite-dimensional vector space over an algebraically closed field $\mathbb{K}$, then f is semisimple if and only if f is diagonalizable.

The well-known Jordan decomposition basically says that any linear transformation f on a finite-dimensional vector space can be uniquely decomposed as

$$f = f_{ss} + f_n \tag{100}$$

where f_{ss} is semisimple, f_n is nilpotent, and $f_{ss} f_n = f_n f_{ss}$. Moreover, f_{ss} and f_n are polynomials in f with zero constant term. We will soon show that the decomposition (100) extends uniformly from individual linear transformations to elements of semisimple Lie algebras, as follows.

Theorem 17. *If $\mathfrak{g}$ is a semisimple Lie algebra, then any $x \in \mathfrak{g}$ admits a unique decomposition*

$$x = x_{ss} + x_n \tag{101}$$

such that $[x_{ss}, x_n] = 0$, and x_{ss} (respectively x_n) acts as a semisimple (respectively nilpotent) operator in any representation of $\mathfrak{g}$. This is true in particular in the adjoint representation, so

$$(\mathrm{ad}_x)_{ss} = \mathrm{ad}_{x_{ss}} \quad \text{and} \quad (\mathrm{ad}_x)_n = \mathrm{ad}_{x_n} \tag{102}$$

*One calls (101) the **abstract** Jordan decomposition.*

Formula (102) is very strong. For example, we claim that it implies that

$$[x, y] = 0 \quad \Leftrightarrow \quad [x_{ss}, y] = [x_n, y] = 0 \tag{103}$$

for any $x, y \in \mathfrak{g}$. To see this, the fact that $(\mathrm{ad}_x)_{ss} = \mathrm{ad}_{x_{ss}}$ implies that the latter operator is a polynomial in ad_x with zero constant term. If $[x, y] = 0$, then the aforementioned polynomial annihilates y, which implies that $[x_{ss}, y] = 0$ (hence also $[x_n, y] = 0$).

8.2

In the remainder of this Lecture, we will prove the results above. For technical reasons, we will start with the Cartan criteria of solvability and semisimplicity.

Proof. of Theorem 14: (a) Since a Lie algebra over $\mathbb{K}$ is solvable if and only if its extension over the algebraic closure of $\mathbb{K}$ is solvable (the property of (83) terminating with 0 is unchanged by field extension), then we will assume the ground field is algebraically closed. In this case, the "only if" statement is an easy case of Lie's theorem 11 for the adjoint representation: indeed, the fact that ad_z is upper triangular for all $z \in \mathfrak{g}$ implies that $\mathrm{ad}_{[x,y]}$ is strictly upper triangular for all $x, y \in \mathfrak{g}$. Therefore, $([x,y],z)_{\mathfrak{g}}$ is the trace of a strictly upper triangular matrix, and thus equal to 0.

Lemma 3. *If a Lie subalgebra $\mathfrak{a} \subseteq \mathfrak{gl}_n$ has the property that*

$$\mathrm{tr}(xy) = 0, \quad \forall x, y \in \mathfrak{a} \tag{104}$$

then $\mathfrak{a}$ is solvable.

Let us first deduce the "if" statement from Lemma 3. By Proposition 11, it suffices to show that $\mathfrak{g}/\mathfrak{z}(\mathfrak{g})$ is solvable. So by replacing $\mathfrak{g}$ with $\mathfrak{g}/\mathfrak{z}(\mathfrak{g})$, we may assume that the adjoint representation provides an injection

$$\mathfrak{g} \hookrightarrow \mathrm{End}(\mathfrak{g})$$

Therefore, the Lie subalgebra $\mathfrak{a} = [\mathfrak{g}, \mathfrak{g}]$ satisfies the hypotheses of Lemma 3, and is thus solvable. Since $\mathfrak{g}/\mathfrak{a}$ is abelian and thus solvable, then Proposition 11 implies that $\mathfrak{g}$ is solvable.

It remains to prove Lemma 3. It suffices to show that $[x,y]$ is a nilpotent matrix for all $x, y \in \mathfrak{a}$, because then the first blue claim of Lecture 7 and Corollary 4 would imply that $[\mathfrak{a}, \mathfrak{a}]$ is nilpotent (hence solvable, hence Proposition 11 would imply that $\mathfrak{a}$ is solvable). Thus, let us pick arbitrary $x, y \in \mathfrak{a}$ and assume that the eigenvalues of $[x,y]$ are $\lambda_1, \dots, \lambda_n$ counted with multiplicities (we are still working over the algebraic closure of the ground field). Our goal is to show that $\lambda_1 = \dots = \lambda_n = 0$, so assume that at least one of the λ_i's is non-zero. Then it is easy to see that there exists a $\mathbb{Q}$-linear functional

$$\zeta : \mathrm{span}_{\mathbb{Q}}(\lambda_1, \dots, \lambda_n) \to \mathbb{Q}$$

such that

$$\sum_{i=1}^{n} \lambda_i \zeta(\lambda_i) \neq 0 \tag{105}$$

(if this ζ seems strange to you, then just assume we are working over $\mathbb{C}$ and we replace $\zeta(\lambda_i)$ by $\overline{\lambda}_i$ from now on). Because ζ is linear, there exists an interpolation polynomial $P(t)$ such that

$$P(\lambda_i - \lambda_j) = \zeta(\lambda_i) - \zeta(\lambda_j) \tag{106}$$

for all $i, j \in \{1, \dots, n\}$. Let

$$A = [x,y]_{ss} = \mathrm{diag}(\lambda_1, \dots, \lambda_n) \quad \Rightarrow \quad \mathrm{ad}_A = \mathrm{diag}(\lambda_i - \lambda_j)_{1 \le i,j \le n}$$

and $B = \mathrm{diag}(\zeta(\lambda_1), \dots, \zeta(\lambda_n))$. Therefore, (106) implies that

$$\mathrm{ad}_B = \mathrm{diag}(\zeta(\lambda_i) - \zeta(\lambda_j))_{1 \le i,j \le n} = P(\mathrm{ad}_A)$$

Moreover, as $\mathrm{ad}_A = \mathrm{ad}_{[x,y]_{ss}} = \left(\mathrm{ad}_{[x,y]}\right)_{ss}$, we conclude that ad_A is itself a polynomial in $\mathrm{ad}_{[x,y]}$. This implies that $\mathrm{ad}_B = Q(\mathrm{ad}_{[x,y]})$ for some polynomial $Q(t)$, and thus

$$\mathrm{ad}_B(\mathfrak{a}) = Q(\mathrm{ad}_{[x,y]})(\mathfrak{a}) \subseteq \mathfrak{a}$$

With this in mind, we have for all $x, y \in \mathfrak{a}$

$$\sum_{i=1}^{n} \lambda_i \zeta(\lambda_i) = \mathrm{tr}(AB) = \mathrm{tr}([x,y]B) = \mathrm{tr}(\mathrm{ad}_B(x)y) = 0$$

with the last equality being precisely the hypothesis (104). We have therefore contradicted (105).

(b) For the "only if" statement, let us consider the kernel $\mathfrak{i}$ of the Killing form of a semisimple Lie algebra $\mathfrak{g}$. By (94), $\mathfrak{i}$ is an ideal of $\mathfrak{g}$. Moreover, for any $x, y, z \in \mathfrak{i}$

$$\mathrm{ad}_{[x,y]} \circ \mathrm{ad}_z$$

calculated in $\mathfrak{g}$ is a block matrix, with one of the blocks being given by the same composition but calculated in $\mathfrak{i}$, and the other block being 0. Thus, the fact that $([x,y],z)_{\mathfrak{g}} = 0$ implies that $([x,y],z)_{\mathfrak{i}} = 0$, for all $x, y, z \in \mathfrak{i}$. Part (a) implies that $\mathfrak{i}$ is solvable, and by the definition of semisimplicity we conclude that $\mathfrak{i} = 0$ (i.e. the Killing form is non-degenerate).

Let us now prove the "if" statement, which relies on the following claim (which we leave to you, using the fact the last non-zero term in the derived series of $\mathrm{rad}(\mathfrak{g})$ would be an abelian ideal of $\mathfrak{g}$).

Lemma 4. *A Lie algebra $\mathfrak{g}$ is semisimple if and only if it has no non-zero abelian ideals.*

Since any abelian ideal $\mathfrak{i}$ of a Lie algebra $\mathfrak{g}$ would lie in the kernel of the Killing form (because ad_x sends $\mathfrak{g}$ to $\mathfrak{i}$ and $\mathfrak{i}$ to 0 for all $x \in \mathfrak{i}$), then the non-degeneracy of the Killing form implies the non-existence of abelian ideals other than 0. □

8.3

Cartan's criterion for semisimplicity has a number of interesting consequences. For one thing, a real Lie algebra is semisimple if and only if its complexification is semisimple (note that this does not hold for simple Lie algebras). More important is the following characterization of Lie algebras.

Lemma 5. *A Lie algebra $\mathfrak{g}$ is semisimple if and only if it is a direct sum of simple Lie algebras.*

Proof. We already showed that simple Lie algebras are semisimple, and so their Killing forms are non-degenerate by Theorem 14. Since the Killing form of a direct sum of Lie algebras is the sum of the Killing forms of its constituents, we conclude that any direct sum of simple algebras has non-degenerate Killing form, hence is semisimple by Theorem 14.

For the opposite direction, consider a semisimple Lie algebra $\mathfrak{g}$ which is not simple. Therefore, if has a proper ideal $\mathfrak{i} \subset \mathfrak{g}$. If we let $\mathfrak{j}$ be the complement of $\mathfrak{i}$ with respect to the Killing form, then show that the non-degeneracy of the latter implies that

$$\mathfrak{g} = \mathfrak{i} \oplus \mathfrak{j}$$

and the invariance of the Killing form implies that $\mathfrak{j}$ is an ideal. Therefore, the decomposition above is a direct sum of Lie algebras, in the sense of Definition 4. Since $\mathfrak{i}$ and $\mathfrak{j}$ are semisimple Lie algebras in turn, we can repeat this algorithm by induction on $\dim \mathfrak{g}$.

□

As a consequence of Lemma 5, we have that

$$[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g} \tag{107}$$

for any semisimple Lie algebra $\mathfrak{g}$, because the property above holds for simple Lie algebras (we saw this in (91)) and it is preserved under direct sums. Therefore, we conclude the following.

Corollary 6. *Any one-dimensional representation of a semisimple Lie algebra is 0.*

8.4

We will now prove Theorem 16 on complete reducibility. The following notion is key.

Proposition 15. *For any non-degenerate s.i.b.f.* $(\cdot, \cdot)$ *on a Lie algebra* $\mathfrak{g}$*, its Casimir element is defined as*

$$C = \sum_i x_i \otimes x^i \in U\mathfrak{g} \tag{108}$$

where $\{x_i\}, \{x^i\}$ *run over dual bases of* $\mathfrak{g}$ *with respect to the form. Then* $C \in Z(U\mathfrak{g})$*.*

By basic algebra, C does not depend on the choice of dual bases $\{x_i\}, \{x^i\}$. If $(\cdot, \cdot)_{\mathfrak{g}}$ is the Killing form, then the corresponding $C_{\mathfrak{g}}$ is usually called the Casimir element of $\mathfrak{g}$. The potential abuse of terminology is basically non-existent for a simple Lie algebra, because Lemma 2 implies that any two Casimir elements are proportional. Because any representation $\mathfrak{g} \curvearrowright V$ is also a representation of $U\mathfrak{g}$, the Casimir element acts in V; we have already encountered this operator for $\mathfrak{sl}_2$ in Subsection 5.4.

Proof. of Proposition 15: Let us assume $x_i = x^i$ is an orthonormal basis with respect to the given symmetric invariant bilinear form. If we let

$$[x_i, x_j] = \sum_k \gamma_{ij}^k x_k$$

then we have $\gamma_{ik}^j = -\gamma_{ki}^j \overset{(98)}{=} \gamma_{kj}^i = -\gamma_{jk}^i$. Therefore, we have for any k

$$C \otimes x_k - x_k \otimes C = \sum_i \Big([x_i, x_k] \otimes x_i + x_i \otimes [x_i, x_k] \Big) = \sum_{i,j} \gamma_{ik}^j \Big(x_j \otimes x_i + x_i \otimes x_j \Big) = 0$$

□

8.5

Because any Casimir element is central, Schur's Lemma implies that it acts by a constant in any irreducible representation. More specifically, for an irreducible representation $\mathfrak{g} \curvearrowright V$, the Casimir element defined with respect to the symmetric invariant bilinear form (95) acts by the constant

$$\frac{\dim \mathfrak{g}}{\dim V} \tag{109}$$

(note that this requires the form $(\cdot, \cdot)_V$ to be non-degenerate; if on the other hand this form has a non-trivial kernel $\mathfrak{i} \subset \mathfrak{g}$, then we simply replace $\mathfrak{g}$ by $\mathfrak{g}/\mathfrak{i}$ in (108)). Even more generally, it is a fact that any Casimir element of a semisimple Lie algebra acts by a non-zero scalar in any irreducible finite-dimensional representation, but we will not need this.

Proof. of Theorem 16: Let us first consider the case when $W \subset V$ has codimension 1. If W is irreducible, then the discussion above shows that there is a Casimir element C which acts by a non-zero constant on W. By Corollary 6, the action of $\mathfrak{g}$ on the one-dimensional quotient V/W is 0, so we conclude that C sends V to W. Thus, $W' = \text{Ker } C|_V$ is non-empty, and because $C|_W$ is a non-zero scalar, we conclude that W' is one-dimensional and complementary to W. The fact that W' is a subrepresentation follows from

$$C \cdot (x \cdot v) = x \cdot (C \cdot v) \quad \Rightarrow \quad x \cdot W' \subseteq W'$$

for all $x \in \mathfrak{g}$, where the first equality is due to the fact that C is central. Having proved the complete reducibility when W is irreducible (but still codimension 1 inside V), let us now prove the case of general W (but still codimension 1 inside V) by induction on dim V. If W is not irreducible, then we may consider a maximal proper subrepresentation $\bar{W} \subsetneq W$ (which exists due to finite-dimensionality) and simply run the discussion above for the representation $\mathfrak{g} \curvearrowright V/\bar{W}$ and its codimension 1 subrepresentation $W/\bar{W}$. We obtain a subrepresentation $\bar{W} \subset \bar{W}' \subset V$ such that

$$V/\bar{W} = W/\bar{W} \oplus \bar{W}'/\bar{W}$$

with $\bar{W}$ having codimension 1 inside $\bar{W}'$. Repeating the argument above gives us a decomposition $\bar{W}' = \bar{W} \oplus W'$ for some one-dimensional subrepresentation $W' \subset \bar{W}'$, and so $V = W \oplus W'$ is the required decomposition in (99).

Having proved the Theorem for any subrepresentation W of codimension 1 in V, let us consider an arbitrary subrepresentation $W \subset V$ and define

$$\widetilde{V} = \left\{ f : V \to W \middle| f|_W = \text{scalar} \right\}$$
$$\widetilde{W} = \left\{ f : V \to W \middle| f|_W = 0 \right\}$$

We may make $\widetilde{V}$ into a representation of $\mathfrak{g}$ via $(x \cdot f)(v) = x \cdot f(v) - f(x \cdot v)$, and it is clear that

$$x \cdot \widetilde{V} = \widetilde{W}$$

for all $x \in \mathfrak{g}$. As $\widetilde{W}$ has codimension 1 inside $\widetilde{V}$, the first part of this proof implies the existence of

$$f \in \widetilde{V} \backslash \widetilde{W}, \quad \text{s.t. } x \cdot f = \lambda(x) f, \ \forall x \in \mathfrak{g}$$

where $\lambda : \mathfrak{g} \to \mathbb{K}$ is some linear functional. Then, $f|_W = \alpha \cdot \text{Id}_W$ for some non-zero α, so $W' = \text{Ker } f$ is a complementary subspace to W inside V. Moreover, W' is a subrepresentation because

$$f(v) = 0 \quad \Rightarrow \quad f(x \cdot v) = x \cdot f(v) - (x \cdot f)(v) = x \cdot f(v) - \lambda(x) f(v) = 0$$

□

8.6

Let us now prove Theorem 17 on the abstract Jordan decomposition in semisimple Lie algebras.

