Lecture notes: Basic group and representation theory

Thomas Willwacher

February 27, 2014

Contents

Chapter 1

Introduction

Organization

Exercise groups:
A-L: Martin Lohmann, HG E22,
M-Z: Emma Hovhannisyan, HG D 7.2.
Start of the exercises: 25.2.2013
Today: 1st exercise sheet, 15 Minutes discussion after the lecture.
Book: Fulton, Harris: Representation Theory: A First Course
Lecture Notes: Probably....

Motivation for Physics students

Essentially there are 3 general ways to obtain information about some physical system:

- Symmetries
- A small parameter (perturbation theory)
- Numerical study on the computer

In this course we will study the first set of tools, i. e., general techniques to extract all information that is contained in the symmetries of a system. The symmetries are organized into a group, and representation theory is used to extract the information.

Since symmetries abound in physics, these techniques are very important, e.g.,

- Classical Mechanics: Symmetries yield conserved quantities.
- Electrodynamics: Shape of Maxwell equations dictated by Lorentz symmetry
- Relativity. Shape of Einstein equations dictated by symmetry.
- Quantum Mechanics: Many QM systems like the hydrogen atom can only be solved because of symmetries. Furthermore, group and representation theory plays an inportant role concerning the statistics (Bose/Fermi) of multi particle systems.
- Solid state physics: Many substances have an approximately translation invariant (crystal) structure. A big part of solid state physics can be understood as the systematic exploitation of that symmetry (band structure, etc.)
- Quantum electronics: The allowed transitions between energy levels of atoms are governed by representation theory.
- Particle physics: Representation theory is essential and determines possible interactions, the particle content and conserved quantities.
- Physical Chemistry: Understanding wavefunctions, understanding vibration patterns of molecules.

Example 1.1. A methane molecule has tetrahedral (T_d) symmetry, what can we say about the vibrational modes (just based on the symmetry)?

Motivation for Math students

It's beautiful. (And the bread and butter of your work.)

1.1 Definitions

Definition 1.1. *A* group *is a set G together with a map*

$$\circ : G \times G \to G$$

called multiplication, such that:

1. $\forall g_1, g_2, g_3 \in G : (g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3)$. *(Associativity)*
2. *There is an element $\mathbb{1} \in G$ (the unit) such that $\forall g \in G : \mathbb{1} \circ g = g \circ \mathbb{1} = g$.*
3. *$\forall g \in G$ there is an element $g^{-1} \in g$ such that $g^{-1} \circ g = g \circ g^{-1} = \mathbb{1}$.*

In the following, we will often omit the symbol $\circ$ and write gh instead of $g \circ h$.

Exercise 1.1. Show that the unit is unique, i.e., if there is another element $\mathbb{1}' \in G$ such that $\forall g \in G$: $1' \circ g = g \circ 1' = g$, then $1 = 1'$. Show that it is enough to ask that $g \circ 1' = g$ (right unit) or $1' \circ g = g$ (left unit).

Exercise 1.2. Similarly, show that the inverse element is unique. Furthermore, show that $(gh)^{-1} = h^{-1}g^{-1}$.

Example 1.2.
- The simplest group is the one element set $\{1\}$.
- $\mathbb{Z}$ is a group with composition given by addition, $\circ = +$, and the unit element being $\mathbb{1} = 0$. The inverse of $n \in \mathbb{Z}$ is $-n$.
- $C_n := \mathbb{Z}/n\mathbb{Z}$ for $n = 1, 2, \ldots$ are groups, the *cyclic groups*.
- The dihedral group D_n ($n \geq 3$) is the symmetry group of the regular n-gon in the plane. It contains $2n$ elements, and is generated by σ (rotation by $2\pi/n$) and τ (reflection along the real axis) with relations $\sigma^n = \tau^2 = 1$, $\sigma\tau\sigma\tau = 1$.
- Let V be a $\mathbb{K}$-vector space, for $\mathbb{K}$ some field. The group $GL_{\mathbb{K}}(V)$ is the group of invertible $\mathbb{K}$-linear maps $V \to V$. We abbreviate $GL(n, \mathbb{K}) := GL(\mathbb{K}^n)$ as the group of invertible $n \times n$ matrices. We will omit the $\mathbb{K}$ if it is clear from the context, usually $\mathbb{K} = \mathbb{C}$. Note also that $GL(1, \mathbb{K}) \cong \mathbb{K}^\times$ is the group of invertible elements in $\mathbb{K}$. The most important case for us: $\mathbb{K} = \mathbb{R}, \mathbb{C}$.
- The orthogonal group $O(n) \subset GL(n, \mathbb{R})$ is the group of matrices R such that $R^T R = \mathbb{1}$, where $\mathbb{1}$ is the $n \times n$ identity matrix. The subgroup $SO(n) \subset O(n)$ is composed of those matrices of determinant 1. (The inverse is the same as the transpose for $O(n)$, $SO(n)$.)

Remark 1.1. There are several basic properties of groups:

- Some groups are finite, i.e., have finitely many elements, others are infinite. In the above list $\{1\}$, C_n, D_n, $O(1) \cong C_2$, $SO(1) \cong \{1\}$, $GL(n, \mathbb{K})$ for $\mathbb{K}$ finite are finite. In general, the cardinality of the underlying set is called the *order* of the group.
- Some groups are *abelian*, i.e., all its elements commute, i.e., $g \circ h = h \circ g$ for all $g, h \in G$. In the above list $\{1\}$, $\mathbb{Z}$, C_n, $O(1), SO(1), SO(2) \cong S^1, GL(1, \mathbb{K})$ are abelian.
- Note that $O(n)$ ($n \geq 2$) and $\mathbb{Z}$ are both finite, but in different ways: $\mathbb{Z}$ is "discrete", while $O(n)$ is "continuous". The proper way to discuss that difference is to consider topological groups. Then $\mathbb{Z}$ is a topological group with the discrete topology (all sets are open and closed) while $O(n)$ is a Lie group, i.e., a group whose underlying topological space is a manifold. (More on that later.)

Remark 1.2. Do not confuse the order of a group G with the *order of an element* $g \in G$, which is the smallest natural number n such that $g^n = \mathbb{1}$, or ∞. (Equivalently, it is the order of the subgroup generated by g.)

Note also that different notations for $\circ, \mathbb{1}$ or the inverse g^{-1} are typically used, for example $+, 0, -g$ (mostly for abelian groups), or $\cdot$ instead of $\circ$ or e or 1 instead of $\mathbb{1}$.

Definition 1.2. *Let G, H be groups. A group homomorphism $f : G \to H$ is a map of the underlying sets that respects the product, i.e., $f(gg') = f(g)f(g')$. It is called isomorphism if the underlying map of sets is a bijection.*

Remark 1.3. Note that if $f : G \to H$ is a group homomorphism $f(\mathbb{1}_G) = \mathbb{1}_H$ and $f(g^{-1}) = f(g)^{-1}$.

Proof: $f(\mathbb{1}_G) = f(\mathbb{1}_G^2) = f(\mathbb{1}_G)^2$. Multiply both sides by $f(\mathbb{1}_G)^{-1}$ to obtain $\mathbb{1}_H = f(\mathbb{1}_G)$. Similarly $\mathbb{1}_H = f(g^{-1}g) = f(g^{-1})f(g)$. Multiply from the right by $f(g)^{-1}$ to obtain $f(g)^{-1} = f(g)$.

Some further definitions that are often useful.

Definition 1.3. *A* subgroup *$H \subset G$ is a subset which is also a group under the multiplication of G. A subgroup $H \subset G$ is called a* normal subgroup *if $\forall g \in G, h \in H : ghg^{-1} \in H$.*

Exercise 1.3. Show that the for $H \subset G$ a subgroup the unit of H and inverse map are those of G.

Example 1.3. $SO(n) \subset O(n) \subset GL(n, \mathbb{R})$ are subgroups. Here $SO(n) \subset O(n)$ is a normal subgroup since $\det(RR'R^{-1}) = \det(R')$. However, for $n \geq 2$, $O(n) \subset GL(n, \mathbb{R})$ is not normal since in general $(ARA^{-1})^T ARA^{-1} = A^{-T} R^T A^T ARA^{-1} \neq \mathbb{1}$.