Proof. of Theorem 17: We begin by claiming that for any semisimple subalgebra

$$\mathfrak{g} \subseteq \mathfrak{gl}(V)$$

the semisimple and nilpotent parts of any element $x \in \mathfrak{g}$ (regarded as linear transformations $V \to V$, see (100)) also lie in $\mathfrak{g}$. Applying this result to the faithful adjoint representation

$$\mathfrak{g} \hookrightarrow \mathrm{End}(\mathfrak{g})$$

(faithfulness, i.e. the injectivity of the Lie algebra homomorphism above, is due to semisimple Lie algebras having trivial center) allows us to construct a decomposition

$$x = x_{ss} + x_n \tag{110}$$

in $\mathfrak{g}$ such that (102) holds. Let us now consider an arbitrary representation $\mathfrak{g} \curvearrowright V$ and the associated Lie algebra homomorphism

$$\mathfrak{g} \xrightarrow{\phi} \mathfrak{gl}(V)$$

We want to show that for any $x \in \mathfrak{g}$, we have

$$\phi(x)_{ss} = \phi(x_{ss}) \quad \text{and} \quad \phi(x)_n = \phi(x_n)$$

i.e. the abstract Jordan decomposition gives rise to the usual Jordan decomposition in V. By replacing $\mathfrak{g}$ with $\mathrm{Im}\ \phi$ (which is also semisimple on account of it being a quotient of a semisimple Lie algebra) we may regard $\mathfrak{g}$ as a subalgebra of $\mathfrak{gl}(V)$. Then any $x \in \mathfrak{g}$ admits a Jordan decomposition in the context of linear transformations $V \to V$

$$x = x'_{ss} + x'_n \tag{111}$$

where the claim at the beginning of the proof establishes the fact that $x'_{ss}, x'_n \in \mathfrak{g}$. You have already shown last week that if $y \in \mathfrak{gl}(V)$ is nilpotent, then ad_y is nilpotent; it is also true that if $y \in \mathfrak{gl}(V)$ is semisimple, then ad_y is semisimple (one of the exercises on this week's exercise sheet will essentially prove this over an algebraically closed field). Therefore,

$$\mathrm{ad}_x = \mathrm{ad}_{x'_{ss}} + \mathrm{ad}_{x'_n} \tag{112}$$

is a Jordan decomposition in $\mathrm{End}(\mathfrak{gl}(V))$. Since the semisimple and nilpotent parts of any operator are polynomials in said operator, then $\mathrm{ad}_{x'_{ss}}$ and $\mathrm{ad}_{x'_n}$ send $\mathfrak{g}$ to $\mathfrak{g}$. It is easy to see that a diagonal block of a semisimple/nilpotent linear transformation is also semisimple/nilpotent, and so the restriction of (112) to $\mathfrak{g}$ is also a Jordan decomposition. By uniqueness of the latter, we conclude that $\mathrm{ad}_{x'_{ss}} = \mathrm{ad}_{x_{ss}}$ and $\mathrm{ad}_{x'_n} = \mathrm{ad}_{x_n}$ as linear transformations of $\mathfrak{g}$. Therefore, the faithfulness of the adjoint representation of a semisimple Lie algebra implies that the decompositions (110) and (111) coincide, i.e. the abstract Jordan decomposition agrees with the Jordan decomposition in V.

To prove the claim at the beginning of this proof, consider the following Lie subalgebras of $\mathfrak{gl}(V)$

- $\mathfrak{a} = \left\{ f : V \to V \middle| [f, \mathfrak{g}] \subseteq \mathfrak{g} \right\}$
- $\mathfrak{b}_W = \left\{ f : V \to V \middle| f(W) \subseteq W \text{ and } \mathrm{tr}(f|_W) = 0 \right\}$ for any $\mathfrak{g}$-invariant subspace $W \subset V$

It is almost obvious to see that

$$\mathfrak{g} \subseteq \mathfrak{g}' := \mathfrak{a} \bigcap_{\mathfrak{g}\text{-invariant } W \subseteq V} \mathfrak{b}_W$$

(the fact that any $x \in \mathfrak{g}$ acts on any $\mathfrak{g}$-invariant subspace $W \subset V$ by a traceless matrix is due to the fact that $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ for a semisimple Lie algebra $\mathfrak{g}$). Consider the Jordan decomposition (100)

$$x = x_{ss} + x_n$$

of any $x \in \mathfrak{g}$, where $x_{ss}, x_n : V \to V$. Because x_{ss} and x_n are polynomial expressions in x, we have

$$x_{ss}, x_n \in \mathfrak{b}_W$$

for any $\mathfrak{g}$-invariant $W \subseteq V$ (a little thought is needed to see that x_{ss} and x_n are traceless on W). Moreover, as we showed in the previous paragraph, $\mathrm{ad}_x = \mathrm{ad}_{x_{ss}} + \mathrm{ad}_{x_n}$ is a Jordan decomposition in $\mathrm{End}(\mathfrak{gl}(V))$. Therefore, $\mathrm{ad}_{x_{ss}}$ and ad_{x_n} are polynomial expressions in ad_x, which implies that

$$x_{ss}, x_n \in \mathfrak{a}$$

Putting the above two displays together implies that

$$x_{ss}, x_n \in \mathfrak{g}' \tag{113}$$

By the complete reducibility Theorem 16, the adjoint representation of $\mathfrak{g}$ on $\mathfrak{g}'$ decomposes as

$$\mathfrak{g}' = \mathfrak{g} \oplus S$$

for some $S \subseteq \mathfrak{gl}(V)$. Because $\mathfrak{g}$ is tautologically an ideal of $\mathfrak{a}$, we conclude that $\mathfrak{g}$ is an ideal of $\mathfrak{g}'$, and so $[\mathfrak{g}, S] = 0$. If we decompose V as a direct sum of irreducible representations of $\mathfrak{g}$

$$V = V_1 \oplus \cdots \oplus V_k$$

then any $f \in S$ is on one hand a $\mathfrak{g}$-intertwiner, while on the other hand f preserves each V_i and acts tracelessly on it. By Schur's lemma, the only option is that any $f \in S$ is actually 0 and so $\mathfrak{g}' = \mathfrak{g}$. Then (113) implies the claim at the beginning of this proof. □

8.7

The Jordan decomposition in the adjoint representation (102) can also be constructed as follows.

Definition 16. *A **derivation** of a Lie algebra $\mathfrak{g}$ is a linear transformation*

$$\zeta : \mathfrak{g} \to \mathfrak{g}$$

which satisfies the following version of the Leibniz rule

$$\zeta([y, z]) = [\zeta(y), z] + [y, \zeta(z)]$$

for all $y, z \in \mathfrak{g}$.

For any $x \in \mathfrak{g}$, the Jacobi identity implies that

$$\xi_x(y) = [x, y] \tag{114}$$

is a derivation. Such derivations are called **inner**, and any other derivation is called **outer**.

Lemma 6. *A semisimple Lie algebra only has inner derivations.*

Proof. The set $\mathrm{Der}(\mathfrak{g})$ of derivations of $\mathfrak{g}$ is itself a Lie algebra, with respect to the Lie bracket

$$[\zeta, \zeta'](x) = \zeta(\zeta'(x)) - \zeta'(\zeta(x))$$

and it is easy to see that the function

$$\mathfrak{g} \to \mathrm{Der}(\mathfrak{g}), \quad x \rightsquigarrow \xi_x \tag{115}$$

is a Lie algebra homomorphism. Because a semisimple Lie algebra $\mathfrak{g}$ has trivial center, the function above is injective. The fact that

$$[\xi_x, \zeta] = \xi_{\zeta(x)}, \quad \forall \zeta \in \mathrm{Der}(\mathfrak{g}), x \in \mathfrak{g} \tag{116}$$

implies that the injection (115) identifies $\mathfrak{g}$ with an ideal of $\mathrm{Der}(\mathfrak{g})$. Moreover, this injection identifies the Killing form on $\mathfrak{g}$ with the one on $\mathrm{Der}(\mathfrak{g})$. The non-degeneracy of the Killing form on $\mathfrak{g}$ means that we have a direct sum decomposition

$$\mathrm{Der}(\mathfrak{g}) = \mathfrak{g} \oplus \mathfrak{i}$$

(as in the proof of Lemma 5), where $\mathfrak{i}$ is an ideal. Because $[\mathfrak{g}, \mathfrak{i}] = 0$, for any $\zeta \in \mathfrak{i}$ we have by (116)

$$\xi_{\zeta(x)} = 0, \forall x \in \mathfrak{g} \quad \Rightarrow \quad \zeta(x) = 0, \forall x \in \mathfrak{g}$$

This shows that $\mathfrak{i} = 0$, and so every derivation is inner.

□

Let us now consider a Lie algebra $\mathfrak{g}$ over an algebraically closed ground field $\mathbb{K}$, and establish (102). For any $x \in \mathfrak{g}$, take the generalized eigenspace decomposition of the operator $\mathrm{ad}_x : \mathfrak{g} \to \mathfrak{g}$

$$\mathfrak{g} = \bigoplus_{\gamma \in \mathbb{K}} \mathfrak{g}_\gamma$$

where

$$\mathfrak{g}_\gamma = \left\{ y \in \mathfrak{g} \middle| (\mathrm{ad}_x - \gamma \cdot \mathrm{Id}_{\mathfrak{g}})^N(y) = 0 \text{ for } N \gg 0 \right\}$$

Define $(\mathrm{ad}_x)_{ss}$ as the operator which acts on $\mathfrak{g}_\gamma$ as multiplication by γ. We claim that

$$\zeta : \mathfrak{g} \to \mathfrak{g}, \quad \zeta(y) = (\mathrm{ad}_x)_{ss}(y) \tag{117}$$

is a derivation. To see this, take any $y \in \mathfrak{g}_\gamma$ and $z \in \mathfrak{g}_\delta$, and note that

$$(\mathrm{ad}_x - (\gamma + \delta) \cdot \mathrm{Id}_{\mathfrak{g}})([y, z]) = [(\mathrm{ad}_x - \gamma \cdot \mathrm{Id}_{\mathfrak{g}})(y), z] + [y, (\mathrm{ad}_x - \delta \cdot \mathrm{Id}_{\mathfrak{g}})(z)] \quad \Rightarrow$$

$$\Rightarrow \quad (\mathrm{ad}_x - (\gamma+\delta)\cdot \mathrm{Id}_{\mathfrak{g}})^N([y,z]) = \sum_{N_1+N_2=N} [(\mathrm{ad}_x - \gamma\cdot \mathrm{Id}_{\mathfrak{g}})^{N_1}(y), (\mathrm{ad}_x - \delta\cdot \mathrm{Id}_{\mathfrak{g}})^{N_2}(z)] = 0$$

if N is large enough. Thus, we conclude that $[y,z] \in \mathfrak{g}_{\gamma+\delta}$, which immediately implies that (117) is a derivation. By Lemma 6, $\zeta = \xi_{x_{ss}}$ for some $x_{ss} \in \mathfrak{g}$ and thus

$$(\mathrm{ad}_x)_{ss} = \mathrm{ad}_{x_{ss}} \tag{118}$$

If we let $x_n = x - x_{ss}$, we conclude that

$$(\mathrm{ad}_x)_n = \mathrm{ad}_{x_n} \tag{119}$$

The fact that x_{ss} and x_n commute follows from the fact that

$$0 = [(\mathrm{ad}_x)_{ss}, (\mathrm{ad}_x)_n] = [\mathrm{ad}_{x_{ss}}, \mathrm{ad}_{x_n}] = \mathrm{ad}_{[x_{ss},x_n]} \quad \Rightarrow \quad 0 = [x_{ss}, x_n]$$

The last implication is due to the fact that the adjoint representation is faithful (has zero kernel) for semisimple Lie algebras.

Lecture 9

9.1

Having developed foundational results on semisimple Lie algebras, we will now describe them explicitly. Throughout the present section, we assume that the ground field is $\mathbb{C}$.

Definition 17. *Let $\mathfrak{g}$ be a semisimple Lie algebra. A Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$ is called **toral** if it is abelian and consists of semisimple elements.*

The standard example of a toral subalgebra is any subspace of the set of diagonal matrices inside $\mathfrak{sl}_n$. Of course, one can replace the set of diagonal matrices by any conjugate thereof, which would lead to many more toral subalgebras. Since it is abelian, any toral subalgebra is isomorphic to $\mathbb{C}^{\oplus n}$ for some n. The space consists of $\mathbb{C}$-linear functions

$$\mathfrak{h}^* = \Big\{\lambda : \mathfrak{h} \to \mathbb{C}\Big\}$$

and it has the same dimension as $\mathfrak{h}$. Throughout the present lecture, $(\cdot,\cdot) : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ denotes an arbitrary symmetric invariant bilinear form, which is non-degenerate (for example the Killing form).

Proposition 16. *For any semisimple Lie algebra $\mathfrak{g}$ over $\mathbb{C}$, and any toral subalgebra $\mathfrak{h} \subset \mathfrak{g}$, we have a decomposition*

$$\mathfrak{g} = \bigoplus_{\lambda \in \mathfrak{h}^*} \mathfrak{g}_\lambda \tag{120}$$

where

$$\boxed{\mathfrak{g}_\lambda = \Big\{x \in \mathfrak{g} \Big| [h,x] = \lambda(h)x, \ \forall h \in \mathfrak{h}\Big\}} \tag{121}$$

Then we have

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta} \tag{122}$$

for all $\alpha, \beta \in \mathfrak{h}^*$, and the non-degenerate form $(\cdot,\cdot)$ has the property that

$$\mathfrak{g}_\alpha \otimes \mathfrak{g}_\beta \xrightarrow{(\cdot,\cdot)} \mathbb{C} \tag{123}$$

is non-degenerate if $\alpha + \beta = 0$ and is 0 if $\alpha + \beta \neq 0$.

Proof. The subspaces (120) are the joint eigenspaces of the commuting operators $\{\mathrm{ad}_x\}_{x\in\mathfrak{h}}$ on $\mathfrak{g}$ (these operators are all semisimple by assumption, hence simultaneously diagonalizable since we are working over $\mathbb{C}$). Property (122) is an immediate consequence of the Jacobi identity

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

(show that the Jacobi identity is equivalent to the above formula) so if $x \in \mathfrak{g}_\alpha \Rightarrow [h,x] = \alpha(h)x$ and $y \in \mathfrak{g}_\beta \Rightarrow [h,y] = \beta(h)y$, then $[h,[x,y]] = (\alpha(h)+\beta(h))[x,y] \Rightarrow [x,y] \in \mathfrak{g}_{\alpha+\beta}$. As far as the restricted pairing (123) is concerned, the invariance implies that

$$\alpha(h)(x,y) = ([h,x],y) = -(x,[h,y]) = -\beta(h)(x,y)$$

for any $x \in \mathfrak{g}_\alpha, y \in \mathfrak{g}_\beta$. This shows that $(x,y) \neq 0$ only if $\alpha+\beta = 0$. The fact that the case $\alpha+\beta = 0$ of the restricted pairing (123) is non-degenerate then follows from the overall non-degeneracy of the pairing $(\cdot,\cdot)$.

□

9.2

In particular, Theorem 13 applied to the non-degenerate pairing (123) for $\alpha = \beta = 0$ implies that $\mathfrak{g}_0$ is reductive. Since $\mathfrak{h}$ is abelian, we have $\mathfrak{g}_0 \supseteq \mathfrak{h}$. The following Definition pertains to the opposite inclusion.

Definition 18. *A toral subalgebra $\mathfrak{h}$ of a complex semisimple Lie algebra $\mathfrak{g}$ is called a* ***Cartan subalgebra*** *if*

$$[x, \mathfrak{h}] \subseteq \mathfrak{h} \quad \Rightarrow \quad x \in \mathfrak{h} \tag{124}$$

Note that (124) implies that $\mathfrak{g}_0 = \mathfrak{h}$. The standard example of a Cartan subalgebra of $\mathfrak{sl}_n$ is the set of diagonal matrices, or any conjugate thereof. From this example, we see that what distinguishes Cartan subalgebras among toral subalgebras is the fact that they are maximal (i.e. the subspace of all diagonal matrices versus some subspace of diagonal matrices). In fact, this is a completely general phenomenon, as we will now see.

Proposition 17. *For a complex semisimple Lie algebra $\mathfrak{g}$, a Cartan subalgebra is the same thing as a maximal toral subalgebra.*

Proof. The fact that a Cartan subagebra is maximal easily follows from (124), since toral subalgebras are by their very nature abelian. For the converse, let us consider a maximal toral subalgebra $\mathfrak{h}$ and prove that

$$\mathfrak{g}_0 = \mathfrak{h} \tag{125}$$

Once we do so, it will follow that $\mathfrak{h}$ is a Cartan subalgebra, as it satisfies the defining property (124) (because of (122), if a certain x has a non-zero component in some $\mathfrak{g}_\alpha$ with $\alpha \neq 0$, then $[x, \mathfrak{g}_0] \not\subseteq \mathfrak{g}_0$). Consider any $x \in \mathfrak{g}_0$ and its Jordan decomposition

$$x = x_{ss} + x_n$$

For any $y \in \mathfrak{h}$, we have $[x, y] = 0$ by definition, hence $[x_{ss}, y] = [x_n, y] = 0$ by (103). Therefore, we have $x_{ss}, x_n \in \mathfrak{g}_0$. If $x_{ss} \notin \mathfrak{h}$, then $\mathfrak{h} \oplus \mathbb{C}x_{ss}$ would be a toral subalgebra, which would contradict the maximality of $\mathfrak{h}$. Therefore, we conclude that $x_{ss} \in \mathfrak{h} \Rightarrow x \in x_n + \mathfrak{h}$, and so

$$\mathrm{ad}_x\Big|_{\mathfrak{g}_0} = \mathrm{ad}_{x_n}\Big|_{\mathfrak{g}_0}$$

is a nilpotent operator on $\mathfrak{g}_0$. By Corollary 4, we conclude that $\mathfrak{g}_0$ is a nilpotent Lie algebra.

Let us first assume that $[\mathfrak{g}_0, \mathfrak{g}_0] \neq 0$. Using Theorem 12, show that nilpotent Lie algebras have the property that their center intersects any non-zero ideal non-trivially. Therefore, there would exist

$$0 \neq z \in \mathfrak{z}(\mathfrak{g}_0) \cap [\mathfrak{g}_0, \mathfrak{g}_0]$$

Since the Killing form is a s.i.b.f., the fact that $z \in [\mathfrak{g}_0, \mathfrak{g}_0]$ implies that

$$(y, z)_{\mathfrak{g}} = 0, \quad \forall y \in \mathfrak{h}$$

On the other hand, since any $x \in \mathfrak{g}_0 \backslash \mathfrak{h}$ is nilpotent and comutes with $z \in \mathfrak{z}(\mathfrak{g}_0)$, then it is an easy fact that

$$(x, z)_{\mathfrak{g}} = \mathrm{tr}_{\mathfrak{g}}(\mathrm{ad}_x \mathrm{ad}_z) = 0 \tag{126}$$

This implies that $(\mathfrak{g}_0, z)_\mathfrak{g} = 0$, which contradicts the non-degeneracy of the restricted pairing (123). We therefore conclude that $[\mathfrak{g}_0, \mathfrak{g}_0] = 0$, i.e. $\mathfrak{g}_0$ is abelian. However, we have already seen that any $x \in \mathfrak{g}_0 \backslash \mathfrak{h}$ would have to be nilpotent, so (126) would hold for all $z \in \mathfrak{g}_0$. By the non-degeneracy of the restricted pairing (123), we conclude that $x = 0$, hence $\mathfrak{g}_0 = \mathfrak{h}$.

□

As a consequence of Proposition 17, Cartan subalgebras of complex semisimple Lie algebras $\mathfrak{g}$ exist: just start from the toral subalgebra 0 and enlarge it as much as possible. All Cartan subalgebras of $\mathfrak{g}$ have the same dimension, which is called the **rank** of $\mathfrak{g}$.

9.3

In light of the discussion in the previous Subsection, it makes sense to consider the decomposition (120) for a maximal toral subalgebra (because then the decomposition would be as fine as possible). Thus, we henceforth let $\mathfrak{h}$ be a Cartan subalgebra, and the decomposition in question takes the form

$$\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha \tag{127}$$

where R simply denotes the set of non-zero linear functionals $\alpha : \mathfrak{h} \to \mathbb{C}$ that have the property that $\mathfrak{g}_\alpha \neq 0$ (and we use the fact that $\mathfrak{g}_0 = \mathfrak{h}$). Since $\mathfrak{g}$ is a finite-dimensional Lie algebra, the set R is finite. It is called the **root system** of $\mathfrak{g}$, and its elements are called **roots**. Because the pairing (123) is non-degenerate when $\alpha + \beta = 0$, we have

$$\alpha \in R \quad \Leftrightarrow \quad -\alpha \in R \tag{128}$$

The $\mathfrak{g}_\alpha$'s that appear in (127) are called the root spaces corresponding the roots α.

Example 6. *Let $\mathfrak{g} = \mathfrak{sl}_n$ and let $\mathfrak{h}$ be the subalgebra of traceless diagonal matrices (with respect to a certain basis). Thus, elements of $\mathfrak{h}$ will be given by*

$$x = (x_1, \dots, x_n) \quad \text{with } x_1 + \dots + x_n = 0$$

We will consider the basis $e_1, \dots, e_n$ of $\mathfrak{h}^*$ given by

$$e_i(x) = x_i$$

With this in mind, the roots are $\{e_i - e_j\}_{1 \leq i \neq j \leq n}$, with

$$(\mathfrak{sl}_n)_{e_i - e_j} = \mathbb{C} E_{ij}$$

Check the previous claim: all it really says is that if x is the diagonal matrix with entries $(x_1, \dots, x_n)$, then we have $[x, E_{ij}] = (x_i - x_j) E_{ij}$. We conclude that the root decomposition is

$$\mathfrak{sl}_n = \mathfrak{h} \bigoplus_{1 \leq i \neq j \leq n} \mathbb{C} E_{ij}$$

A number of properties one can observe from Example 6 are quite general, for example the following.

Proposition 18. *The set of roots spans $\mathfrak{h}^*$.*

Proof. If the set of roots failed to span $\mathfrak{h}^*$, then there would be a non-zero element $x \in \mathfrak{h}$ such that $\alpha(x) = 0$ for all $\alpha \in R$. This implies that $[x, \mathfrak{g}_\alpha] = 0$ for all $\alpha \in R$. However, the fact that x commutes with $\mathfrak{h}$ implies that $x \in \mathfrak{z}(\mathfrak{g})$. Since semisimple algebras have trivial center, then $x = 0$.

□

9.4

Another property that one can observe from Example 6 is that all the root spaces are one-dimensional. To prove that this is in fact a general phenomenon, recall the non-degenerate s.i.b.f. $(\cdot,\cdot) : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$. Its restriction to $\mathfrak{h}$ is also non-degenerate, and so it provides an isomorphism

$$\boxed{\mathfrak{h} \cong \mathfrak{h}^*} \tag{129}$$

We will write $h_\alpha \in \mathfrak{h}$ for the element corresponding to any root α under the above isomorphism. Moreover, we may define the non-degenerate pairing.

$$(\cdot,\cdot) : \mathfrak{h}^* \times \mathfrak{h}^* \to \mathbb{C}$$

via the isomorphism (129).

Lemma 7. *For any root α and any $e_\alpha \in \mathfrak{g}_\alpha$, $f_\alpha \in \mathfrak{g}_{-\alpha}$, we have*

$$[e_\alpha, f_\alpha] = (e_\alpha, f_\alpha) h_\alpha \tag{130}$$

Proof. For any $h \in \mathfrak{h}$, we have

$$([e_\alpha, f_\alpha], h) = ([h, e_\alpha], f_\alpha) = \alpha(h)(e_\alpha, f_\alpha)$$

By the non-degeneracy of $(\cdot,\cdot)$ restricted to $\mathfrak{h}$, this implies (130).

□

Proposition 19. *For any root α, we have $(\alpha, \alpha) \neq 0$, so we may define*

$$H_\alpha = \frac{2h_\alpha}{(\alpha, \alpha)} \tag{131}$$

If $E_\alpha \in \mathfrak{g}_\alpha$ and $F_\alpha \in \mathfrak{g}_{-\alpha}$ are chosen so that

$$(E_\alpha, F_\alpha) = \frac{2}{(\alpha, \alpha)}$$

then we have the commutation relations

$$[H_\alpha, E_\alpha] = 2E_\alpha, \qquad [H_\alpha, F_\alpha] = -2F_\alpha, \qquad [E_\alpha, F_\alpha] = H_\alpha \tag{132}$$

In other words, $E_\alpha, F_\alpha, H_\alpha$ provide a Lie algebra homomorphism

$$\mathfrak{sl}_2 \hookrightarrow \mathfrak{g}$$

*so they are called a $\mathfrak{sl}_2$-**triple**.*

The reason for the normalization (131) is that such a H_α is independent of the choice of s.i.b.f. We will soon see that the $\mathfrak{sl}_2$-triple corresponding to a root α is unique, up to rescaling E_α, F_α by opposite constants.

Proof. of Proposition 19: Let us assume for the purpose of contradiction that $(\alpha, \alpha) = 0$, which would mean that $\alpha(h_\alpha) = 0$. We may choose $e_\alpha \in \mathfrak{g}_\alpha$ and $f_\alpha \in \mathfrak{g}_{-\alpha}$ such that $(e_\alpha, f_\alpha) = 1$ (by the non-degeneracy of the s.i.b.f.) and so Lemma 7 would imply that the subalgebra $\mathfrak{a} \subset \mathfrak{g}$ generated by $e_\alpha, f_\alpha, h_\alpha$ satisfies the Lie bracket relations

$$[h_\alpha, e_\alpha] = [h_\alpha, f_\alpha] = 0, \quad [e_\alpha, f_\alpha] = h_\alpha$$

This subalgebra is solvable (due to the abelian ideal $\mathbb{C}h_\alpha$) and so Lie's Theorem 11 implies that there is a basis of $\mathfrak{g}$ in which ad_{e_α} and ad_{f_α} are upper triangular. Being the commutator of upper triangular matrices, ad_{h_α} would be strictly upper triangular. However, because $\mathfrak{h}$ is a toral subalgebra, ad_{h_α} is also semisimple. Therefore, we conclude that $\text{ad}_{h_\alpha} = 0$, which implies that $h_\alpha \in \mathfrak{z}(\mathfrak{g})$, which is impossible because semisimple Lie algebras have no center. Having showed that $(\alpha, \alpha) \neq 0$, properties (132) are straightforward computations, which we leave to you.

□

9.5

$\mathfrak{sl}_2$-triples give a powerful tool for the study of semisimple Lie algebras.

Proposition 20. *For any root α, the subspaces $\mathfrak{g}_{\pm\alpha}$ are one-dimensional.*

Proof. Consider any root α, fix a corresponding $\mathfrak{sl}_2$-triple $E_\alpha, F_\alpha, H_\alpha$ and let

$$V_\alpha = \mathbb{C}H_\alpha \bigoplus_{\ell \in \mathbb{Z} \setminus 0} \mathfrak{g}_{\ell\alpha} \tag{133}$$

By (122), the operators ad_{E_α}, ad_{F_α}, ad_{H_α} provide a representation $\mathfrak{sl}_2 \curvearrowright V_\alpha$. The weights of this representation, i.e. the eigenvalues of ad_{H_α}, are equal to the numbers 2ℓ in (133). Thus, V_α is a representation of $\mathfrak{sl}_2$ with all even weights and a one-dimensional 0 weight subspace, so (58) implies that V_α must be irreducible. In particular, this implies that all the root subspaces $\mathfrak{g}_{\pm\alpha}$ are one-dimensional. □

Proposition 21. *For any two roots α and β, the number*

$$c_{\alpha\beta} = \frac{2(\alpha, \beta)}{(\alpha, \alpha)}$$

is an integer, and $\beta - c_{\alpha\beta}\alpha$ is also a root.

Proof. The numbers in question are equal to the weights of the adjoint action of an $\mathfrak{sl}_2$-triple $E_\alpha, F_\alpha, H_\alpha$ on $\mathfrak{g}_\beta$. Since any finite-dimensional representation of $\mathfrak{sl}_2$ has integer weights, we conclude that $c_{\alpha\beta} \in \mathbb{Z}$. As we have seen in Lecture 5, in any finite-dimensional representation of $\mathfrak{sl}_2$, the operators E^n and F^n provide isomorphisms between the subspaces of weight n and $-n$. In the case at hand, if $0 \neq x \in \mathfrak{g}_\beta$ and $c_{\alpha\beta} < 0$ (respectively $c_{\alpha\beta} > 0$), then $\text{ad}_{E_\alpha}^{-c_{\alpha\beta}}(x) \neq 0$ (respectively $\text{ad}_{F_\alpha}^{c_{\alpha\beta}}(x) \neq 0$). Since the latter elements lie in $\mathfrak{g}_{\beta - c_{\alpha\beta}\alpha}$, we conclude that $\beta - c_{\alpha\beta}\alpha$ is also a root. □

Proposition 22. *The only multiples of a root α which are also roots are α and $-\alpha$.*

Proof. The fact that we have a non-degenerate pairing between $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ implies that $-\alpha$ is a root whenever α is a root. On the other hand, if αt were also a root for some $t \in \mathbb{C}\backslash\{\pm 1\}$, then Proposition 21 would imply that $2t \in \mathbb{Z}$. However, switching the roles of α and αt would also imply that $2t^{-1} \in \mathbb{Z}$, which only leaves the possibility that $t \in \{\pm\frac{1}{2}, \pm 2\}$. Let us assume without loss of generality that $t = \pm 2$. Then as we saw in Proposition 20, V_α must be an irreducible representation with respect to the $\mathfrak{sl}_2$ triple $E_\alpha, F_\alpha, H_\alpha$. This would imply that $\mathfrak{g}_{2\alpha} \subseteq \mathrm{ad}_{E_\alpha}(\mathfrak{g}_\alpha)$, which is impossible since we already showed that $\mathfrak{g}_\alpha$ is the one-dimensional space $\mathbb{C}E_\alpha$. $\square$

Lecture 10

10.1

Propositions 18, 20, 21, 22 can be unified in the following abstract definition of a root system (which we will shortly see is a model for all complex semisimple Lie algebras).

Definition 19. *An **abstract root system** is a $\mathbb{R}$-vector space U endowed with an inner product*

$$U \times U \xrightarrow{(\cdot,\cdot)} \mathbb{R} \tag{134}$$

together with a finite set $R \subset U \backslash 0$ such that

$$R \text{ spans } U \tag{135}$$

$$\text{if } \alpha \in R \text{ then } k\alpha \begin{cases} \in R & \text{if } k = -1 \\ \notin R & \text{otherwise} \end{cases} \tag{136}$$

$$\text{if } \alpha, \beta \in R \text{ then } c_{\alpha\beta} := \frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{Z} \tag{137}$$

$$\text{if } \alpha, \beta \in R \text{ then } s_\alpha(\beta) = \beta - c_{\alpha\beta}\alpha \in R \tag{138}$$

The **rank** of a root system is defined to be the dimension of U.

10.2

Axioms (137) and (138) might seem contrived at first, but they have a geometric meaning in terms of reflections. The former of these axioms is a statement about the angle between the vectors α and β. Meanwhile, the latter axiom concerns the function

$$s_\alpha : U \to U, \qquad s_\alpha(\lambda) = \lambda - \frac{2(\alpha,\lambda)}{(\alpha,\alpha)}\alpha \tag{139}$$

which is none other than the reflection across the hyperplane $\alpha^\perp$ perpendicular to α: thus (138) merely says that the root system is preserved by such reflections.