Definition 1.4. *For $H \subset G$ a subgroup, the left cosets of H in G are the sets*

$$gH := \{g \circ h \mid h \in H\}$$

for $g \in G$ and the right cosets are the sets

$$Hg := \{h \circ g \mid h \in H\}.$$

We denote the set of left (right) cosets by G/H ($H \setminus G$).

In general G/H and $H \setminus G$ are not groups, since

$$ghg'h' \notin (gg')H$$

in general. However:

Lemma 1.1. *If $H \subset G$ is a normal subgroup, then the left and right agree and $G/H (\cong H \setminus G)$ forms a group, the* quotient group.

Proof. First let us show that the left and right cosets of any $g \in G$ agree. We have to show that for all $h \in H$ $hg \in gH$ and $gh \in Hg$. But by normality $g^{-1}hg =: h' \in H$ and $ghg^{-1} =: h'' \in H$, hence

$$hg = gh' \in gH$$

and

$$gh = h''g \in Hg.$$

Now the multiplication $G/H \times G/H \to G/H$

$$gH \circ g'H := gg'H$$

is well defined since for any $h, h' \in H$

$$(gh)H \circ (g'h')H := ghg'h'H = gg'h''h'H = gg'H$$

where $h'' := (g')^{-1}hg' \in H$. The unit is the coset $\mathbb{1}H$. The inverse element of gH is the coset $g^{-1}H$. (Simple exercise: Verify the associativity, unit and inverse axioms.) □

Note that each coset has size $|H|$, since $hg = h'g \Rightarrow h = h'$. Since the group decomposes into its cosets, we arrive at the following statement.

Proposition 1.1. *Let G be a finite group and H be a subgroup. Then $|H|$ divides $|G|$. In particular, the order of any element divides $|G|$*

A corollary is the following.

Lemma 1.2. *Any group of prime order p is cyclic, i.e., isomorphic to C_p.*

Proof. The order of any element g must divide the group order, hence it must be 1 (only possible for the identity element) or p (otherwise). But in the latter case, the map $G \to C_p$, $g \mapsto \sigma$ is an isomorphism. □

Exercise 1.4.
- Given two groups G, H, the product group $G \times H$ is the set of pairs (g, h) together with the product $(g, h)(g', h') = (gg', hh')$.
- Given a group homomorphism $f : G \to H$ the kernel $\ker f := \{g \in G \mid f(g) = \mathbb{1}_H\} \subset G$ is a normal subgroup. Similarly, the image $Imf := \{f(g) \mid g \in G\} \subset H$ is a subgroup (not necessarily normal).

Exercise 1.5. Work out subgroups and normal subgroups and right/left cosets for D_4.

Exercise 1.6. Show that any group of prime order p is cyclic, i.e., isomorphic to C_p.

Solution. The order of any element must divide the group order, hence it must be 1 (for the identity element) or p (otherwise).

Exercise 1.7. Show that a group of prime power order p^n has a nontrivial center.

Solution. Consider the conjugacy classes $C_1, \dots C_m$. Each $|C_j|$ must divide the group order and hence $|C_j| = p^{n_j}$ for some $n_j < n$. The conjugacy classes of central elements have size 1, and there is at least one such, given by the identity. But since

$$p^n = |G| = \sum_j |C_j| = |Z(G)| + \sum_j{}' |C_j| = |Z(G)| + \sum_j{}' p^{n_j}$$

where the prime indicates that the ranges only over j such that $n_j > 1$. It follows that $|Z(G)| \equiv 0$ modulo p and since $|Z(G)| > 0$, $Z(G) \geq p$.

Exercise 1.8. Show that any group of order p^2 (p prime) is abelian.

Solution. Two proofs: 1. If there is an element of order p^2 we are done. Otherwise, pick an element $x \neq 1$ (necessarily of order p) in the center. Pick another element y not in the group generated by x. We claim that the elements $x^a y^b$, $0 \leq a, b < p$ are all distinct. Indeed, if $x^a y^b = 1$ then by taking an appropriate power $y = x^{a'}$ contradicting the choice of y. Hence we exhaust all p^2 elements of the group and since the elements $x^a y^b$ commute, the group is abelian.

2. $Z(G)$ must have order p or p^2. In the latter case we are done. In the former case $G/Z(G)$ has order p and is hence cyclic, say generated by $[x]$. Hence any element of G has the form $x^a z$, with z in the center. But all such elements commute.

1.2 Actions and the orbit-stabilizer Theorem

Let S be a set. The set of bijective maps $S \to S$, $\mathrm{Aut}(S)$, forms a group. An *action* of a group G on the set S is a group homomorphism

$$\rho : G \to \mathrm{Aut}(S).$$

Let $s \in S$ and $g \in G$. We will abbreviate

$$g \cdot s := \rho(g)(s)$$

The *orbit* of $s \in S$ is the set

$$G \cdot s := \{g \cdot s \mid g \in G\}.$$

The *stabilizer* of $s \in S$ is the subgroup

$$\mathrm{Stab}_s := \{g \in G \mid g \cdot s = s\}$$

of elements of G that fix s.

Example 1.4. Let $H \subset G$ a subgroup. H acts on the set G by left multiplication, and by right multiplication by the inverse, i.e., $h \cdot g = hg$ or $h \cdot g = gh^{-1}$ for $g \in G, h \in H$. The orbits are the right and left cosets. The stabilizer of any element $g \in G$ is trivial, since $hg = g \Rightarrow h = \mathbb{1}$.

Definition 1.5. *The action of a group G on itself by the formula*

$$g' \cdot g := g'g(g')^{-1}$$

is called the adjoint action. The orbits of $g \in G$ under this action is called the conjugacy class of g. Two elements g, g' of a group G are conjugate *if they live in the same conjugacy class, i.e., if there is an element $h \in G$ such that*

$$g = hg'h^{-1}.$$

The stabilizer of an element $g \in G$ under this action is called the centralizer *of G.*

Remark 1.4. A function $f : G \to S$, S some set, is called a *class function* if it is constant on conjugacy classes.

The following is a simple but very often needed result:

Theorem 1.1 (Orbit-Stabilizer Theorem)**.** *Let G be a finite group, S be a finite set and $s \in S$ be arbitrary. Then*

$$|G| = |\mathrm{Stab}_s||G \cdot s|.$$

Proof. Denote the elements in the orbit of s by $s = s_1, \dots s_n$, where $n = |G \cdot s|$. Claim: The sets

$$G_j := \{g \in G \mid g \cdot s = s_j\}$$

all have the same cardinality. Clearly, left multiplication by one element $g_j \in G_j$ induces a map between $G_1 = \mathrm{Stab}_s$ and G_j. This map is invertible with inverse given by left multiplication by g_j^{-1} and hence a bijection. But since since any element of G must map s somewhere, $G = \cup_j S_j$ and hence

$$|G| = \sum_{j=1}^{n} |G_j| = n|G_1| = |G \cdot s||\mathrm{Stab}_s|$$

□

Also the Theorem seems trivial, be advised that it is used very often, in often in surprising and not so trivial ways. Here is an example from combinatorics:

Exercise 1.9. Compute the number of isomorphism classes of simple (i.e., undirected, no multiple edges or short loops) graphs with n vertices.

Example 1.5. Apply the Theorem to the action of a subgroup $H \subset G$ by left multiplication. Since the stabilizers are trivial, we learn that any right coset has $|H|$ elements.

1.3 Generators and relations

Definition 1.6. *Let S be a set. The free group generated by S, $\mathcal{F}_{\mathrm{Grp}}(S)$, is the set of words in symbols $\{s, s^{-1} \mid s \in S\}$, modulo the identification that consecutive symbols ss^{-1} or $s^{-1}s$ can be removed from the words. Multiplication is concatenation of words.*

For example, for $S = \{a, b\}$, some words are

$$\emptyset, a, b, a^{-1}, ab^{-1}a, \dots.$$

The multiplication is concatenation, where we identify a word with s followed or preceded by s^{-1} with the word with these two letters deleted. For example

$$ab^{-1}a \circ a^{-1}baa = aaa.$$

The unit is the empty word and the inverse is obtained by replacing the letter s by s^{-1} and vice versa, and flipping the order of letters, i.e.,

$$(a^{-1}baa)^{-1} := a^{-1}a^{-1}b^{-1}a.$$

Definition 1.7. *Let S be a set (of* generators*), and let $R \subset \mathcal{F}_{\mathrm{Grp}}(S)$ be a set (of* relations*). Then the group generated by S with relations R is the group*

$$G = \langle S \mid R \rangle := \mathcal{F}_{\mathrm{Grp}}(S)/\langle R \rangle$$

where $\langle R \rangle \subset \mathcal{F}_{\mathrm{Grp}}(S)$ is the smallest normal subgroup of $\mathcal{F}_{\mathrm{Grp}}(S)$ containing R.