Definition 20. *The **abstract Weyl group** is the subgroup of $GL(U)$ (actually of the orthogonal group, because reflections preserve the inner product* (134)*) generated by the reflections $\{s_\alpha\}_{\alpha \in R}$.*

Note that the Weyl group is always finite, because any element of $GL(U)$ which fixes every root must be the identity due to (135). The reason for the word "abstract" in Definition 20 is to differentiate it from the Weyl group of a complex semisimple Lie algebra $\mathfrak{g}$, which is defined as

$$W = N_G(H)/H \tag{140}$$

where G is the complex Lie group with Lie algebra $\mathfrak{g}$, and H is a closed Lie subgroup with Lie algebra given by a Cartan subalgebra of G (such a H is called a maximal torus). As you can expect, W in (140) is isomorphic to the abstract Weyl group corresponding to the root system of $\mathfrak{g}$, though we will not prove it.

Example 7. *For the root system of $\mathfrak{sl}_n$, we have that*

$$s_{e_i - e_j}(\ldots, x_i, \ldots, x_k, \ldots, x_j, \ldots) = (\ldots, x_j, \ldots, x_k, \ldots, x_i, \ldots)$$

The corresponding Weyl group is easily seen to be the symmetric group S_n.

10.3

It turns out that the axioms of a root system are very restrictive: in rank 1, this is probably not so surprising, since the only root system in $\mathbb{R}$ is the root system associated to the Lie algebra $\mathfrak{sl}_2$: this root system is called "type A_1", where the subscript denotes the rank.

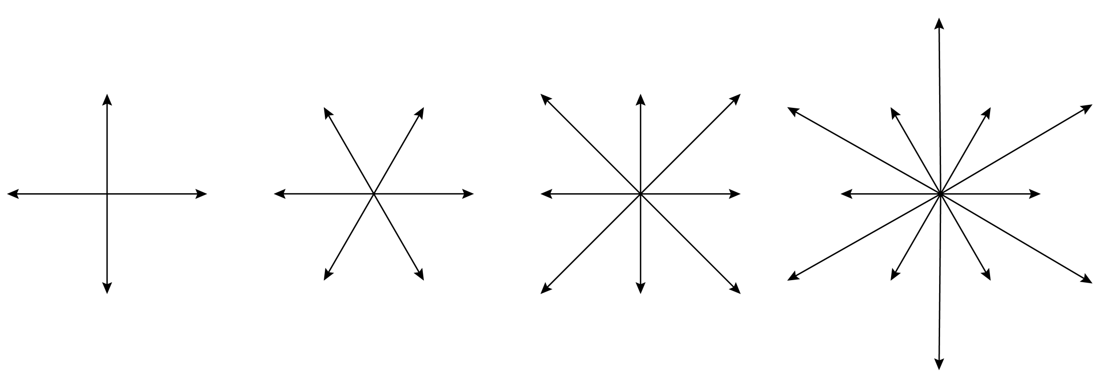
Things are a bit more interesting in rank 2, i.e. $U = \mathbb{R}^2$. In this case, the angle θ between two non-collinear roots $\alpha, \beta \in R$ is given by the following formula (we let $|\alpha| = \sqrt{(\alpha, \alpha)}$)

$$\cos\theta = \frac{(\alpha, \beta)}{|\alpha||\beta|} = \frac{c_{\alpha\beta}}{2} \cdot \frac{|\alpha|}{|\beta|} = \frac{c_{\beta\alpha}}{2} \cdot \frac{|\beta|}{|\alpha|} \quad \Rightarrow \quad (\cos\theta)^2 = \frac{c_{\alpha\beta}c_{\beta\alpha}}{4}$$

Therefore, we must have $0 \leq c_{\alpha\beta}c_{\beta\alpha} < 4$. As axiom (137) requires that $c_{\alpha\beta}$ and $c_{\beta\alpha}$ be integers, we only have the following options (we assume without loss of generality that $|\alpha| \leq |\beta|$, so $|c_{\alpha\beta}| \geq |c_{\beta\alpha}|$):

- $c_{\alpha\beta} = 0 \Leftrightarrow c_{\beta\alpha} = 0$, which implies $\theta = \frac{\pi}{2}$
- $c_{\alpha\beta} = c_{\beta\alpha} = 1$, which implies $|\alpha| = |\beta|$ and $\theta = \frac{\pi}{3}$
- $c_{\alpha\beta} = c_{\beta\alpha} = -1$, which implies $|\alpha| = |\beta|$ and $\theta = \frac{2\pi}{3}$
- $c_{\alpha\beta} = 2$ and $c_{\beta\alpha} = 1$, which implies $|\alpha|\sqrt{2} = |\beta|$ and $\theta = \frac{\pi}{4}$
- $c_{\alpha\beta} = -2$ and $c_{\beta\alpha} = -1$, which implies $|\alpha|\sqrt{2} = |\beta|$ and $\theta = \frac{3\pi}{4}$
- $c_{\alpha\beta} = 3$ and $c_{\beta\alpha} = 1$, which implies $|\alpha|\sqrt{3} = |\beta|$ and $\theta = \frac{\pi}{6}$
- $c_{\alpha\beta} = -3$ and $c_{\beta\alpha} = -1$, which implies $|\alpha|\sqrt{3} = |\beta|$ and $\theta = \frac{5\pi}{6}$

With this in mind, the following are easily seen to be root systems, because the angle between and length of any two roots are admissible by the discussion above (the notation $A_1 \times A_1$, A_2, $B_2 \cong C_2$ and G_2 will be explained in Theorem 19).



$A_1 \times A_1$ $\quad$ A_2 $\quad$ $B_2 = C_2$ $\quad$ G_2

What is more interesting is that the above are all the rank 2 root systems, up to linear transformations. Indeed, consider any rank 2 root system R and look at the most obtuse angle between two non-collinear roots: if this angle is $\frac{\pi}{2}$, $\frac{2\pi}{3}$, $\frac{3\pi}{4}$, $\frac{5\pi}{6}$, the root system is $A_1 \times A_1$, A_2, $B_2 = C_2$, G_2, respectively (prove this yourselves; the idea is that once you draw two vectors with the most obtuse angle between them, the fact that R is preserved under the reflections (139) means that R must contain a copy R' of the root system of type $A_1 \times A_1$, A_2, $B_2 = C_2$, G_2, respectively; but if R contained any other root α, you could find some root $\beta \in R'$ which would violate the angle and length conditions in the bullets above).

10.4

What about root systems of arbitrary rank? The bulleted list in the previous Subsection actually pertains to any pair of non-collinear elements of any root system, and thus controls the angle between and lengths of any pair of roots. For instance, we ask you to prove the following Lemma by examining all the rank 2 systems above.

Lemma 8. *If $\alpha \neq \beta \in R$ have the property that $(\alpha, \beta) > 0$, then $\alpha - \beta \in R$.*

Let us now consider arbitrary root systems R, and develop some further tools to classify them.

Definition 21. *Any hyperplane in U that does not intersect R determines a decomposition*

$$R = R^+ \sqcup R^- \tag{141}$$

*into **positive** and **negative** roots, depending on which side of U they lie. Clearly, $R^- = -R^+$. A **simple** root is a positive root which cannot be written as a sum of two or more positive roots.*

For the root system associated to $\mathfrak{sl}_n$, the usual choice is to let

$$R^+ = \left\{ e_i - e_j \middle| 1 \leq i < j \leq n \right\}$$
$$R^- = \left\{ e_i - e_j \middle| 1 \leq j < i \leq n \right\}$$

The simple roots are then $\alpha_i = e_i - e_{i+1}$, with $i \in \{1, \dots, n-1\}$.

Proposition 23. *(a) Every positive root can be written uniquely as a sum of simple roots.*

(b) The simple roots determine a basis of U (so there are as many of them as the rank of R).

Proof. (a) We may successively decompose any positive root α into sums of positive roots. This process must terminate after finitely many steps (since there are finitely many positive roots, and all of them are at least a fixed distance away from the hyperplane separating R^+ from R^-) and when it terminates, we will have written α as a sum of simple roots.

(b) By the previous part and axiom (135), the simple roots span U. To prove that they are linearly independent, note that

$$\alpha \neq \beta \text{ simple} \quad \Rightarrow \quad (\alpha, \beta) \leq 0 \tag{142}$$

(indeed, otherwise Lemma 8 would imply that either $\alpha - \beta$ or $\beta - \alpha$ is a positive root, which would contradict the simplicity of α and β). However, it is a classic and easy exercise to show that any set of vectors which have all non-acute angles between them must be linearly independent. $\square$

10.5

The following result, which will occupy the remainder of this lecture, shows that a set of simple roots determines the entire R. In what follows, we fix a set of simple roots $\alpha_1, \ldots, \alpha_r$ of R.

Theorem 18. *The Weyl group W is generated by the* ***simple reflections***

$$\boxed{\left\{s_i = s_{\alpha_i}\right\}_{i \in \{1,\ldots,r\}}} \tag{143}$$

and any root can be obtained by some element of W acting on some simple root α_i.

The theorem above says that a Weyl group is a particular case of a so-called Coxeter group. The first step in proving Theorem 18 is to systematize the freedom we had in choosing the decomposition (141).

Definition 22. *Consider the hyperplanes $\alpha^\perp$ perpendicular to the the roots $\alpha \in R$. A connected component $\mathcal{C}$ of*

$$U \Big\backslash \bigcup_{\alpha \in R} \alpha^\perp$$

is called a ***Weyl chamber****. The boundary hyperplanes of a chamber are often called its walls.*

By definition, a Weyl chamber is the set of all $x \in U$ such that (x, α) has a given sign for all roots $\alpha \in R$. As the decomposition (141) partitions the roots into two sets, depending on whether their pairing with a given $x \in U$ is positive or negative, we conclude that the decomposition itself only depends on the Weyl chamber of x. In other words, a choice of positive/negative roots is equivalent to a choice of positive/negative Weyl chamber $\mathcal{C}^\pm$ (namely the chamber consisting of x's whose inner product with the positive/negative roots is > 0).

Recall that the reflections s_α of (139) are by definition given by reflecting in the hyperplanes $\alpha^\perp$. By the axiom that the Weyl group action takes any root β to a root β' it follows that W takes the hyperplane $\beta^\perp$ to ${\beta'}^\perp$ and thus W takes Weyl chambers to Weyl chambers.

Proposition 24. *The Weyl group action on the Weyl chambers is transitive.*

Proof. It is a common feature of hyperplane arrangements that any two chambers $\mathcal{C}$ and $\mathcal{C}'$ can be connected by a sequence of Weyl chambers

$$\mathcal{C} = \mathcal{C}_0, \mathcal{C}_1, \ldots, \mathcal{C}_{k-1}, \mathcal{C}_k = \mathcal{C}'$$

such that $\mathcal{C}_i$ and $\mathcal{C}_{i+1}$ are adjacent chambers separated by a single hyperplane $\alpha^\perp$. That means that the reflection s_α takes $\mathcal{C}_i$ to $\mathcal{C}_{i+1}$, which implies that some element of the Weyl group takes $\mathcal{C}$ to $\mathcal{C}'$. $\square$

Corollary 7. *For any two sets of positive roots $R = R^+ \sqcup R^- = {R'}^+ \sqcup {R'}^-$, there exists an element of the Weyl group taking $R^\pm$ to ${R'}^\pm$.*

Proof. Let $\mathcal{C}$ and $\mathcal{C}'$ be the positive Weyl chambers with respect to R^+ and R'^+, respectively. Proposition 24 says that there exists an element w of the Weyl group which sends $\mathcal{C}$ to $\mathcal{C}'$. Since elements of the Weyl group act by orthogonal matrices, they preserve the scalar product, so a positive root with respect to $\mathcal{C}$ will be sent to a positive root with respect to $\mathcal{C}'$ (check this fact). Since the number of positive roots is always the same, this proves that $w(R^+) = R'^+$. □

It is also easy to see that a Weyl group element taking R^+ to R'^+ must also take the simple roots inside R^+ to the simple roots inside R'^+.

10.6

Let us now fix a decomposition (141). By definition, the positive Weyl chamber $\mathcal{C}^+$ is defined by the property $(x, \alpha) > 0$ for all $\alpha \in R^+$. By Proposition 23, this is equivalent to

$$(x, \alpha_i) > 0$$

where $\alpha_1, \dots, \alpha_r$ are the simple roots inside R^+. This implies that the walls of the chamber $\mathcal{C}^+$ are actually $\alpha_1^\perp, \dots, \alpha_r^\perp$, since it is impossible to encounter any other wall $\alpha^\perp$ (i.e. have $(x, \alpha) = 0$) without first encountering one of the walls $\alpha_1^\perp, \dots, \alpha_r^\perp$ (i.e. have $(x, \alpha_i) = 0$ for some $i \in \{1, \dots, r\}$).

Proposition 25. *For any chamber $\mathcal{C}$, there exist simple reflections $i_1, \dots, i_k$ such that*

$$\mathcal{C} = s_{i_1} \dots s_{i_k}(\mathcal{C}^+) \tag{144}$$

Proof. By Proposition 24, there exists a sequence of Weyl chambers

$$\mathcal{C}^+ = \mathcal{C}_0, \mathcal{C}_1, \dots, \mathcal{C}_{k-1}, \mathcal{C}_k = \mathcal{C}$$

such that $\mathcal{C}_{\ell-1}$ is separated from $\mathcal{C}_\ell$ by a single hyperplane. We will prove that each $\mathcal{C}_\ell$ can be written in the form (144) by induction on ℓ. So assume that we have

$$\mathcal{C}_{\ell-1} = s_{i_1} \dots s_{i_{\ell-1}}(\mathcal{C}^+) \tag{145}$$

and $\mathcal{C}_\ell$ is separated from $\mathcal{C}_{\ell-1}$ by the hyperplane $\alpha^\perp$. This hyperplane must be of the form $s_{i_1} \dots s_{i_{\ell-1}}(\alpha_{i_\ell}^\perp)$ for some $i_\ell \in I$, because the walls of $\mathcal{C}^+$ are the $\alpha_i^\perp$'s. We conclude that

$$\alpha^\perp = s_{i_1} \dots s_{i_{\ell-1}}(\alpha_{i_\ell}^\perp) \quad \Leftrightarrow \quad \alpha = s_{i_1} \dots s_{i_{\ell-1}}(\alpha_{i_\ell})$$

The obvious formula

$$\boxed{s_{w(\alpha)} = w s_\alpha w^{-1}} \tag{146}$$

$\forall w \in W, \alpha \in R$, then allows us to deduce the required identity $\mathcal{C}_\ell = s_{i_1} \dots s_{i_\ell}(\mathcal{C}^+)$ from (145). □

Proof. of Theorem 18: Since a choice of Weyl chamber is equivalent to a choice of positive roots, Proposition 25 means that for any two collections of positive roots, we can find a product of simple reflections that takes one to the other. By (146) then, any Weyl group element can be written as a product of simple reflections. To see that any root can be obtained by some element of W acting on some simple root, it suffices to show that any root can be made to be a simple root with respect to some hyperplane (just choose the hyperplane very close to the simple root in question).

□

Lecture 11

11.1

In the previous lecture, we showed that one can reconstruct a root system from a set of simple roots $\{\alpha_i\}_{i\in I}$. In turn, we will soon see that such a set of simple roots is completely determined (up to an angle-preserving linear transformation of U) by the following notion that encodes the lengths of simple roots and the angles between them.

Definition 23. *A* ***Cartan matrix*** *of rank r is a square matrix with integer entries*

$$C = (c_{ij})_{1\le i,j\le r} \tag{147}$$

such that

- $c_{ii} = 2$ *and* $c_{ij} \le 0$ *for all* $i \neq j$.
- $c_{ij} = 0 \Leftrightarrow c_{ji} = 0$.
- $C = DS$, *where D is a diagonal matrix with positive entries on the diagonal and S is a positive-definite symmetric matrix.*

We can associate a Cartan matrix to any root system R, by letting $c_{ij} = c_{\alpha_i\alpha_j}$ for some set of simple roots $\alpha_1, \dots, \alpha_r$. One chooses the matrices D and S to have entries $\frac{2}{(\alpha_i,\alpha_i)}$ and (α_i, α_j), respectively, with the notation in Definition 19. Then the positive-definiteness of S is equivalent to the fact that $\alpha_1, \dots, \alpha_r$ are a basis of E. For instance, the Cartan matrix of $\mathfrak{sl}_n$ is

$$\begin{pmatrix} 2 & -1 & 0 & \vdots & 0 & 0 & 0 \\ -1 & 2 & -1 & \vdots & 0 & 0 & 0 \\ 0 & -1 & 2 & \vdots & 0 & 0 & 0 \\ \dots & \dots & \dots & \ddots & \dots & \dots & \dots \\ 0 & 0 & 0 & \vdots & 2 & -1 & 0 \\ 0 & 0 & 0 & \vdots & -1 & 2 & -1 \\ 0 & 0 & 0 & \vdots & 0 & -1 & 2 \end{pmatrix} \tag{148}$$

Note that positive-definiteness implies that in a Cartan matrix we have the equation

$$0 \le c_{ij}c_{ji} < 4$$

just like we saw in Subsection 10.3. We typically refer to <u>the</u> Cartan matrix of a root system because changing the choice of simple roots merely has the effect of permuting the rows and columns of C.

Proposition 26. *Up to an angle-preserving linear transformation of U, a root system R is completely determined by its Cartan matrix.*

Proof. The key insight is that given a Cartan matrix C, the decomposition

$$C = DS$$

is unique up to rescaling D and S by inverse amounts (this is because multiplying D on the right by a non-scalar diagonal matrix D' would have to be counterbalanced by multiplying S on the left by D'^{-1}, but this would spoil the symmetry of S). Therefore, once the Cartan matrix of a root system is given, the inner products (α_i, α_j) are all determined up to constant multiple. It is easy to show that this determines the simple roots $\alpha_1, \ldots, \alpha_r$ up to an angle-preserving linear transformation. But once a collection of simple roots is fixed, Theorem 18 implies that any root can be obtained from them by successively applying reflections $s_i = s_{\alpha_i}$. □

11.2

Beside Cartan matrices, which provide a numerical characterization of simple roots in a root system, we also have the following graphical realization of the same information.

Definition 24. *The* ***Dynkin diagram*** *associated to a root system R is the graph with vertex set $\{1, \ldots, r\}$, and*

- *0 edges between i and j if the angle between α_i and α_j is $\frac{\pi}{2}$*
- *1 edge between i and j if the angle between α_i and α_j is $\frac{2\pi}{3}$*
- *2 edges between i and j if the angle between α_i and α_j is $\frac{3\pi}{4}$*
- *3 edges between i and j if the angle between α_i and α_j is $\frac{5\pi}{6}$*

(by the discussion in Subsections 10.3 and 10.4, the above are the only possibilities for angles between simple roots). If there are multiple edges between two vertices, we draw an arrow from the one corresponding to a longer root to the one corresponding to a shorter root.

It is easy to see that there is a one-to-one correspondence

$$\Big(\text{Cartan matrices}\Big) \leftrightarrow \Big(\text{Dynkin diagrams}\Big) \tag{149}$$

wherein the set of rows/columns of a Cartan matrix is identified with the vertex set $\{1, \ldots, r\}$ of a Dynkin diagram. For any two vertices $i \neq j \in \{1, \ldots, r\}$, the number of edges between i and j in the Dynkin diagram are perfectly encoded in the non-positive integer entries c_{ij} and c_{ji} of the Cartan matrix, as explained in the bulleted list of Subsection 10.3. More explicitly, since filling out a Cartan matrix and drawing a Dynkin diagram are rank 2 tasks (i.e. ones which you perform by considering any principal 2×2 submatrix and any 2-vertex subgraph at a time) then the correspondence (149) is completely determined by the assignment

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \quad \leftrightarrow \quad A_1 \times A_1$$

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad \leftrightarrow \quad A_2$$

$$\begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix} \quad \leftrightarrow \quad B_2 = C_2$$

$$\begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \quad \leftrightarrow \quad G_2$$

(see below for the notation $A_n, B_n, C_n, D_n, E_{6,7,8}, F_4, G_2$ of Dynkin diagrams).

Definition 25. *Given root systems $R \subset U$ and $R' \subset U'$, their* ***direct sum*** *is the root system*

$$(R, 0) \sqcup (0, R') \subset E \oplus E'$$

A root system which is not isomorphic to a direct sum of root systems is called ***irreducible****.*

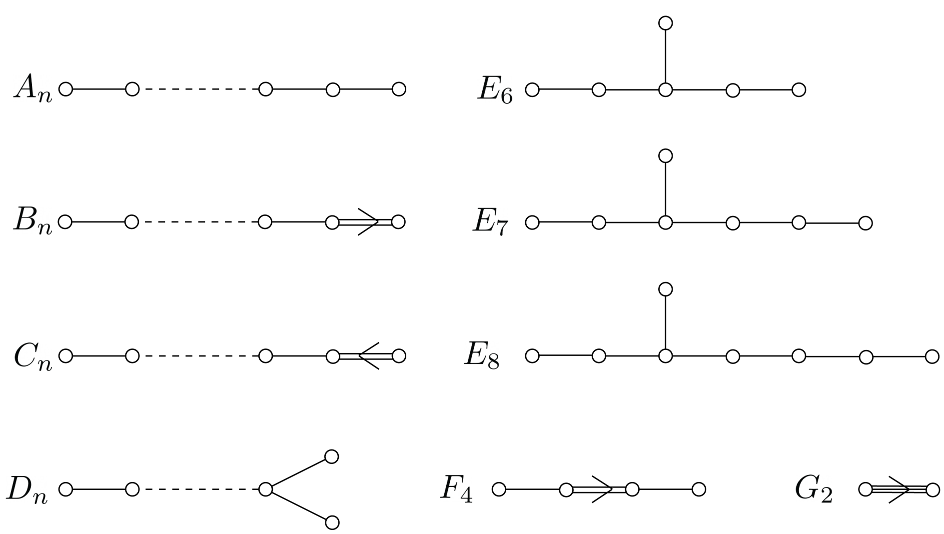
Proposition 27. *In an irreducible root system R, either all roots have the same length (in which case R is called* ***simply laced****) or there are only two possible values for the root lengths (in which case they are called* ***short*** *and* ***long*** *roots, the latter being $\sqrt{2}$ or $\sqrt{3}$ times longer than the fomer).*

Proof. In an irreducible root system, the action $W \curvearrowright U$ is irreducible (hence the terminology), because the orthogonal complement of any W-invariant subspace with respect to the inner product (134) would also be W-invariant. As a consequence, the W-orbit of any root α spans U, so for any other root β there must exist $w \in W$ such that $(w(\alpha), \beta) \neq 0$. The bulleted list in Subsection 10.3 then implies that the length of β and the length of α (which is equal to the length of $w(\alpha)$ for any $w \in W$) must differ by a ratio of $1, \sqrt{2}, \sqrt{3}$. Thus, if the roots could have 3 or more lengths, we could always find a pair of them which differ by a ratio other than $1, \sqrt{2}, \sqrt{3}$, thus contradicting the previous sentence. □

11.3

It is easy to see that a Cartan matrix (respectively Dynkin diagram) corresponds to an irreducible root system if and only if it is not block diagonal (respectively connected). Therefore, the task is to classify irreducible Dynkin diagrams.

Theorem 19. *Any irreducible Dynkin diagram is one of the following list*

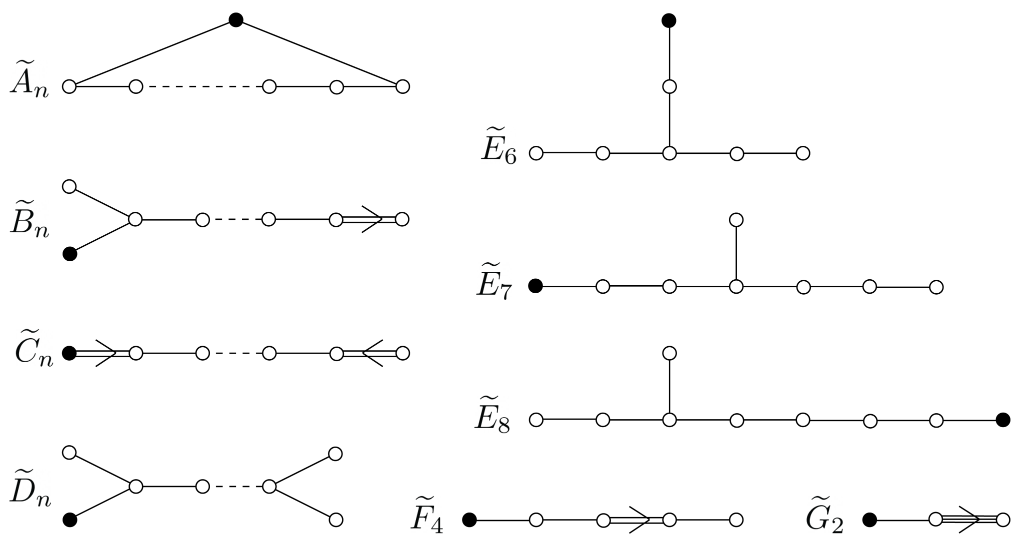


where the index denotes the number of vertices in the diagram.

Proof. The fact that the pictures above represent Dynkin diagrams comes from the fact that the corresponding Cartan matrices $C = DS$ have positive determinant

$$\det A_n = n+1$$
$$\det B_n = 2$$
$$\det C_n = 2$$
$$\det D_n = 4$$
$$\det E_{6,7,8} = 3, 2, 1$$
$$\det F_4 = 1$$
$$\det G_2 = 1$$

and that the top left corners of the corresponding S matrices also have positive determinant (since they are also Cartan matrices of the Dynkin diagrams listed above). By Sylvester's criterion, these matrices S are positive definite. On the other hand, the following pictures are <u>not</u> Dynkin diagrams, as the corresponding matrices S have determinant 0 [1].



These are called **extended Dynkin diagrams**, and they reflect the representation theory of affine Lie algebras (these are infinite-dimensional Lie algebras obtained from $\mathfrak{g}[t^{\pm 1}]$ where $\mathfrak{g}$ is a complex semisimple Lie algebra, that we will not study in the present course). But the relevance of the pictures above to us is that no Dynkin diagram can contain an extended Dynkin diagram as a subdiagram, or else its symmetrized Cartan matrix S would have a principal minor of determinant 0 (and thus fail to be positive definite). Thus, we have the following observations:

- A Dynkin diagram cannot contain a cycle: indeed, if we had a cycle $i_1, i_2, \dots, i_k, i_{k+1} = i_1$ with $k \geq 3$, then we would be able to contradict the positive-definiteness of $S = (d_{ij})_{1 \leq i,j \leq r}$ as follows

$$\sum_{s=1}^{k} (d_{i_s i_s} + 2 d_{i_s i_{s+1}}) \leq 0$$

[1] The convention for $\widetilde{A}_1$ is that its Cartan matrix is $\begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$.

whch holds because $C = DS$ implies that $c_{ij} = \frac{2d_{ij}}{d_{ii}}$ for all $i \neq j \in \{1, \dots, r\}$, and we have

$$d_{ii} + d_{jj} \leq -4d_{ij} \Leftrightarrow \frac{1}{-c_{ij}} + \frac{1}{-c_{ji}} \leq 2$$

for all negative integers c_{ij}, c_{ji}.

- If a Dynkin diagram contains a triple edge, then it is just G_2 (otherwise it would contain a copy of the extended Dynkin diagram $\widetilde{G}_2$, or one whose matrix S is negative-definite)

- if a Dynkin diagram contains a double edge, then it is just B_n, C_n or F_4, because otherwise it would contain a copy of $\widetilde{A}_1, \widetilde{B}_n$, $\widetilde{C}_n$ or $\widetilde{F}_4$.

- if a Dynkin diagram has only single edges and no cycles, then it is either one of $A_n, D_n, E_{6,7,8}$, or else it would contain a copy of $\widetilde{D}_n$ or $\widetilde{E}_{6,7,8}$.

□

11.4

We have already seen that A_n is the Dynkin diagram of the Lie algebra $\mathfrak{sl}_{n+1}$. We also have

B_n is the Dynkin diagram of $\mathfrak{o}_{2n+1}$
C_n is the Dynkin diagram of $\mathfrak{sp}_{2n}$
D_n is the Dynkin diagram of $\mathfrak{o}_{2n}$ for $n > 1$

As for the Lie algebras that correspond to types E, F, G, we will construct them abstractly in the next lecture. As the Dynkin diagrams above are all irreducible, the corresponding Lie algebras are simple. For a semisimple Lie algebra $\mathfrak{g}$, its Dynkin diagram is simply the disconnected union of the Dynkin diagrams of the direct summands of $\mathfrak{g}$.

Lecture 12

12.1

In the previous lectures, we showed how to perform the following operations

$$\text{semisimple Lie algebras} \rightsquigarrow \text{root systems} \rightsquigarrow \text{Cartan matrices} \leftrightarrow \text{Dynkin diagrams}$$

We will now show how to reconstruct a semisimple Lie algebra from the Cartan matrix / Dynkin diagram of the corresponding root system. Before we do so, we must be able to define Lie algebras by generators and relations. The following discussion is completely analogous to that of groups, that you encountered in Math 211. Let $\mathbb{K}$ be any field of characteristic 0.

Definition 26. *Let S be any set called an **alphabet**. The free associative algebra on S is*

$$A_S = \bigoplus_{s_1 \dots s_k \text{ word in } S} \mathbb{K} s_1 \dots s_k$$

*with the operation given by concatenation of words. We can think of A_S as a Lie algebra with respect to commutator, and we define the **free Lie algebra** on S*

$$\boxed{\mathfrak{f}_S} \tag{150}$$

as the Lie subalgebra of A_S generated by all one-letter words.

If you like universal properties, the free Lie algebra is determined up to isomorphism by the fact that for any Lie algebra $\mathfrak{g}$, a choice of elements $\{x_s \in \mathfrak{g}\}_{s \in S}$ extends uniquely to a Lie algebra homomorphism $\mathfrak{f}_S \to \mathfrak{g}$. But perhaps more explicitly, you should think of $\mathfrak{f}_S$ as consisting of all $\mathbb{K}$-linear combination of symbols

$$[\dots [[s_1, s_2], s_3], [s_4, s_5] \dots] \tag{151}$$

(for any $s_1, s_2, \dots \in S$ and any distribution of square brackets) that satisfy antisymmetry and the Jacobi identity. While this may seem complicated, it is controlled by beautiful combinatorics. For instance, let us fix a total order on the set S, which determines a lexicographic order on the set of all words written with the alphabet S. We call a word w **Lyndon** (also known as **Shirshov**) if it is lexicographically smaller than all of its proper suffixes. Then a classic result is that

$$\mathfrak{f}_S = \bigoplus_{w \text{ Lyndon}} \mathbb{K} x_w$$

where $x_w \in \mathfrak{f}_S$ are defined recursively by $x_s = s$ for any $s \in S$, while for any Lyndon word w of length ≥ 2 we set

$$x_w = [x_{w'}, x_{w''}]$$

where w'' is the longest suffix of $w = w'w''$ which is also a Lyndon word (with this choice, it is not hard to show that the prefix w' is also a Lyndon word).

Definition 27. *Let R denote any set of **relations**, i.e. $\mathbb{K}$-linear combinations of symbols* (151)*. Then*

$$\boxed{\mathfrak{f}_{S|R} = \mathfrak{f}_S \Big/ \Big(\text{ideal generated by } R\Big)} \tag{152}$$

is called the Lie algebra generated by S modulo relations R.