Specifying a set of generators and relations is called a *presentation* of a group.

Concretely, we build $\mathcal{F}_{\mathrm{Grp}}(S)/\langle R \rangle$ by identifying the substring $r \in R$ in any word with the empty substring. For example, if we impose the relation a^2 then

$$a^{-1}baa = a^{-1}b = a^2a^{-1}b = ab.$$

The advantage of giving a presentation is that it is a very compact way of describing the group. The disadvantage is that in general there problem of deciding whether a given string of elements of $S \sqcup S$ is the identity or not (the word problem) is undecidable.

Exercise 1.10.
- A presentation of C_n is $\langle \sigma \mid \sigma^n = \mathbb{1} \rangle$.
- A presentation of D_n is $\langle \sigma, \tau \mid \sigma^n = \tau^2 = (\sigma\tau)^2 = \mathbb{1} \rangle$.
- A presentation of S_n is

$$\langle \sigma_1, \ldots, \sigma_{n-1} \mid \sigma_1^2 = \cdots = \sigma_{n-1}^2 = \mathbb{1}, \sigma_i\sigma_j = \sigma_j\sigma_i, \text{for } |i-j| > 1, \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1} \rangle.$$

1.4 Representations

Definition 1.8. *Let G be a group and V be a $\mathbb{K}$-vector space. Then a representation of G on V is a group homomorphism $\rho : G \to GL(V)$.*

A morphism (or intertwiner) between representations ρ and ρ' on V and V' is a linear map $f : V \to V'$ such that for all $g \in G$:

$$f \circ \rho(g) = \rho'(g) \circ f.$$

f is called an isomorphism of representations if the underlying map on vector spaces is an isomorphism.

A representation is an action on a vector space by linear maps.

Example 1.6.
- The representation $G \to GL(V)$, $g \mapsto \mathbb{1} \forall g$ is called the trivial representation.
- The defining representation of $GL(V)$ on V is the identity map $\rho : GL(V) \to GL(V)$. Similarly, there is the defining representation of $O(n)$ on $\mathbb{R}^n$.
- There is a representation of C_n on $\mathbb{C}$ by sending the generator σ to any n-th root of unity.
- There is a representaion of D_n on the real vector space $\mathbb{C} \cong \mathbb{R}^2$ by sending the generator σ to any n-th root of unity and the generator τ to complex conjugation. (Exercise: Check that this is a representation.)
- There are representations ρ of D_{2n} on the complex vector space $\mathbb{C}$ by setting $\rho(\omega) = \pm 1$ and $\rho(\tau) = \pm 1$.

Remark 1.5. Note that our definition of $GL(V)$ depends on $\mathbb{K}$ and hence so does the notion of representation. The relevant cases for us are $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

Definition 1.9. *Given a representation ρ of G on a vector space V, one may form the dual representation ρ^* on V^* by setting, for $g \in G$, $l \in V^*$ and $x \in V$*

$$(\rho^*(g)(l))(x) := l(\rho(g^{-1})x).$$

Let us check that this defines indeed a representation, i.e.,

$$(\rho^*(gh)(l))(x) = l(\rho((gh)^{-1})x) = l(\rho(h^{-1})\rho(g^{-1})x) = \rho^*(g)(l \circ \rho(h^{-1}))(x) = \rho^*(g)(\rho^*(h)(l))(x).$$

Definition 1.10. *Given two representations ρ, ρ' of G on vector spaces V, V', one may form*

- *The direct sum representation $\rho \oplus \rho'$ on $V \oplus V'$, where*
$$(\rho \oplus \rho')(g)(v + v') := \rho(g)(v) + \rho'(g)(v')$$
for $v \in V, v' \in V'$.

- *The tensor product representation $\rho \otimes \rho'$ on $V \otimes V'$ such that*
$$(\rho \otimes \rho')(g)(v \otimes v') := \rho(g)(v) \otimes \rho'(g)(v')$$
for $v \in V, v' \in V'$.

- *The homomorphism representation on* $\mathrm{Hom}(V, V')$ *such that*
$$\rho_{\mathrm{Hom}}(g)f := \rho'(g) \circ f \circ \rho(g^{-1}).$$

Exercise 1.11. Verify that these are representations.

Exercise 1.12. Verify that $f \in \mathrm{Hom}(V, V')$ is an intertwiner iff it is fixed by all $g \in G$, i.e., $\rho_{\mathrm{Hom}}(g)f = f$ for all $g \in G$.

Exercises

Exercise 1.13. Show that every finitely generated abelian group G is a product of copies of the cyclic groups C_n and of copies of $\mathbb{Z}$, i.e.,
$$G \cong \mathbb{Z}^n \times C_{n_1} \times \cdots \times C_{n_k}$$
for some numbers $k, n, n_1, \ldots, n_k$.

Exercise 1.14. Show that the group of symmetries of a regular n-gon in $\mathbb{R}^3$ is $D_n \times C_2$.

1.5 Basic properties of representations, irreducibility and complete reducibility

Definition 1.11. *Fix a representation $\rho : G \to GL(V)$. A subspace $W \subset V$ is called* invariant *if $\rho(G)W \subset W$. The representation ρ is called* irreducible *if it does not have a non-trivial (i.e., other than $W = 0, W = V$) invariant subspace W.*

The representation ρ is called completely reducible *if if splits into a direct sum of irreducible representations.*

Example 1.7.

- The (real) representation of $SO(2)$ on $\mathbb{R}^2$ is irreducible. Proof: Suppose there was an invariant one dimensional subspace, say $W = \mathbb{R}(a, b)$. For W to be invariant $(a, b)^T$ would need to be an eigenvector of all rotation matrices. A contrdiction.

- However, the complexification of this representation on $\mathbb{C}^2$ is not irreducible, but completely reducible. Proof: Consider the vectors $(1, i)^T$ and $(1, -i)^T$. They form a bsis of C^2 and both are eigenvectors of the rotation R of radius ϕ with eigenvalue $e^{\mp i\phi}$. Hence the representation on $\mathbb{C}^2$ decomposes into the sum of two one-dimensional representations.

- The representation of $\mathbb{R}$ on $\mathbb{R}^2$ given by
$$\rho(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$
is not irreducible and not completely reducible. Proof: Consider the subspace $W = \mathbb{R}(a, b)^T$. It is invariant iff $b = 0$. Hence there is one invariant subspace but the representation cannot split into the sum of two invariant subspaces.

In the ideal world, all representations of the group G were completely reducible. In this course, we will mostly work under assumptions that make the ideal world a reality.

Proposition 1.2. *Let G be a finite group and ρ be a representation on the finite dimensional $\mathbb{K}$-vector space V, where $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Then V is completely reducible.*

Lemma 1.3. *Let G be a finite group and ρ be a representation on the finite dimensional $\mathbb{K}$-vector space V, where $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Let $W \subset V$ be an invariant subspace. Then there is an invariant subspace $U \subset V$ such that*

$$V = U \oplus W.$$

I.e., "every invariant subspace has an invariant complement".