For example, if R is the set of $\{ss' - s's\}_{s,s' \in S}$, then (152) is called the free abelian Lie algebra on S, and it is simply isomorphic to $\oplus_{s \in S} \mathbb{K} s$ with trivial Lie bracket.

12.2

We henceforth work over the ground field $\mathbb{C}$. For any Cartan matrix $C = (c_{ij})_{i,j \in \{1,\dots,r\}}$, we define

$$\boxed{\mathfrak{g}_C} \tag{153}$$

to be the Lie algebra generated by symbols $\{E_i, F_i, H_i\}_{i \in \{1,\dots,r\}}$ modulo the relations

$$[H_i, H_j] = 0 \tag{154}$$

$$[H_i, E_j] = c_{ij} E_j \tag{155}$$

$$[H_i, F_j] = -c_{ij} F_j \tag{156}$$

$$[E_i, F_j] = \delta_{ij} H_i \tag{157}$$

for all $i, j \in \{1, \dots, r\}$, as well as

$$\mathrm{ad}_{E_i}^{1-c_{ij}}(E_j) = 0 \tag{158}$$

$$\mathrm{ad}_{F_i}^{1-c_{ij}}(F_j) = 0 \tag{159}$$

for all distinct $i, j \in \{1, \dots, r\}$. The main result of this Lecture is the following theorem of Serre.

Theorem 20. *For any irreducible Cartan matrix C, the Lie algebra $\mathfrak{g}_C$ is finite-dimensional and simple. Its root system has associated Cartan matrix precisely equal to C.*

Moreover, it is easy to see that if $C = C_1 \oplus C_2$, then $\mathfrak{g}_C \cong \mathfrak{g}_{C_1} \oplus \mathfrak{g}_{C_2}$. Coupling this with Lemma 5 allows us to extend Theorem 20 to arbitrary Cartan matrices, by replacing the word "simple" with "semisimple".

Example 8. *When C is the Cartan matrix (148) of type A_{n-1}, the isomorphism $\mathfrak{g}_C \cong \mathfrak{sl}_n$ is given by*

$$E_i \rightsquigarrow E_{i,i+1}, \quad F_i \rightsquigarrow E_{i+1,i}, \quad H_i \rightsquigarrow E_{ii} - E_{i+1,i+1}$$

It is easy to check by hand that relations (154)-(159) hold in $\mathfrak{sl}_n$, which gives a homomorphism $\mathfrak{g}_C \to \mathfrak{sl}_n$. It is also easy to check that this homomorphism is surjective (because the matrices $E_{i,i+1}$ and $E_{i+1,i}$ generate $\mathfrak{sl}_n$ as a Lie algebra) and so it must be an isomorphism due to $\mathfrak{g}_C$ being simple.

12.3

We start by identifying relations (154)-(159) in any complex semisimple Lie algebra $\mathfrak{g}$. Fix a s.i.b.f. $(\cdot,\cdot)$ and a set of simple roots $\alpha_1, \dots, \alpha_r$ of $\mathfrak{g}$. As in Proposition 19, we can pick elements

$$E_{\alpha_i} \in \mathfrak{g}_{\alpha_i} \quad \text{and} \quad F_{\alpha_i} \in \mathfrak{g}_{-\alpha_i}$$

such that $[E_{\alpha_i}, F_{\alpha_i}] = H_{\alpha_i}$ determine an $\mathfrak{sl}_2$-triple, and so satisfy relations (154)-(157) with E_i replaced by E_{α_i} etc. To prove that these elements also satisfy (158), consider any $i \neq j$ and make

$$\bigoplus_{\ell \in \mathbb{Z}} \mathfrak{g}_{\alpha_j + \ell \alpha_i}$$

into a representation of $\mathfrak{sl}_2$ via the operators $\text{ad}_{E_{\alpha_i}}, \text{ad}_{F_{\alpha_i}}, \text{ad}_{H_{\alpha_i}}$. As we have seen in the proof of Propositions 20 and 21, the weight of the ℓ-th direct summand in the formula above is

$$\frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)} + 2\ell = c_{ij} + 2\ell$$

(check this fact). Since $\text{ad}_{F_{\alpha_i}}(E_{\alpha_j}) = 0$ by (157), then we conclude that E_{α_j} is a vector of lowest weight c_{ij}. By Corollary 1, this implies that

$$\text{ad}_{E_{\alpha_i}}^{1-c_{ij}}(E_j) = 0$$

which is precisely (158). Relation (159) is proved analogously. Therefore, we conclude that the assignments $E_i \mapsto E_{\alpha_i}$, $F_i \mapsto F_{\alpha_i}$, $F_i \mapsto F_{\alpha_i}, \forall i \in \{1, \dots, r\}$ determine a Lie algebra homomorphism

$$\mathfrak{g}_C \to \mathfrak{g} \tag{160}$$

where C is the Cartan matrix associated to $\mathfrak{g}$.

Proposition 28. *The E_{α_i} and F_{α_i} defined above generate any complex semisimple Lie algebra $\mathfrak{g}$, i.e. the homomorphism* (160) *is surjective.*

Proof. Since any root is a non-negative integer combination of positive roots, it suffices to prove the following statement in any complex semisimple Lie algebra $\mathfrak{g}$: if α and β are roots such that $\alpha + \beta$ is also a root, then

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta} \tag{161}$$

(the inclusion $\subseteq$ between the sets above is quite general, see (122)). To see this, one picks an $\mathfrak{sl}_2$ triple $E_\alpha, F_\alpha, H_\alpha$ for the positive root α, and uses it to construct a representation of $\mathfrak{sl}_2$ on

$$\bigoplus_{\ell \in \mathbb{Z}} \mathfrak{g}_{\beta + \ell\alpha}$$

In any representation of $\mathfrak{sl}_2$, higher weight subspaces are obtained from the action of E on lower weight subspaces. In the case at hand, since $\mathfrak{g}_{\alpha+\beta}$ has higher weight than $\mathfrak{g}_\beta$, then we must have

$$\mathfrak{g}_{\alpha+\beta} = \text{ad}_{E_\alpha}(\mathfrak{g}_\beta)$$

which precisely implies (161). □

12.4

Having showed that the homomorphism (160) is surjective, it will follow from Theorem 20 that it is an isomorphism: thus, there exists a unique simple complex Lie algebra with any given irreducible root system. As a stepping stone to proving Theorem 20, let us understand the Lie algebra

$$\widetilde{\mathfrak{g}}_C$$

freely generated by $\{E_i, F_i, H_i\}_{1 \leq i \leq r}$ modulo relations (154)-(157), as per Definition 27. This Lie algebra is graded by the **root lattice**

$$Q = \left\{ n_1\alpha_1 + \dots + n_r\alpha_r \middle| n_1, \dots, n_r \in \mathbb{Z} \right\} \tag{162}$$

via

$$\deg E_i = \alpha_i, \quad \deg F_i = -\alpha_i, \quad \deg H_i = 0 \tag{163}$$

We will write $Q^{\pm}$ for the non-negative integer span of the positive/negative roots.

Proposition 29. *Consider the following subalgebras of* $\widetilde{\mathfrak{g}}_C$

$$\widetilde{\mathfrak{n}}_C^+ \text{ is spanned by arbitrary iterated Lie brackets of } E_i\text{'s}$$
$$\widetilde{\mathfrak{n}}_C^- \text{ is spanned by arbitrary iterated Lie brackets of } F_i\text{'s}$$
$$\widetilde{\mathfrak{h}}_C \text{ is spanned by the } H_i\text{'s}$$

Then we have

$$\widetilde{\mathfrak{g}}_C = \widetilde{\mathfrak{n}}_C^+ \oplus \widetilde{\mathfrak{h}}_C \oplus \widetilde{\mathfrak{n}}_C^-$$

with $\widetilde{\mathfrak{n}}_C^{\pm}$ *lying in degrees* $Q^{\pm}\backslash 0$ *and* $\widetilde{\mathfrak{h}}_C^+$ *in degree 0.*

Proof. Any element $x \in \widetilde{\mathfrak{g}}_C$ is a linear combination of iterated Lie brackets of E's, F's and H's. Let us consider any such Lie bracket, and reduce it via anti-symmetry and the Jacobi identity and relations (154)-(157), so that it has the minimal number of E's, F's and H's. We must show that if so reduced, then x must either be just an iterated Lie bracket of E's, or an iterated Lie bracket of F's, or a single H. Indeed, assume for the purpose of contradiction that some part of the iterated Lie bracket in question involved

$$\dots [F_j, [E_{i_1}, \dots, [E_{i_{k-1}}, E_{i_k}] \dots]] \dots$$

(the reason why we do not assume there are any H's instead of the E's in the formula above is that they could be readily simplified by (155)). Then by repeated applications of the Jacobi identity, we could ensure that the innermost Lie bracket is $[F_j, E_{i_a}]$ for some a, which can be simplified using (157). □

Proposition 30. $\widetilde{\mathfrak{n}}_C^+$ *and* $\widetilde{\mathfrak{n}}_C^-$ *are freely generated by* $\{E_i\}_{1\leq i\leq r}$ *and* $\{F_i\}_{1\leq i\leq r}$*, respectively, while*

$$\widetilde{\mathfrak{h}}_C = \bigoplus_{i=1}^{r} \mathbb{C} H_i$$

Proof. Let us consider the tensor algebra TV of the vector space $V = \bigoplus_{i=1}^r \mathbb{C}v_i$. There is an action

$$\widetilde{\mathfrak{g}}_C \curvearrowright TV$$

given by

$$E_i \cdot (v_{j_1} \otimes \dots \otimes j_n) = \sum_{1\leq s\leq n \text{ s.t. } j_s = i} (c_{ij_{s+1}} + \dots + c_{ij_n})(v_{j_1} \otimes \dots \otimes v_{j_{s-1}} \otimes v_{j_{s+1}} \otimes \dots \otimes v_{j_n})$$
$$F_i \cdot (v_{j_1} \otimes \dots \otimes v_{j_n}) = v_i \otimes v_{j_1} \otimes \dots \otimes v_{j_n}$$
$$H_i \cdot (v_{j_1} \otimes \dots \otimes v_{j_n}) = -\sum_{s=1}^{n} (c_{ij_1} + \dots + c_{ij_n})(v_{j_1} \otimes \dots \otimes v_{j_n})$$

Check that the above action is well-defined, by verifying the Lie bracket relations (154)-(157). From the formula above, it is clear that the F_i do not satisfy any other Lie algebra relations between themselves other than the ones prescribed by the free Lie algebra (by the very Definition 26). The analogous statement for the E_i is proved likewise. Finally, because

$$H_i \cdot v_j = -c_{ij} v_j$$

then any linear relation $\sum_{i=1}^{r} \gamma_i H_i = 0$ would imply

$$\sum_{i=1}^{r} \gamma_i c_{ij} = 0$$

for all j. This is impossible, as the Cartan matrix C has positive determinant. □

12.5

Let us now consider the ideals

$$\mathfrak{i}^{\pm} \subset \widetilde{\mathfrak{n}}_C^{\pm}$$

generated by relations (158) and (159), respectively.

Proposition 31. *The direct sum* $\mathfrak{i} = \mathfrak{i}^+ \oplus \mathfrak{i}^-$ *is an ideal in* $\widetilde{\mathfrak{g}}_C$*, and we have*

$$\mathfrak{g}_C = \widetilde{\mathfrak{g}}_C \Big/ \mathfrak{i}$$

Proof. Let S_{ij}^+ and S_{ij}^- denote the LHS of (158) and (159), respectively. Prove the formulas

$$[F_k, S_{ij}^+] = 0, \quad \forall i, j, k \in \{1, \dots, r\} \tag{164}$$

$$[E_k, S_{ij}^-] = 0, \quad \forall i, j, k \in \{1, \dots, r\} \tag{165}$$

in $\widetilde{\mathfrak{g}}_C$ using repeated applications of the Jacobi identity and relations (155) and (157) (if you prefer, you can work in the universal enveloping algebra by the injectivity of (80), where $S_{ij}^{\pm}$ can be expressed as an alternating sum of binomial coefficients times $E_i^k E_j E_i^{1-c_{ij}-k}$ for $k \in \{0, \dots, 1-c_{ij}\}$). By (164)-(165) and the fact that $\mathfrak{i}^{\pm}$ are ideals inside $\widetilde{\mathfrak{n}}_C^{\pm}$, we conclude that $\mathfrak{i}^{\pm}$ are preserved under Lie bracket with all E's and F's. Because of (157) and the Jacobi identity, then $\mathfrak{i}^{\pm}$ are ideals of $\widetilde{\mathfrak{g}}_C$, and therefore so is their direct sum. We therefore obtain a surjective Lie algebra homomorphism

$$\widetilde{\mathfrak{g}}_C / \mathfrak{i} \twoheadrightarrow \mathfrak{g}_C \tag{166}$$

However, anything in the kernel of the above function would be a combination of iterated commutators of $S_{ij}^{\pm}$'s with E's and F's. By (164)-(165), any such commutator would already be in $\mathfrak{i}^{\pm}$, so (166) is an isomorphism. □

Proof. of Theorem 20: Let us start with a technical observation: consider the adjoint action of any $\mathfrak{sl}_2$-triple E_i, F_i, H_i on $\mathfrak{g}_C$. The Serre relations (158)-(159) precisely imply that the subrepresentation generated by any E_j or F_j is finite-dimensional, with weights in $\{c_{ij}, \dots, -c_{ij}\}$. However, if the representations generated by x and y are finite-dimensional, then so is the representation generated by $[x, y]$ (specifically, it would be spanned by Lie brackets of the basis vectors of the aforementioned two representations). We conclude that any element $\mathfrak{g}_C$ generates a finite-dimensional subrepresentation with respect to any $\mathfrak{sl}_2$-triple. By Proposition 31, we have a decomposition

$$\mathfrak{g}_C = \mathfrak{h} \bigoplus_{\beta \in Q} \mathfrak{g}_{C,\beta}$$

with respect to the grading (163). By Proposition 29, all the direct summands above are finite-dimensional, and moreover

$$\mathfrak{g}_{C,\beta} = 0 \quad \text{if } \beta \notin Q^{\pm} \tag{167}$$

and

$$\mathfrak{g}_{C,k\alpha_i} = \begin{cases} \mathbb{C}E_i & \text{if } k = 1 \\ 0 & \text{if } k > 1 \end{cases} \quad \text{and} \quad \mathfrak{g}_{C,-k\alpha_i} = \begin{cases} \mathbb{C}F_i & \text{if } k = 1 \\ 0 & \text{if } k > 1 \end{cases} \tag{168}$$

We want to show that

$$\dim \mathfrak{g}_{C,\beta} = \begin{cases} 1 & \text{if } \beta \in R^{\pm} \\ 0 & \text{otherwise} \end{cases} \tag{169}$$

To this end, choose any $\beta \in Q$ and consider the adjoint action of the $\mathfrak{sl}_2$ triple E_i, F_i, H_i on

$$T_{\beta,i} = \bigoplus_{\ell \in \mathbb{Z}} \mathfrak{g}_{C,\beta+\ell\alpha_i} \tag{170}$$

The ℓ-th direct summand above has weight with respect to H_i equal to

$$\frac{2(\alpha_i, \beta)}{(\alpha_i, \alpha_i)} + 2\ell$$

As explained in the first paragraph of the proof, any vector of $T_{\beta,i}$ generates a finite-dimensional $\mathfrak{sl}_2$ representation. Therefore, Corollary 2 implies that $T_{\beta,i}$ is finite-dimensional. As a consequence, Corollary 1 implies that its subspaces of opposite weights have the same dimension, so in particular

$$\dim \mathfrak{g}_{C,\beta} = \dim \mathfrak{g}_{C,s_i(\beta)} \tag{171}$$

where s_i is the simple reflection corresponding to α_i. We leave it to you to show that if β is not a multiple of a root, then there exists a sequence of reflections s_i that will land it in $Q \backslash (Q^+ \cup Q^-)$; in this case (167) would imply the bottom option in (169). On the other hand if β is a multiple of a root, then the last sentence in Theorem 18 implies that there is a sequence of reflections s_i that will make it into a multiple of a simple root; in this case (168) would imply the top option in (169).

We showed that $\mathfrak{g}_C$ is finite-dimensional, and that the dimensions of its graded subspaces are given by (169). It remains to prove that $\mathfrak{g}_C$ is simple, and to this end consider a non-zero ideal $\mathfrak{i} \subset \mathfrak{g}_C$. Because the Cartan matrix is invertible, the operators $\{\mathrm{ad}_{H_i}\}_{i \in \{1,\dots,r\}}$ act with disjoint spectrum on the root spaces of $\mathfrak{g}_C$. Since $\mathfrak{i}$ is preserved by the aforementioned operators, then if some element of $\mathfrak{i}$ has a non-zero coefficient in some $\mathfrak{g}_{C,\alpha}$, then we can assume that $\mathfrak{i}$ contains the subspace $\mathfrak{g}_{C,\alpha}$ in question. By the same logic as in (171), this implies that $\mathfrak{i}$ contains $\mathfrak{g}_{C,s_i(\alpha)}$ for all i. Because the Weyl group acts transitively on the set of roots, then we conclude that $\mathfrak{i} = \mathfrak{g}_C$. ☐

Lecture 13

13.1

We will now use the root system R associated to a complex semisimple Lie algebra $\mathfrak{g}$ to describe its complex representations. Fix a Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$ as in Lecture 9, a choice of positive/negative roots $R = R^+ \sqcup R^-$, and write $\alpha_1, \ldots, \alpha_r$ for the corresponding simple roots. Show that the following is an immediate consequence of Proposition 8, using the various $\mathfrak{sl}_2$-triples in $\mathfrak{g}$.

Proposition 32. *Any finite-dimensional representation $\mathfrak{g} \curvearrowright V$ has a* ***weight decomposition****, i.e.*

$$V = \bigoplus_{\lambda \in P} V_\lambda \tag{172}$$

where its ***weight subspaces*** *are*

$$V_\lambda = \left\{ v \in V \middle| x \cdot v = \lambda(x) v, \forall x \in \mathfrak{h} \right\} \tag{173}$$

and the direct sum in (172) *goes over the* ***(integral) weight lattice***

$$P = \left\{ \lambda \in \mathfrak{h}^* \middle| \frac{2(\lambda, \alpha_i)}{(\alpha_i, \alpha_i)} \in \mathbb{Z}, \forall i \in \{1, \ldots, r\} \right\} \tag{174}$$

Note that by the very definition of a root system, the weight lattice contains the root lattice (162)

$$P \supseteq Q \tag{175}$$

The two lattices are in general not equal (in fact, the quotient P/Q is a finite group whose order is equal to the determinant of the Cartan matrix). It is an exercise that the integrality condition on λ from (174) is equivalent to the a priori stronger condition $\frac{2(\lambda,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}$ for all roots α.

Example 9. *When $\mathfrak{g} = \mathfrak{sl}_n$, the weight lattice is*

$$P = \left\{ (k_1, \ldots, k_n) \in \mathbb{C}^n \middle| k_1 + \cdots + k_n = 0,\ k_i - k_{i+1} \in \mathbb{Z},\ \forall i \in \{1, \ldots, n-1\} \right\} \tag{176}$$

while the root lattice is

$$Q = \left\{ (k_1, \ldots, k_n) \in \mathbb{C}^n \middle| k_1 + \cdots + k_n = 0,\ k_i \in \mathbb{Z},\ \forall i \in \{1, \ldots, n\} \right\} \tag{177}$$

The fact that $|P/Q| = n$ comes about by noting that each k_i in (176) *must be congruent to $\frac{d}{n}$ modulo $\mathbb{Z}$, for one and the same value of $d \in \{0, \ldots, n-1\}$.*

13.2

Consider now the root decomposition

$$\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha$$

By definition, $\mathfrak{h}$ preserves the weight subspaces V_λ of any representation $\mathfrak{g} \curvearrowright V$. Moreover, an easy consequence of the Jacobi identity implies that

$$\boxed{\mathfrak{g}_\alpha \cdot V_\lambda \subseteq V_{\lambda+\alpha}} \tag{178}$$

Associated to our choice of positive roots $R = R^+ \sqcup R^-$, we write

$$\mathfrak{g} = \mathfrak{n}^+ \oplus \mathfrak{h} \oplus \mathfrak{n}^-$$

where $\mathfrak{n}^\pm = \oplus_{\alpha \in R^\pm} \mathfrak{g}_\alpha$. We will write

$$\mathfrak{b} = \mathfrak{n}^+ \oplus \mathfrak{h} \tag{179}$$

which is called a **Borel subalgebra** (Prove that $\mathfrak{b}$ is a solvable subalgebra of $\mathfrak{g}$; it is actually a Theorem of Borel-Morozov that it is a maximal solvable subalgebra).

Definition 28. *We say that $\lambda \in P$ is a **highest weight** for a representation $\mathfrak{g} \curvearrowright V$ if*

$$V_\lambda \neq 0 \quad \text{and} \quad V_{\lambda+\alpha_1} = \cdots = V_{\lambda+\alpha_r} = 0$$

A highest weight vector of P will be some non-zero $v \in V_\lambda$ as above.

Because of (178), any highest weight vector v satisfies

$$\mathfrak{n}^+ \cdot v = 0 \tag{180}$$

If the highest weight of v is λ, then we have

$$h \cdot v = \lambda(h)v, \quad \forall h \in \mathfrak{h} \tag{181}$$

Since one can always find a highest weight in any finite-dimensional representation, it is elementary to obtain the following.

Proposition 33. *Any finite-dimensional irreducible representation $\mathfrak{g} \curvearrowright V$ is a **highest weight representation**, i.e. it is generated by a highest weight vector.*

We will use the notion of highest weight to classify irreducible representations. As we have seen from Theorem 16, this would completely characterize the representation theory of complex semisimple Lie algebras, since any such representation uniquely decomposes as a direct sum of irreducible representations (moreover, in the next lecture, we will learn how to use characters in order to determine which particular irreducibles show up in the decomposition of any given representation).

13.3

Since irreducible representations are generated by highest weight vectors (as per Proposition 33), the first step in constructing them is to construct the universal representation satisfying (180) and (181).

Definition 29. *The **Verma module** with highest weight λ is*

$$M(\lambda) = U\mathfrak{g} \bigotimes_{U\mathfrak{b}} \mathbb{C} \tag{182}$$

where the tensor product is defined with respect to

- *the injection $U\mathfrak{b} \hookrightarrow U\mathfrak{g}$ of universal enveloping algebras corresponding to $\mathfrak{b} \hookrightarrow \mathfrak{g}$*

- *the surjection $U\mathfrak{b} \twoheadrightarrow \mathbb{C}$ which sends $\mathfrak{n}^+$ to 0 and every $h \in \mathfrak{h}$ to $\lambda(h)$.*

As (182) is an $U\mathfrak{g}$ module with respect to the left action, Subsection 6.1 implies that it is also a representation of $\mathfrak{g}$. Note that it is infinite-dimensional.

Remark. *You may recognize $M(\lambda)$ as being the induced representation $\mathrm{Ind}_{U\mathfrak{b}}^{U\mathfrak{g}}(\mathbb{C}_\lambda)$, where $\mathbb{C}_\lambda$ is the one-dimensional representation of $\mathfrak{b}$ corresponding to $\mathfrak{n}^+$ acting by 0 and $\mathfrak{h}$ acting by the weight λ. Indeed, in Math 314 you studied induced representations of finite groups, but the situation for infinite-dimensional algebras such as $U\mathfrak{g}$ is analogous.*

If we let v_λ denote a non-zero vector in the one-dimensional representation $\mathbb{C}_\lambda$, then any element of $M(\lambda)$ is of the form xv_λ for some $x \in U\mathfrak{g}$. However, because of the PBW Theorem 10, we have an isomorphism

$$U\mathfrak{g} = U\mathfrak{n}^- \otimes U\mathfrak{h} \otimes U\mathfrak{n}^+ = U\mathfrak{n}^- \otimes U\mathfrak{b}$$

Therefore, any element $x \in U\mathfrak{g}$ is a linear combination of products of elements from $U\mathfrak{n}^-$ and elements of $U\mathfrak{b}$. Since any element of $U\mathfrak{b}$ acts on v_λ by multiplying it with a constant, while elements of $U\mathfrak{n}^-$ act on V_λ freely, the assignment $xv_\lambda \to x$ yields a vector space isomorphism

$$\boxed{M(\lambda) \cong U\mathfrak{n}^-} \tag{183}$$

Moreover, the dimension of the weight spaces of the above isomorphism match up: any element $x_1 \dots x_n v_\lambda$ (for various $x_k \in \mathfrak{g}_{-\beta_k}$) lies in the $\lambda - \beta_1 - \dots - \beta_n$ weight subspace of $M(\lambda)$, by repeated applications of (178). Therefore, (183) and the PBW Theorem 10 tell us that

$$\dim M(\lambda)_\mu = \left| \left\{ \text{unordered positive roots } \beta_1, \dots, \beta_n \text{ with sum } \lambda - \mu \right\} \right| \tag{184}$$

In particular, $M(\lambda)$ has finite-dimensional weight subspaces.

13.4

Let $L(\lambda)$ be any irreducible representation of $\mathfrak{g}$ generated by a vector of highest weight $\lambda \in \mathfrak{h}^*$. There exists a homomorphism of $\mathfrak{g}$ representations

$$\pi : M(\lambda) \twoheadrightarrow L(\lambda)$$

defined by sending v_λ to a highest weight vector of $L(\lambda)$ (check that this homomorphism is well-defined and surjective, using (180) and (181)). By definition, the kernel of π is a proper $\mathfrak{g}$ subrepresentation of $M(\lambda)$ that is graded by weights $\mathfrak{h}^*$, and the irreducibility of $L(\lambda)$ implies that it is a maximal such proper graded subrepresentation.

Proposition 34. *Up to isomorphism, there exists a unique irreducible representation $\mathfrak{g} \curvearrowright L(\lambda)$ generated by a vector of highest weight $\lambda \in \mathfrak{h}^*$.*

The Proposition is an easy consequence of the fact that $M(\lambda)$ has a unique maximal graded subrepresentation (simply take the sum of all graded proper subrepresentations, which does not coincide with $M(\lambda)$, because it cannot contain v_λ). Therefore, we will refer to the representation $L(\lambda)$.

Theorem 21. *$L(\lambda)$ is finite-dimensional if and only if λ lies in*

$$P^+ = \left\{ \lambda \in P \ \Big| \ \frac{2(\lambda, \alpha_i)}{(\alpha_i, \alpha_i)} \in \mathbb{Z}_{\geq 0}, \forall i \in \{1, \dots, r\} \right\} \tag{185}$$

Such weights are called ***dominant****.*

For $\mathfrak{g} = \mathfrak{sl}_n$, a weight as in (176) is dominant if and only if $k_i - k_{i+1} \in \mathbb{Z}_{\geq 0}$ for all $i \in \{1, \dots, n-1\}$.

Proof. of Theorem 21: Assume that $L(\lambda)$ is finite-dimensional. Take an $\mathfrak{sl}_2$-triple E_i, F_i, H_i corresponding to any $i \in \{1, \dots, r\}$. Then $L(\lambda)$ is a finite-dimensional representation with respect to this $\mathfrak{sl}_2$, with the highest weight vector of $L(\lambda)$ having weight

$$\frac{2(\lambda, \alpha_i)}{(\alpha_i, \alpha_i)}$$

As we saw in Lecture 5, the numbers above must be non-negative integers in order to have a finite-dimensional representation of $\mathfrak{sl}_2$, so we conclude that $\lambda \in P^+$. Conversely, assume $\lambda \in P^+$. We will actually prove the following stronger claim on the weight subspaces of $L(\lambda)$

$$\dim L(\lambda)_\mu = \dim L(\lambda)_{w(\mu)} \tag{186}$$

for all $\mu \in P$ and $w \in W$. Indeed, because $L(\lambda)$ is a quotient of $M(\lambda)$, then its weight subspaces are finite-dimensional and only non-zero in the cone $\{\lambda - m_1\alpha_1 - \dots - m_r\alpha_r\}_{m_1,\dots,m_r \geq 0}$. If there existed such a non-zero root subspace with $m_1 + \dots + m_r$ arbitrarily large, then by applying formula (186) for the element $w \in W$ which sends the positive roots to negative roots (Corollary 7), then we would conclude the existence of a non-zero subspace with weight $w(\lambda) + m'_1\alpha_1 + \dots + m'_r\alpha_r$ for arbitrarily large $m'_1 + \dots + m'_r$. As this is impossible, the only option is for $L(\lambda)$ to only have finitely many non-zero weight subspaces, hence it must be finite-dimensional.

Let us now prove (186). Let us consider an $\mathfrak{sl}_2$-triple E_i, F_i, H_i for every $i \in \{1, \dots, r\}$, and define

$$V \subseteq L(\lambda)$$

to consist of all vectors v on which the E_i's and F_i's act **locally nilpotently**, i.e.

$$E_i^N v = F_i^N v = 0$$

for all $i \in \{1, \dots, r\}$ and for some $N \geq 0$ which may repend on i and v. Firstly, the highest weight vector v_λ lies in V because for all $i \in \{1, \dots, r\}$ we have

$$E_i v_\lambda = 0 \quad \text{and} \quad F_i^{k+1} v_\lambda = 0, \text{ where } k = \frac{2(\lambda, \alpha_i)}{(\alpha_i, \alpha_i)}$$

(note that the second equality is non-trivial, but we leave it to you to show that $F_i^{k+1} v_\lambda$ is annihilated by all the E_j's, and thus would generate a proper highest weight subrepresentation of $L(\lambda)$ if it weren't zero). Secondly, V is preserved by the action of $\mathfrak{g}$, because for any $x \in \mathfrak{g}$ and any $i \in \{1, \dots, r\}$ we have

$$[E_i^N, x] = \sum_{M=0}^{N-1} \binom{N}{M} \underbrace{[E_i, [E_i, \dots, [E_i, x] \dots]]}_{N-M \text{ copies of } E_i} E_i^M$$

in $U\mathfrak{g}$, and analogously for F_i. Thus, if a vector v is annihilated by large enough powers of every E_i, then so if xv because the iterated commutators in the equation above will all be 0 if $N - M$ is large enough. The remarks labeled firstly and secondly above imply that V is a subrepresentation of $L(\lambda)$, hence $V = L(\lambda)$ due to the latter's irreducibility. Then let us consider any weights μ and $s_i(\mu)$ and define the subrepresentation (with respect to the $\mathfrak{sl}_2$-triple E_i, F_i, H_i) generated by the corresponding weight subspaces

$$L(\lambda)_\mu \oplus \cdots \oplus L(\lambda)_{s_i(\mu)}$$

By the discussion above, this subrepresentation is finite-dimensional, with $L(\lambda)_\mu$ and $L(\lambda)_{s_i(\mu)}$ having H_i-weights

$$\frac{2(\mu, \alpha_i)}{(\alpha_i, \alpha_i)} \quad \text{and} \quad \frac{2(s_i(\mu), \alpha_i)}{(\alpha_i, \alpha_i)} = -\frac{2(\mu, \alpha_i)}{(\alpha_i, \alpha_i)}$$

By Corollary 1, the two weight subspaces in question must have the same dimension, which yields (186) for $w = s_i$. Since the simple reflections generate the Weyl group, then (186) holds for all w. □

Together with Theorem 16, we conclude the following.

Corollary 8. *Any finite-dimensional representation $\mathfrak{g} \curvearrowright V$ is isomorphic to*

$$V \cong L(\lambda_1) \oplus \cdots \oplus L(\lambda_k)$$

for $\lambda_1, \ldots, \lambda_k \in P^+$.

13.5

Motivated by Theorem 21, we have the following.