Proof. Pick some inner product $\langle , \rangle$ on V. Then the averaged inner product

$$\langle u, v \rangle' = \sum_{g \in G} \langle \rho(g)u, \rho(g)v \rangle'$$

is G-invariant, i.e.,

$$\langle \rho(h)u, \rho(h)v \rangle' = \sum_{g \in G} \langle \rho(g)\rho(h)u, \rho(g)\rho(h)v \rangle = \sum_{g \in G} \langle \rho(gh)u, \rho(gh)v \rangle = \sum_{g \in G} \langle \rho(g)u, \rho(g)v \rangle = \langle u, v \rangle' .$$

Now the $\langle , \rangle'$-orthogonal complement of W, $W^\perp$, is again invariant. Hence setting $U = W^\perp$ shows the Lemma. □

Proof of the Proposition. We do an induction on the dimension of V. If V is irreducible, we are done. (This is in particular true for $\dim V = 0, 1$.) Otherwise, there is some invariant subspace W. By the Lemma, it has an invariant complement U, so $V = W \oplus U$. W and U have smaller dimension, so applying the induction hypothesis we are done. □

Remark 1.6. The assumption on finitieness of G is important as we saw. However, in the exercises you will lift the assumption that V is finite dimensional and weaken the assumption that $\mathbb{K} = \mathbb{R}, \mathbb{C}$ to $char\mathbb{K} \nmid |G|$.

Our goal for the next weeks will be: (i) Given a group G, list all irreducible representations of G. (ii) Produce an algorithm to decompose a given representation into irreducibles. (iii) Examine how the irreducibles behave under the operations on representations we saw above (mostly tensor product).

For now, let us also note the following corollary of the above proof.

Lemma 1.4. *Let G be a finite group acting on a vector space V (over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$). Then there is an inner product on V invariant under the G action. In other words, the representation matrices can be assumed to be orthogonal (or unitary).*

Exercise 1.15. Show that the dual of a finite dimensional irreducible representation is irreducible.

1.6 Schur's Lemmata

Proposition 1.3 (First Schur Lemma)**.** *Let $\rho : G \to GL(V), \rho' : G \to GL(V')$ be irreducible representations of the group G. Let $f \in \mathrm{Hom}(V, V')$ be an intertwiner. Then either $f \equiv 0$ or f is invertible and hence V and V' are isomorphic.*

Proof. The kernel and image of f are invariant subspaces (exercise). Hence they are either $= 0$ or the whole space. □

Proposition 1.4 (Second Schur Lemma)**.** *Let $\rho : G \to GL(V)$ be an irreducible representations of the group G on the finite dimensional complex vector space V (i.e., $\mathbb{K} = \mathbb{C}$). Let $f \in \mathrm{Hom}(V, V)$ be an intertwiner. Then f is a multiple of the identity $\mathbb{1}_V$.*

Proof. f has an eigenvector λ over $\mathbb{C}$, hence $f - \lambda \mathbb{1}_V$ is an intertwiner and not invertible. Hence by the first Schur Lemma $f - \lambda \mathbb{1}_V \equiv 0$. □

Corollary 1.1. *Every finite dimensional irreducible complex representation of an abelian group is 1-dimensional.*

Proof. Every element of the group acts as an intertwiner, and hence acts by a multiple of the identity. Hence if the representation is not 1-dimensional, it decomposes. □

Remark 1.7. Note that it is important here that $\mathbb{K} = \mathbb{C}$ (or algebraically closed). For example, for $\mathbb{K} = \mathbb{R}$ we saw that the defining representation of $SO(2)$ is irreducible, but not 1-dimensional.

Remark 1.8. For general $\mathbb{K}$, the first Schur Lemma says that $\mathrm{End}(V)$ is a division algebra, i.e., an algebra in which every element other than 0 is invertible. The classical Frobenius Theorem states that for $\mathbb{K} = \mathbb{R}$ there are only three finite dimensional division algebras (up to isom.), namely $\mathbb{R}, \mathbb{C}$ and the quaternions $\mathbb{H}$. Hence (as we will see later) real irreducible representations come in three sorts: real, complex and quaternionic.

Also note Wedderburn's little Theorem that any finite division algebra is a finite field.

Remark 1.9. The Schur Lemmata are key to many "physical" application of representation theory as follows: Often the goal is to diagonalize a certain (possible complicated) matrix H, typically the Hamiltonian. We know that the system has G symmetry, i.e., that we have a representation of G on the underlying vector space V and that H intertwines that representation. Decompose V according to irreducible components:

$$V \cong \oplus_j V_j \otimes \mathbb{C}^{n_j}.$$

Here the sum ranges over (isomorphism classes of) all irreducible representations of G. Then the Schur Lemmata tell us that H is block diagonal with respect to this composition, i.e., that H has the form

$$\oplus_j \mathbb{1}_{V_j} \otimes M_j$$

with some yet undetermined $n_j \times n_j$ matrices M_j. Note that this reduces the original problem of diagonalising H to diagonalising each much smaller matrix M_j. If one is lucky, $n_j = 1$ and M_j is already a (diagonal) 1×1 matrix. Note furthermore that if one has more symmetries in general (i) the V_j become higher dimensional and (ii) the multiplicities n_j become smaller. Hence one has more information that can be extracted from symmetries and the eigenvalues of H in general have to be more degenerate.

Remark 1.10. The subspace $V_j \otimes \mathbb{C}^{n_j}$ spanned by irreducible components isomorphic to the irrep V_j is also called the *isotypical component* corresponding to V_j.

Chapter 2

Finite groups and finite dimensional representations

In this chapter we assume without further mention that the occurring groups are finite, and all occuring representations are finite dimensional over $\mathbb{C}$.

2.1 Character theory

Definition 2.1. *Let $\rho : G \to GL(V)$ be a representation of the group G on the vector space G The character of ρ_V is the function $\chi : G \to \mathbb{K}$, $\chi_V(g) = tr(\rho(g))$.*

The character is clearly a class function, i.e., constant on conjugacy classes. Furthermore

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

and

$$\chi_{V \otimes V'} = \chi_V \chi_{V'},$$

while

$$\chi_{V^*}(g) = \overline{\chi}(V).$$

(To show the last point, assume the representation is unitary.)

The character table of the finite group G is a table with rows the isomorphism classes of irreducible representations of G, columns the conjugacy classes and entries the values of the character function.

Example 2.1. Take $G = C_3$. We saw that all irreducibles are one-dimensional. The character table is

1	σ	σ^3
1	1	1
1	ζ	ζ^2
1	ζ^2	ζ

where ζ is a non-trivial 3rd root of unity.

Exercise 2.1. Show that knowledge of the character of a representation ρ suffices to determine the eigenvalues of $\rho(g)$ for all $g \in G$. (Note also that $\rho(g)$ may be assumed to be unitary and hence diagonalizable, and all eigenvalues have norm 1.)

On the space of functions on G put the inner product

$$\langle u, v \rangle = \frac{1}{|G|} \sum_{g \in G} \bar{u}(g) v(g).$$

It induces an inner product on class functions

$$\langle u, v \rangle = \frac{1}{|G|} \sum_{C} |C| \bar{u}(C) v(C)$$

where the sum is over all conjugacy classes.

Theorem 2.1 (First orthogonality). *Let V, V' be irreducible complex representations of G. Then*

$$\langle \chi_V, \chi_V' \rangle = \begin{cases} 1 & \text{if } V \cong V' \\ 0 & \text{otherwise} \end{cases}$$

Remark 2.1. This Theorem solves the problem of decomposing a representation ρ (on V) into irreducibles. Namely if

$$V \cong \oplus_j V_j \otimes \mathbb{C}^{n_j}$$

then $\chi_V = \sum_j n_j \chi_j$. Hence the multiplicities n_j may be recovered as $n_j = \langle \chi_j, \chi_V \rangle$. This also shows that the character completely determines the representation.

We will show the Theorem by first proving another orthogonality result.