Definition 30. *The **fundamental weights** $\omega_1, \ldots, \omega_r$ are defined such that*

$$\frac{2(\omega_j, \alpha_i)}{(\alpha_i, \alpha_i)} = \delta_{ij} \tag{187}$$

for all $i, j \in \{1, \ldots, r\}$.

Fundamental weights form a $\mathbb{Z}_{\geq 0}$-basis of the cone of dominant weights, meaning that any dominant weight is of the form $n_1\omega_1 + \cdots + n_r\omega_r$ for some $n_1, \ldots, n_r \in \mathbb{Z}_{\geq 0}$. This has the following effect on the representation theory: show that the tensor product

$$\mathfrak{g} \curvearrowright L(\omega_1)^{\otimes n_1} \otimes \cdots \otimes L(\omega_r)^{\otimes n_r}$$

has highest weight $n_1\omega_1 + \cdots + n_r\omega_r = \lambda$. By Corollary 8, we have

$$L(\omega_1)^{\otimes n_1} \otimes \cdots \otimes L(\omega_r)^{\otimes n_r} \cong L(\lambda) \bigoplus_{\beta \in Q^+ \backslash 0} L(\lambda - \beta)^{\oplus \text{multiplicities}}$$

where the multiplicities above can be construed as a generalization of the Clebsch-Gordan rule (71). The formula above implies that $L(\lambda)$ can be recursively constructed (up to irreducible representations of the form $L(\lambda - \beta)$ with $\beta \in Q^+ \backslash 0$) from tensor products of the irreducible representations corresponding to the fundamental weights. Thus it is in this sense that the fundamental weights "generate" the finite-dimensional representation theory of $\mathfrak{g}$.

Example 10. *Whereas the simple roots of* $\mathfrak{sl}_n$ *are* $\{e_i - e_{i+1}\}_{1 \le i \le n-1}$*, the fundamental weights are*

$$\omega_i = \frac{n-i}{n} \cdot (e_1 + \dots + e_i) - \frac{i}{n} \cdot (e_{i+1} + \dots + e_n)$$

The irreducible representation corresponding to ω_i *is none other than*

$$\boxed{\wedge^i \mathbb{C}^n}$$

Indeed, the highest weight vector of $\wedge^i \mathbb{C}^n$ *is* $v_1 \wedge \dots \wedge v_i$*, which is an eigenvector for the action of any* $x = (x_1, \dots, x_n) \in \mathfrak{h}$ *with eigenvalue*

$$x_1 + \dots + x_i = (\omega_i, x)$$

To see that $\wedge^i \mathbb{C}^n$ *is irreducible, take any linear combination of tensors*

$$v_{t_1} \wedge \dots \wedge v_{t_i} + \dots$$

where $t_1 < \dots < t_i$ *and the ellipsis stands for sequences lexicographically smaller than* $(t_1, \dots, t_i)$*. Then applying the operators* $\{E_{dt_d}\}_{1 \le d \le i, t_d \ne d}$ *in succession to the above linear combination will simply produce* $v_1 \wedge \dots \wedge v_i$*.*

Lecture 14

14.1

In Math 314, you saw that characters are certain functions on a finite group that completely determine its representations. For representations of semisimple Lie algebras, the analogous role is taken by the following notion.

Definition 31. *The **character** of a representation $\mathfrak{g} \curvearrowright V$ is the sum*

$$\chi_V = \sum_{\lambda \in P} (\dim V_\lambda) e^\lambda \tag{188}$$

where $\{e^\lambda\}_{\lambda \in P}$ are formal symbols.

Although e^λ is a formal symbol, it arises from the following construction. As per Subsection 3.5, the representation $\mathfrak{g} \curvearrowright V$ lifts to a representation of the simply connected Lie group

$$G \curvearrowright V$$

with Lie algebra $\mathfrak{g}$ (this G is also called semisimple). There is an abelian subgroup called **maximal torus**

$$H \subset G$$

with Lie algebra $\mathfrak{h}$, and integral weights lift to characters

$$\lambda : H \to \mathbb{C}^*$$

Then we have for all $t = e^x \in H$

$$\chi_V(t) = \sum_{\lambda \in P} (\dim V_\lambda) e^{\lambda(x)} = \mathrm{tr}(t|_V) \tag{189}$$

This is now closer to the usual definition of characters as traces of group elements acting in the representation V. Of course, you may object that (189) only measures the trace on elements of H and not of G. But because the trace is conjugation invariant, the formula above actually measures the trace on any conjugates of H, which are dense in G (think about SL_n and arbitrary conjugates of diagonal matrices).

14.2

Another reason why we prefer formal expressions like (188) to actual numbers like (189) is that the former also applies to infinite-dimensional representations V (with finite-dimensional weight spaces) while the latter only applies to finite-dimensional representations. For example, (184) implies that

$$\chi_{M(\lambda)} = \frac{e^\lambda}{\prod_{\alpha \in R^+}(1 - e^{-\alpha})} = \frac{e^{\lambda+\rho}}{\prod_{\alpha \in R^+}(e^{\frac{\alpha}{2}} - e^{-\frac{\alpha}{2}})} \tag{190}$$

(the reason for the shift in the numerator by e^ρ, where $\rho = \frac{1}{2}\sum_{\alpha \in R^+} \alpha$, will be made apparent in Theorem 22). To make the above formula precise, we expand the denominator as a power series

$$\frac{1}{1 - e^{-\alpha}} = 1 + e^{-\alpha} + e^{-2\alpha} + \dots$$

and use the following operations on the formal symbols e^λ:

$$e^\lambda e^\mu = e^{\lambda+\mu}$$
$$\overline{e^\lambda} = e^{-\lambda}$$

The motivation behind these operations is given by the following formulas, which we invite you to prove

$$\chi_{V \oplus V'} = \chi_V + \chi_{V'} \tag{191}$$
$$\chi_{V \otimes V'} = \chi_V \chi_{V'} \tag{192}$$
$$\chi_{V^\vee} = \overline{\chi_V} \tag{193}$$

with respect to direct sums, tensor products and dual representations (see (30), (31), (32))

14.3

We will now calculate the character of irreducible representations $L(\lambda)$ of a semisimple Lie algebra $\mathfrak{g}$. The key result is the following formula of Freudenthal, which allows one to recursively compute the dimensions of the weight spaces of any irreducible representation starting from the obvious

$$\dim L(\lambda)_\lambda = 1$$

Proposition 35. *For any $\mu \in P$, we have*

$$\Big((\lambda+\rho, \lambda+\rho) - (\mu+\rho, \mu+\rho)\Big) \dim L(\lambda)_\mu = 2 \sum_{\alpha \in R^+} \sum_{k=1}^{\infty} (\mu + k\alpha, \alpha) \dim L(\lambda)_{\mu+k\alpha} \tag{194}$$

where $\rho = \frac{1}{2} \sum_{\alpha \in R^+} \alpha$ (note that the sum in the RHS is actually finite).

Proof. Let us consider the Casimir element associated to the non-degenerate s.i.b.f. of $\mathfrak{g}$

$$C = \sum_{i=1}^{r} H_i H^i + \sum_{\alpha \in R^+} \frac{(\alpha, \alpha)}{2} (E_\alpha F_\alpha + F_\alpha E_\alpha)$$

where H_i and H^i are dual bases of $\mathfrak{h}$ (the latter basis can be readily expressed in terms of the former basis using the formulas $(H_\alpha, H_\beta) = \frac{4(\alpha,\beta)}{(\alpha,\alpha)(\beta,\beta)}$, but we will not need this).

Lemma 9. *C acts on $L(\lambda)$ via the scalar $(\lambda+\rho, \lambda+\rho) - (\rho, \rho)$.*

Proof. Since C is central, we know that it acts on $L(\lambda)$ as a scalar, so it remains to identify this scalar by calculating how C acts on the highest weight vector v_λ. Any $\mathfrak{sl}_2$-triple $E_\alpha, F_\alpha, H_\alpha$ will have the property that

$$E_\alpha v_\lambda = 0 \quad \Rightarrow \quad F_\alpha E_\alpha v_\lambda = 0$$

hence

$$E_\alpha F_\alpha v_\lambda = H_\alpha v_\lambda = \frac{2(\lambda, \alpha)}{(\alpha, \alpha)} v_\lambda$$

On the other hand, it is an easy manipulation with symmetric bilinear forms that

$$\sum_{i=1}^{r} H_i H^i v_\lambda = (\lambda, \lambda) v_\lambda \tag{195}$$

Since $(\lambda, \lambda) + \sum_{\alpha \in R^+} (\lambda, \alpha) = (\lambda+\rho, \lambda+\rho) - (\rho, \rho)$, the Lemma follows. □

Let us now consider any positive root $\alpha \in R^+$, and decompose $L(\lambda)$ into subrepresentations of the $\mathfrak{sl}_2$-triple $E_\alpha, F_\alpha, H_\alpha$. These will all be of the form

$$L(\lambda)_{\beta+\frac{m}{2}\alpha} \oplus L(\lambda)_{\beta+\frac{m-2}{2}\alpha} \oplus \dots \oplus L(\lambda)_{\beta-\frac{m-2}{2}\alpha} \oplus L(\lambda)_{\beta-\frac{m}{2}\alpha} \tag{196}$$

where we assume that m is maximal such that the above weight spaces actually appear in $L(\lambda)$. The shift by β is chosen so that $(\alpha, \beta) = 0$, which implies that as a representation of the $\mathfrak{sl}_2$-triple $E_\alpha, F_\alpha, H_\alpha$, the $\beta + \frac{d}{2}\alpha$ direct summand above has weight d. By the inclusion-exclusion principle, it is easy to see that the number of copies of the irreducible representation $\mathfrak{sl}_2 \curvearrowright L(n)$ in the above representation is

$$\dim L(\lambda)_{\beta+\frac{n}{2}\alpha} - \dim L(\lambda)_{\beta+\frac{n+2}{2}\alpha}$$

for all $n \geq 0$. Meanwhile, (64), (65), (66) tell us how $E_\alpha F_\alpha$ and $F_\alpha E_\alpha$ act on the direct summands above. Specifically, since $E_\alpha F_\alpha + F_\alpha E_\alpha$ acts on the d-th weight subspace of an irreducible representation $\mathfrak{sl}_2 \curvearrowright L(n)$ by the constant

$$\frac{n(n+2)}{2} - \frac{d^2}{2}$$

Let's assume $d \geq 0$ for simplicity. We conclude that $E_\alpha F_\alpha + F_\alpha E_\alpha$ acts on $L(\lambda)_{\beta+\frac{d}{2}\alpha}$ with trace

$$\sum_{n=d}^{\infty} \frac{n(n+2)}{2} \Big(\dim L(\lambda)_{\beta+\frac{n}{2}\alpha} - \dim L(\lambda)_{\beta+\frac{n+2}{2}\alpha} \Big) - \frac{d^2}{2} \dim L(\lambda)_{\beta+\frac{d}{2}\alpha}$$

If we write $\mu = \beta + \frac{d}{2}\alpha \Leftrightarrow d = \frac{2(\mu,\alpha)}{(\alpha,\alpha)}$ and manipulate the telescoping sum above, we conclude that

$$\frac{(\alpha,\alpha)}{2}(E_\alpha F_\alpha + F_\alpha E_\alpha)$$

acts on $L(\lambda)_\mu$ with trace

$$(\mu,\alpha) \dim L(\lambda)_\mu + 2\sum_{k=1}^{\infty} (\mu + k\alpha, \alpha) \dim L(\lambda)_{\mu+k\alpha}$$

Summing over all $\alpha \in R^+$ and adding to the mix the fact (analogous to (195)) that $\sum_{i=1}^{r} H_i H^i$ acts in $L(\lambda)_\mu$ as multiplication with (μ, μ), we conclude that C acts on $L(\lambda)_\mu$ with trace

$$2(\mu,\rho) \dim L(\lambda)_\mu + 2\sum_{\alpha \in R^+} \sum_{k=1}^{\infty} (\mu + k\alpha, \alpha) \dim L(\lambda)_{\mu+k\alpha}$$

Comparing this with Lemma 9 implies (194). ☐

14.4

A purely algebraic manipulation (which you may find in §25.2 of Fulton-Harris) allows one to deduce from Freudenthal's formula (194) the following so-called **Weyl character formula**.

Theorem 22. *The character of any irreducible representation $\mathfrak{g} \curvearrowright L(\lambda)$ is given by the formula*

$$\chi_{L(\lambda)} = \frac{\sum_{w\in W} \operatorname{sgn}(w) e^{w(\lambda+\rho)}}{\prod_{\alpha \in R^+} (e^{\frac{\alpha}{2}} - e^{-\frac{\alpha}{2}})} \tag{197}$$

where $\operatorname{sgn} : W \to \{\pm 1\}$ *is the group homomorphism that sends the simple reflections s_i to -1 (it is a generalization of the sign homomorphism of S_n).*

Certain observations about (197) are in order.

1. The $w = e$ summand in the numerator of (197) yields precisely the character of the Verma module in (190), and this corresponds to the fact that $M(\lambda)$ contains a copy of the irreducible representation $L(\lambda)$ generated by the highest weight vector. Conversely, we can interpret the right-hand side of (197) as an alternating sum of the right-hand sides of (190). This underlies the famous BGG (Bernstein-Gelfand-Gelfand) resolution of $L(\lambda)$ as a complex of Verma modules.

2. The numerator of (197) is an antisymmetric expression with respect to the Weyl group action, i.e. the operation $\{e^\mu \rightsquigarrow e^{w(\mu)}\}_{w \in W}$, has the effect of multiplying the numerator of (197) by $\mathrm{sgn}(w)$. It is a general property of Coxeter groups that the denominator of (197) is also antisymmetric, and it divides the numerator, thus revealing the fact that the right-hand side of (197) is a linear combination of e^μ's (as expected).

3. As per the previous point, the right-hand side of (197) is a symmetric expression with respect to the Weyl group action, which gives an equivalent proof of (186).

4. We can get a formula for the dimension of $L(\lambda)$ by taking the evaluation of the right-hand side of (197) as $e^\mu \rightsquigarrow e^{\mu(0)}$ where 0 is the origin of $\mathfrak{h}$. This is strictly speaking ill-defined, since we get $\frac{0}{0}$. The way to resolve this issue is to evaluate the right-hand side of (197) as $e^\mu \rightsquigarrow e^{\mu(x)}$ for $x \in \mathfrak{h}$ tending to 0 along a generic line. If you do so appropriately, you will find that

$$\dim L(\lambda) = \prod_{\alpha \in R^+} \frac{(\lambda + \rho, \alpha)}{(\rho, \alpha)} \tag{198}$$

14.5

Finally, let us consider the whole discussion above for $\mathfrak{g} = \mathfrak{sl}_n$. For a weight $\lambda = (k_1, \dots, k_n)$ with $k_1 + \dots + k_n = 0$, we will write

$$e^\lambda = z_1^{k_1} \dots z_n^{k_n}$$

The Weyl group $W = S_n$ acts on monomials above simply by permuting the variables $z_1, \dots, z_n$. For the positive root $\alpha = e_i - e_j$, we have

$$e^\alpha = \frac{z_i}{z_j}$$

and $\rho = \frac{1}{2}(n-1, n-3, \dots, 3-n, 1-n)$. Therefore, the Weyl character formula reads (after some algebraic manipulations)

$$\chi_{L(\lambda)} = \frac{\sum_{w \in S_n} \mathrm{sgn}(w) \prod_{i=1}^n z_{w(i)}^{k_i+n-i}}{\prod_{1 \le i < j \le n} (z_i - z_j)}$$

One recognizes the denominator of the right-hand side as the Vandermonde determinant, and the entire right-hand side as the Schur function associated to the partition $(k_1, k_2, \dots, k_{n-1}, k_n)$. This underlies the Schur-Weyl duality between irreducible representations of $\mathfrak{sl}_n$ and those of symmetric groups, which as you have seen in Math 314, also have characters which can be interpreted as Schur functions.