Theorem 2.2 (Great Orthogonality Theorem). *Let G be a finite group and ρ, ρ' be inequivalent irreducible representations on the finite dimensional $\mathbb{K} = \mathbb{C}$-vector spaces V, V' which we assume to be unitary (w.l.o.g. by 1.4). Pick some orthonormal bases and denote the entries of the representation matrices by $\rho(g)_{ij}, \rho'(g)_{kl}$ for $g \in G$. Then*

$$\frac{1}{|G|} \sum_{g \in G} \bar{\rho}(g)_{ij} \rho'(g)_{kl} = 0$$

for all i, j, k, l. Furthermore,

$$\frac{1}{|G|} \sum_{g \in G} \bar{\rho}(g)_{ij} \rho(g)_{kl} = \frac{1}{\dim V} \delta_{ik} \delta_{jl}$$

Proof. For fixed i, k the matrix M_{ik} with j, l entry

$$M_{ik;jl} = \frac{1}{|G|} \sum_{g \in G} \bar{\rho}(g)_{ij} \rho'(g)_{kl}$$

defines an intertwiner between the representations ρ and ρ':

$$\begin{aligned} \sum_{g \in G} \bar{\rho}(g)_{ij} \rho'(g)_{kl} \rho'(h)_{lm} &= \sum_{g \in G} \bar{\rho}(g)_{ij} \rho'(gh)_{km} \\ &= \sum_{g \in G} \bar{\rho}(gh^{-1})_{ij} \rho'(g)_{km} \\ &= \sum_{g \in G} \bar{\rho}(g)_{il} \bar{\rho}(h^{-1})_{lj} \rho'(g)_{km} \\ &= \rho(h)_{jl} \sum_{g \in G} \bar{\rho}(g)_{il} \rho'(g)_{km}. \end{aligned}$$

For the last equality we used the unitarity of ρ, and we used summation conventions throughout. It follows by the first Schur Lemma and the assumption that ρ, ρ' are inequivalent irreps that $M_{ik} = 0$. Furthermore, replacing ρ' by ρ we obtain an intertwiner M_{ik} of the representation ρ, which is a multiple of the identity by the second Schur Lemma,

$$M_{ik;jl} = \lambda_{ik} \delta_{jl}.$$

The same derivation with the roles of i, k and j, l interchanged yields

$$M_{ik;jl} = \lambda'_{jl} \delta_{ik}$$

for some constants λ'_{jl} and hence

$$M_{ik;jl} = \lambda \delta_{ik} \delta_{jl}$$

for some constant λ. This constant may be obtained by setting $i = k$ and summing over all i,

$$\lambda \dim V \delta_{jl} = M_{ii;jl} = \frac{1}{|G|} \sum_{g \in G} \rho(g^{-1} g)_{jl} = \frac{1}{|G|} \sum_{g \in G} \delta_{jl} = \delta_{jl}.$$

Hence $\lambda = \frac{1}{\dim V}$. ∎

Proof of Theorem 2.1. We set $i = j$ and $k = l$ in Proposition 2.2 and sum over i and k to obtain

$$\frac{1}{|G|}\sum_{g\in G}\bar{\chi}(g)\chi'(g) = \begin{cases} 0 & \text{if } \rho, \rho' \text{ are inequivalent} \\ \frac{1}{\dim V}\delta_{ik}\delta_{ik} = 1 & \text{if } \rho = \rho'. \end{cases}$$

Now neither the left nor the right hand side change when replacing ρ or ρ' by an equivalent (isomorphic) representation, and hence the last statement holds also for ρ, ρ' isomorphic (but maybe $\rho \neq \rho'$). □

Remark 2.2. The fact that the characters of two distinct irreps are othogonal holds for any field as long as $charG$ does not divide the group order. A single irrep may well have a norm distinct from 1 however.

Exercise 2.2. Let V be a finite dimensional complex representation of the finite group G. Show that the projector onto the isotypical component corresponding to the j-th irrep V_j is

$$\pi_j = \dim V_j \sum_{g\in G}\bar{\chi}_j(g)\rho(g).$$

2.2 Algebras

Definition 2.2. *Let $\mathbb{K}$ be a field. A $\mathbb{K}$-algebra is a $\mathbb{K}$-vector space A together with a bilinear map*

$$\mu : A \times A \to A$$

the multiplication *and a distinguished element $\mathbb{1} \in A$ (the* unit*) such that*

- *(associativity) For all $a, b, c \in A$, $\mu(a, \mu(b, c)) = \mu(\mu(a, b), c)$.*
- *(unit) For all $a \in A$, $\mu(\mathbb{1}, a) = \mu(a, \mathbb{1}) = a$.*

In the following we will just write ab instead of $\mu(a, b)$ for the product.

Example 2.2.
- The space of $n \times n$ matrices $Mat_{n\times n}(\mathbb{K})$ is a $\mathbb{K}$-algebra. In particular $\mathbb{K}$ itself is a $\mathbb{K}$ algebra.
- The polynomials $\mathbb{K}[X]$ form an algebra.
- The smooth functions on $\mathbb{R}^n$, $C^\infty(\mathbb{R}^n)$ form an algebra.
- The space of functions $S \to \mathbb{K}$, for S any set, forms a $\mathbb{K}$-algebra.

The algebra we are mainly interested in is the group algebra

$$\mathbb{K}[G]$$

of a group G. Its elements are formal $\mathbb{K}$-linear combinations of elements of G, i.e.,

$$\mathbb{K}[G] = \{\sum_{g\in G} c_g g \mid c_g \in \mathbb{K}\,\forall g \in G\}.$$

There is a multiplication on $\mathbb{K}[G]$ inherited from the multiplication in G:

$$(\sum_{g\in G} c_g g)(\sum_{h\in G} d_h h) := \sum_{g\in G, h\in H} c_g d_h (gh) = \sum_{g\in G}\left(\sum_{h\in H} c_{gh^{-1}} d_h\right) g.$$

The unit is just the element $\mathbb{1} \in \mathbb{K}[G]$.

Example 2.3. Consider $G = C_3$. Then elements of $\mathbb{C}[G]$ are of the form $\lambda_0\mathbb{1} + \lambda_1\sigma + \lambda_2\sigma^2$. The action of $\sigma \in C_3$ maps this element to

$$\lambda_0\sigma + \lambda_1\sigma^2 + \lambda_2\mathbb{1}.$$

For A, B algebras one can build thge direct sum $A \oplus B$ which is the vector space direct sum, with multiplication $(a + b)(a' + b') = aa' + bb'$.

Example 2.4. The direct sum $Mat_{n\times n}(\mathbb{K}) \oplus Mat_{m\times m}(\mathbb{K})$ are the n, m blockdiagonal matrices.

2.3 Existence and classification of irreducible representations

One immediate corollary of Theorem 2.1 is that there can be at most as many inequivalent irreducible representations as conjugacy classes in G, say c. We will see now that there are exactly c inequivalent irreducible representations, by an explicit construction. Consider the group algebra

$$\mathbb{C}[G].$$

$\mathbb{C}[G]$ is both a left and a right G-module. We consider at as a representation by multiplication from the left. (And get the *regular representation.*)

Proposition 2.1. *Let ρ be an irreducible complex representation of the finite group G on V. Then the isotypical component of the representation $\mathbb{C}[G]$ corresponding to ρ contains ρ exactly* $\dim V$ *times.*

Proof. Let the characters of ρ and $\mathbb{C}[G]$ be χ and χ' respectively. We have to show that

$$\langle \chi, \chi' \rangle = \dim V.$$

Note that

$$\chi'(g) = \sum_{\substack{h \\ gh=h}} 1 = \begin{cases} |G| & g = \mathbb{1} \\ 0 & \text{otherwise.} \end{cases}$$

Hence

$$\langle \chi, \chi' \rangle = \frac{1}{|G|}\bar{\chi}(1)|G| = \dim V.$$

This shows the Proposition. □

Let us list several immediate corollaries:

Corollary 2.1. *Let $V_1, \dots V_m$ be the (isomorphism classes of) irreps of G. Then*

$$|G| = \dim \mathbb{C}[G] = \sum_{j=1}^{c} (\dim V_j)^2.$$

In particular, there are only finitely many irreducible representations, up to isomorphism.

Corollary 2.2 (Peter-Weyl Theorem, finite group version)**.** *The matrix elements of the representation matrices of the irreducible representations (assumed to be unitary) form an orthogonal basis of the space of functions on G.*

Proof. We have seen before that they are non-zero and orthogonal and by the corollary that there are $|G|$ many, hence they form an orthogonal basis. □

Corollary 2.3 (Maschke's Theorem, $\mathbb{K} = \mathbb{C}$)**.** *Let G be a finite group, and $V_1, V_2, \dots, V_m$ be the irreducible complex representations of G. Then the canonical map of algebras*

$$\mathbb{C}[G] \to \oplus_{j=1}^{m} \mathrm{Hom}_{\mathbb{C}}(V_j, V_j)$$

is an isomorphism, and hence also an isomorphism as representations of G by left and right multiplication.

Proof for $\mathbb{K} = \mathbb{C}$. (Sketch) The map $\mathbb{C}[G] \to \oplus_j \mathrm{Hom}_{\mathbb{C}}(V_j, V_j)$ is clearly injective, for otherwise there were some non-zero $x \in \mathbb{C}[G]$ acting as 0 on $\mathbb{C}[G] \overset{\text{as rep.}}{\cong} \oplus V_j \otimes \mathbb{C}^{\dim V_j}$. But this is impossible, for example $x \cdot 1 = x \neq 0$.

To see surjectivity, consider the element

$$\sum_g \bar{\rho}_j(g)_{ab} g.$$

By the Great Orthogonality Theorem, the corresponding matrix $\sum_g \bar{\rho}_j(g)_{ab}\rho(g)$ has all entries 0, except for the one in the a, b position in the j-th block. But such matrices span $\oplus_j \mathrm{Hom}_{\mathbb{C}}(V_j, V_j)$ and hence we are done. □

Remark 2.3. A similar Theorem holds over a field $\mathbb{K}$ such that $char\mathbb{K} \nmid |G|$. Here $\mathrm{Hom}_{\mathbb{C}}(V_j, V_j)$ has to be replaced by $\mathrm{Hom}_D(V_j, V_j)$, where D is a division ring ($\mathrm{End}_G(V_j)$). Note further that $\mathrm{Hom}_D(V_j, V_j) \cong Mat_{n\times n}(D)$ where n is the dimension of V_j over D. Exercise: Prove this for $charK \nmid |G|$.

We may also consider $\mathbb{C}[G]$ as a representation with the adjoint action:

$$\rho(g)x = gxg^{-1}.$$

Let us compute the multiplicity n_1 of the trivial representation in this representation in two ways:

- Using Maschke's Theorem: By Schur's Lemma and an earlier exercise we know that $\mathrm{Hom}_{\mathbb{C}}(V_j, V_j)^G \cong \mathbb{C}$. Hence $n_1 = n_{irr}$, the number of irreducible representations.

- Using character theory: The character of the representation is

$$\chi(g) = \sum_{\substack{h\in G \\ ghg^{-1}=h}} 1 = |Z_g|$$

  where Z_g is the centralizer of g. Hence

$$n_1 = \langle \chi_1, \chi\rangle = \frac{1}{|G|}\sum_{j=1}^{n_{conj}} |C_j|\chi(C_j) = \frac{1}{|G|}\sum_{j=1}^{n_{conj}} |C_j||Z_{g_j}| = \sum_{j=1}^{n_{conj}} 1 = n_{conj}.$$

  Here the sum is over conjugacy classes, and g_j is a representative of the j-th conjugacy class. We also used the identity (orbit stabilizer theorem)

$$|G| = |C_j||Z_{g_j}|.$$

Hence we arrive at the following conclusion:

Proposition 2.2. *The number of complex irreducible representations of a finite group G is the same as the number of conjugacy classes, $n_{irr} = n_{conj}$.*

Corollary 2.4 (Second orthogonality relation)**.** *The columns of the character table are orthogonal. Moreover the column corresponding to the conjugacy class C has squared length $\frac{|G|}{|C|}$, which is the cardinality of the centralizer of one element of the conjugacy class.*

Proof. Recall from linear algebra that the columns of a square matrix M with orthonormal rows are automatically orthonormal (i.e., U unitary is equivalent to U^T unitary). Since the rows of the character table are orthonormal after rescaling the column corresponding to the conjugacy class C by $\sqrt{|C|/|G|}$, and since there are c (number of conjugacy classes) rows and hence the character table is square, the rescaled columns need also be orthonormal. □

Exercise 2.3. Show that the normal subgroups of G is the intersections of one or more subgroups of the form $\{g \in G \mid \chi(g) = \chi(\mathbb{1})\}$ for χ the character of an irreducible representation.

Solution. Since the eigenvalues of $\rho(g)$ have norm one, the condition $\chi(g) = \chi(\mathbb{1})$ implies that $\rho(g)$ is the identity. This in turn immediately implies that the subset above is a normal subgroup.

Now note that if $\sum_g \lambda_g \rho'(g) = \sum_g \lambda'_g \rho'(g)$ then $\lambda_g = \lambda'_g$ for ρ' the representation on $\mathbb{C}[G]$.

Let $G' \subset \tilde{G} \subset G$, where G' is the above subgroup and $\tilde{G}$ is the intersection of kernels of representations of G that are trivial on G'. Then the projectors

$$\sum_{g\in G'} \rho'(g), \sum_{g\in\tilde{G}} \rho'(g)$$

must be proportional. Hence, by the remark, $G = \tilde{G}$.

2.4 How to determine the character table – Burnside's algorithm

The easisest way to determine the character table in small examples is by using "ad hoc heuristics" to fill its entries, using (i) the number of conj. classes is the number of irreps (ii) the 1-entry is the dimension (iii) dimensions squared sum to the group order (iv) first and second orthogonality (v) knowledge of some representations, in particular the 1 dimensional ones including the trivial representation.

Example 2.5. Consider D_4. We saw that there are 5 conjugacy classes $\{1\}, \{\sigma^2\}, \{\sigma, \sigma^3\}, \{\tau, \tau\sigma^2\}, \{\tau\sigma, \tau\sigma^3\}$. Hence there are 5 irreps. One of them has to be the trivial representation. We furthermore know by Corollary 2.1 that the sum of squares of the dimensions has to be $|G| = 8$. Hence the only possible dimensions are $1, 1, 1, 1, 2$. Furthermore, we saw that 1 dimensional representations can be formed by setting $\rho(\tau) = \pm 1$, $\rho(\sigma) = \pm 1$. This yields 4 1 dimensional representations ad the character table looks like this:

1	σ^2	σ, σ^3	$\tau, \tau\sigma^2$	$\tau\sigma, \tau\sigma^3$
1	1	1	1	1
1	1	-1	1	-1
1	1	-1	-1	1
1	1	1	-1	-1
2	a	b	c	d

By second orthogonality we determine that $a = -2$, $b = c = d = 0$.

Exercise 2.4. Compute the character table of the quaternion group $\{-1, i, j, k| - (-1)^2 = i^2 = j^2 = k^2 = ijk = -1\}$. Show that it is identical to that of D_4 and conclude that the character table does not uniquely determine the group.

Exercise 2.5. Show that the character of a representation determines the representation up to isomorphism. (Find an algorithm to recover the representation.)

There is also a more systematic way to determine the character table. For A, B, C conjugacy classes of G define the numbers

$$M_{ABC} = \frac{1}{\sqrt{|B|}\sqrt{|C|}} |\{a \in A, b \in B | ab \in C\}|.$$

Also fix some ordering on the m conjugacy classes $C_1, \dots, C_m$ and define m many $m \times m$ matrices M_A with entries M_{ABC}. Note that the numbers M_{ABC} may be computed solely from knowing the conjugacy classes and group structure on G.

Proposition 2.3 (Burnside's algorithm)**.** *Let G be a finite group, $\rho : G \to GL(V)$ a (complex) d-dimensional irrep with character χ. Then for any conjugacy class A:*

$$M_A \begin{pmatrix} \sqrt{|C_1|}\chi(C_1) \\ \vdots \\ \sqrt{|C_m|}\chi(C_m) \end{pmatrix} = \frac{|A|\chi(A)}{d} \begin{pmatrix} \sqrt{|C_1|}\chi(C_1) \\ \vdots \\ \sqrt{|C_m|}\chi(C_m) \end{pmatrix}$$

In other words the vector on the right hand side is a joint eigenvector of the matrices M_A.

Remark 2.4. The proposition says that the rescaled columns of the character table are simultaneous eigenvectors of $M_{C_1}, \dots, M_{C_m}$. Furthermore note that the vectors of eigenvalues $(|C_j|\chi(C_j)/d)_j$, $(|C_j|\chi'(C_j)/d')_j$ are distinct for distinct irreducible charcters χ, χ' by first orthogonality.

Hence the problem of finding the character table is reduced to the standard numerical problem of jointly diagonalising m (necessarily commuting) matrices.

Also note that I chose the normalization so that the eigenvectors are orthogonal by the first orthogonality Theorem.

Proof. Note that the operators

$$F_C := \sum_{c \in C} \rho(c)$$

(for C a conjugacy class, ρ irreducible) commute with all representation matrices and hence must be multiples of the identity by the second Schur Lemma. Taking the trace we find

$$F_C = \frac{|C|\chi(C)}{d} \mathbb{1}$$

where $d = \dim V$. Compute

$$\frac{1}{d^2}|A||B|\chi(A)\chi(B)\mathbb{1} = F_A F_B = \sum_{a\in A}\sum_{b\in B}\rho(ab) = \sum_C M_{ABC}\sqrt{|B|}\sqrt{|C|}\frac{1}{|C|}\sum_{c\in C}\rho(c)$$
$$= \frac{\sqrt{|B|}}{d}\sum_C M_{ABC}\sqrt{|C|}\chi(C)\mathbb{1}.$$

Hence the stated eigenvector equality follows. Note that the vectors occurring on the right (as χ becomes the characters of the different irreps) are just the rescaled rows of the character table. Hence their $\frac{1}{\sqrt{|G|}}$-multiples form an orthonormal basis by first orthogonality. The corresponding eigenvalues form vectors which are different rescalings of the same rows. Hence they are independent and the eigenvectors are uniquely determined by the joint eigenvalue problem.

Since in the joint eigenbasis all matrices become diagonal, they also have to commute. □

2.5 Real and complex representations

One may ask which irreducible complex representations are real, i.e., are of the form $V_0 \otimes \mathbb{C}$ for V_0 a real representation. In other words, we ask whether a basis can be found such that all representation matrices become real.

Clearly, a necessary condition is that the character is real. However:

Exercise 2.6. Show that the quaternion group (see exc. **??**) has an irrep with real character, which is not the complexification of a real representation. (Here $V = \mathbb{C}^2$.)

Solution. The real rep. matrices would be in SO(2), which is abelian. But then the representation decomposes.

Theorem 2.3. *Let G be a finite group, and let V be a complex irreducible representation. Then exactly one of the following holds:*

- *(real case) $V \cong V_0 \otimes \mathbb{C}$, where V_0 is a real irreducible representation. The character is real in this case and* $\mathrm{End}_G(V_0) \cong \mathbb{R}$.
- *(complex case) $V \oplus V^* \cong V_0 \otimes \mathbb{C}$, where V_0 is a real irreducible representation. The character of V is not real and* $\mathrm{End}_G(V_0) \cong \mathbb{C}$.
- *(quaternionic case) $V \oplus V \cong V_0 \otimes \mathbb{C}$, where V_0 is a real irreducible representation. The character of V is real and* $\mathrm{End}_G(V_0) \cong \mathbb{H}$.

*The following quantity (*Frobenius-Schur indicator*) may be used to decide the case:*

$$FS(V) = \frac{1}{|G|}\sum_{g\in G}\chi(g^2) = \begin{cases} 0 & \textit{complex case} \\ 1 & \textit{real case} \\ -1 & \textit{quaternionic case} \end{cases}$$

Furthermore, any real irreducible representation of G is obtained in this manner.

Proof. We assume we have picked a basis for which the representation V is unitary. If the character of V is not real, then V is not isomorphic to V^*. Consider the representation $V \oplus V^*$ in the basis chosen. The representation matrices have the form

$$\begin{pmatrix} \rho(g) & 0 \\ 0 & \bar\rho(g) \end{pmatrix}.$$

Hence the following operator intertwines the representation matrices:

$$C\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \bar v \\ \bar u \end{pmatrix}.$$

Clearly $C^2 = \mathbb{1}$ and C is anti-linear. Hence, using C as the complex conjugation we obtain that

$$V \oplus V^* \cong V_0 \otimes \mathbb{C}$$

where V_0 is the fixed point set of C. The intertwiners of V_0 are intertwiners X of $V \oplus V^*$ such that $XC = CX$. By Schur and because X is an intertwiner:

$$X = \begin{pmatrix} \lambda \mathbb{1} & 0 \\ 0 & \bar{\mu}\mathbb{1} \end{pmatrix}.$$

Hence $XC = CX$ implies $\lambda = \mu$. So $\mathrm{End}(V_0) \cong \mathbb{C}$. Hence V_0 must be irreducible for otherwise $\mathrm{End}(V_0)$ contained non-invertible elements other than 0.

Next assume the character is real. In this case V and V^* are isomorphic. In yet other words, there is some matrix A such that for all $g \in G$

$$\bar{\rho}(g) = A\rho(g)A^{-1}.$$

It immediately follows that

$$\rho(g) = \bar{A}A\rho(g)(\bar{A}A)^{-1}.$$

Hence, by Schur $\bar{A}A = \lambda\mathbb{1}$. multiplying A by some number a changes λ to $\lambda|a|^2$. Hence we may assume $\lambda \in \{\pm 1\}$. Suppose $\lambda = 1$. Then the operator

$$C\colon v \mapsto \bar{A}\bar{v}$$

is antilinear square one and commutes with the representation matrices. Hence $V \cong V_0 \otimes \mathbb{C}$. The intertwiner of the real representation V_0 are matrices $\mu\mathbb{1}$ (by Schur) such that $C\mu = \mu C$, i.e., $\mu \in \mathbb{R}$. It follows that $\mathrm{End}(V_0) \cong \mathbb{R}$ and by the same reason as above V_0 is irreducible.

Finally, if $\lambda = -1$, then consider the operator

$$C\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \bar{A}\bar{v} \\ \bar{A}\bar{u} \end{pmatrix}$$

on $V \oplus V$. Clearly $C^2 = \mathbb{1}$ and C is antilinear and commutes with all representation matrices (exercise). Hence $V \oplus V \cong V_0 \otimes \mathbb{C}$. Let us compute the intertwiners. By Schur they correspond to matrices $Y = X \otimes \mathbb{1}$, where X is some 2×2 complex matrix. Reality demands (by$CY = YC$) that furthermore (exercise)

$$X = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$$

where $a, b \in \mathbb{C}$. These algebra of such matrices is called the quaternion algebra $\mathbb{H}$. Irreducibility follows as before.

Also note that A above is actually unitary: Since for all $g \in G$

$$\rho(g) = \rho(g)^{-\dagger} = (A\rho(g)A^{-1})^{-T} = A^{-T}\rho(g)^{-T}A^T = A^{-T}\bar{\rho}(g)A^T = A^{-T}A\rho(g)A^{-1}A^T$$

we have (by Schur) that $A^{-T}A$ is a multiple of $\mathbb{1}$. Using that $A^{-1} = \lambda\bar{A}$ we find that also $A^\dagger A = \lambda'\mathbb{1}$ must be a multiple of the identity. Taking the trace we find that $\lambda > 0$. Taking the determinant in the equation and in $\bar{A}A = \lambda\mathbb{1}$ we find that

$$(\lambda')^{\dim V} = |\det A|^2 = \lambda^{\dim V}.$$

Hence we learn $\lambda = 1$. (And actually that if $\lambda = -1$ then $\dim V$ must be even.)

Next we check the character formula. Use V and $V' = V^*$ in the great orthogonality theorem and sum over $j = k$ and $i = l$. For $V \not\cong V^*$, i.e., χ not real, we find that the result is zero (first case inthe character formula). For real V we may assume the representation matrices are real and hence find the second case in the character formula. In the last case, we note that the great orthogonality theorem says that the sum equals $-tr(A^\dagger A)/\dim V$ (exercise). By unitarity of A we are done.

The final statement is that those representations cover all real irreps of G. Indeed, suppose V_0 is some real irrep with (necessarily real) character χ. Assume that V_0 is not isomorphic to any irrep constructed above. Denote the $\mathbb{C}$-irreps of G by $V_1, V_2, \ldots$. By (real) orthogonality of the characters, χ has to be orthogonal to all χ_j (for V_j real or quaternionic) and to $\chi_j + \bar{\chi}_j$ for V_j complex. The orthogonal complement of these spaces is d-dimensional, where $2d$ is the number of non-real characters. It is spanned (over $\mathbb{C}$) by vectors $\chi_j - \bar{\chi}_j$ for j such that V_j is complex. Hence

$$\lambda = \sum_j \lambda_j(\chi_j - \bar{\chi}_j)$$

where we agree that only if $V_j^* = V_i$ only one of i, j appears in the sum. Hence λ_j is the number of times V_j appears in $V_0 \otimes \mathbb{C}$, while $-\lambda_j$ is the number of times V_j^* appears. Both have to be nonnegative, hence $\lambda_j = 0$. Hence $\lambda = 0$. Contradiction. □

An element $g \in G$ is called *real* if g^{-1} is conjugate to g.

Exercise 2.7. Check that if g is real, all elements in its conjugacy class are real. Check that the value of any character on g is real.

Proposition 2.4. *Let G be a finite group. The number of real characters is equal to the number of real conjugacy classes.*

Proof. Sending a conjugacy class C to C^{-1} yields a permutation P of the columns of the character table. Since $\chi(g^{-1}) = \bar{\chi}(g)$,

$$\bar{M} = MP$$

where M is the (matrix of) the character table. Similarly, sending an irrep V to V^* yields a permutation Q of the rows of the character table, sending each character to its complex conjugate.

$$\bar{M} = QM.$$

But then $MP = QM$ or $P = M^{-1}QM$. Hence P and Q are similar and have the same trace. But the trace of P is the number of real conjugacy classes, while the trace of Q is the number of real characters (or self-dual representations). □

Exercise 2.8 (Frobenius-Schur)**.** $\sum_V \dim V FS(V)$ is the number of involutions in G (i.e., elements $g \in G$ such that $g^2 = \mathbb{1}$).

Solution. Inserting the formula for FS gives

$$\sum_V \dim V FS(V) = \frac{1}{|G|} \sum_g tr_{\mathbb{C}[G]}(g^2).$$

And $tr_{\mathbb{C}[G]}(g^2) = 0$ unless $g^2 = \mathbb{1}$.

2.6 Induction, restriction and characters

Let G be a finite group and let $H \subset G$ be a subgroup. Let ρ be a representation of G on the vector space V. The restriction of ρ to H is the homomorphism

$$Res^G_H \rho : H \to G \to GL(V).$$

The character of the restricted representation is obviously the character of ρ, restricted to H.

Conversely, let ρ be a representation of H on V. The *induced representation* is the representation $Ind^G_H(\rho)$ of G on

$$V' := \mathbb{K}[G] \otimes_{\mathbb{K}[H]} V := (\mathbb{K}[G] \otimes V)/\langle xh \otimes v - x \otimes \rho(h)v \mid x \in \mathbb{K}[G], h \in H, v \in V\rangle$$

by left multiplication.

Exercise 2.9. Show that

$$Ind^G_H(\rho) \cong \mathrm{Hom}_H(\mathbb{K}[G], V) = \{f : \mathbb{K}[G] \to V \mid f(hx) = \rho(h)f(x) \forall h \in H\}$$

where the right hand side is a G-representation via $(g \cdot f)(x) = f(xg)$. Hint: The isomorphism from right to left sends $f \in \mathrm{Hom}_H(\mathbb{K}[G], V)$ to

$$\sum_{c \in H\backslash G} g_c^{-1} \otimes f(g_c).$$

where g_c are right coset representatives. (The isomorphism does not depend on their choice.) For $char\mathbb{K} = 0$ you can also take

$$\sum_{g \in G} g^{-1} \otimes f(g).$$

Proposition 2.5 (Character of induced representation). *Let G be a finite group and let $H \subset G$ be a subgroup. Let ρ be a representation of H on the $\mathbb{K}$-vector space V (with char$\mathbb{K} \nmid |G|$). The character χ' of the induced representation $Ind_H^G(\rho)$ satisfies*

$$\chi'(g) = \frac{1}{|H|} \sum_{\substack{x\in G\\ xgx^{-1}\in H}} \chi(xgx^{-1}).$$

Proof. Pick left coset representatives $g_1, \dots g_n \in G$, where $n = |G|/|H|$. Then a basis of V' is given by $g_i \otimes e_j$, where the e_j form a basis of V. Let g act on a basis element and suppose gg_i is in the coset of g_k, i.e., $g_k^{-1}gg_i \in H$

$$gg_i \otimes e_j = g_k(g_k^{-1}gg_i) \otimes e_j = g_k \otimes (g_k^{-1}gg_i) \cdot e_j.$$

Hence there is a contribution to the trace only if $k = i$. Overall, we obtain the formula

$$\chi(g) = {\sum_i}' \chi(g_i^{-1}gg_i)$$

where the prime indicates that we should sum only over those i such that $g_i^{-1}gg_i \in H$. But instead of just having g_i in the sum, we may as well sum over all x in the coset and divide by the size of the coset $|H|$.

$$\chi(g) = \frac{1}{|H|} \sum_{\substack{x\in G\\ xgx^{-1}\in H}} \chi(x^{-1}gx)$$

□

Remark 2.5. If *char*$\mathbb{K} \mid |G|$ one has at least the Mackey formula

$$\chi'(g) = \sum_{\substack{c\in H\backslash G\\ x_c g x_c^{-1}\in H}} \chi(x_c g x_c^{-1})$$

where the sum is over right cosets c and x_c is a representative of the coset c.

The most important special case is that V is the trivial representation, i.e., $\chi \equiv 1$.

Corollary 2.5. *The character of the induced representation of the trivial representation is:*

$$\chi(g) = \frac{|G|}{|H||C_g|}|C_g \cap H|$$

where C_g is the conjugacy class of g.

Proof. Consider the adjoint action of the group G on C_g. The number of $x \in G$ that map g to any fixed $g' \in C_g$ is $|\mathrm{Stab}_g|$, cf. the proof ofthe Orbit-Stabilizer Theorem. (Note also that $\mathrm{Stab}_g = Z_g$ is the centralizer.) Hence we have

$$\sum_{\substack{x\in G\\ xgx^{-1}\in H}} 1 = |\mathrm{Stab}_g||C_g \cap H| = \frac{|G|}{|C_g|}|C_g \cap H|$$

using the orbit stabilizer theorem again. □

An important result is the following:

Theorem 2.4 (Frobenius Reciprocity). *Let G be a finite group, $H \subset G$ a subgroup. Let χ be a character of G and $\tilde{\chi}$ be a character of H. Then*

$$\left\langle \chi, Ind_H^G \tilde{\chi} \right\rangle_G = \left\langle Res_H^G \chi, \tilde{\chi} \right\rangle_H .$$

Proof. It is a straightforward verification, using the character fomula above

$$\left\langle \chi, Ind_H^G \tilde{\chi} \right\rangle_G = \frac{1}{|G||H|} \sum_{g \in G} \sum_{\substack{x \in G \\ xgx^{-1} \in H}} \chi(g) \tilde{\chi}(xgx^{-1}) = \frac{1}{|G||H|} \sum_{h \in H} \sum_{g \in G} \sum_{\substack{x \in G \\ xgx^{-1} = h}} \chi(h) \tilde{\chi}(h)$$
$$= \frac{1}{|G||H|} \sum_{h \in H} \chi(h) \tilde{\chi}(h) |G| = \left\langle Res_H^G \chi, \tilde{\chi} \right\rangle_H .$$

□

Remark 2.6. If we think of Ind and Res as maps between class functions on G and H, the Theorem states that they are adjoint to each other.

Remark 2.7. In the standard form Frobenius reciprocity states that Ind_H^G and Res_H^G are adjoint functors between the categories of G representations and H representations. In other words that there is a natural isomorphism

$$\mathrm{Hom}_G(V, Ind_H^G(W)) \cong \mathrm{Hom}_H(Res_H^G(V), W)$$

where V is a representation of G and W is a representation of H.

Proof. The isomorphisms are as follows: $\phi \in \mathrm{Hom}_G(V, Ind_H^G(W))$ is sent to

$$v \mapsto \phi(v)(\mathbb{1})$$

and $\psi \in \mathrm{Hom}_H(Res_H^G(V), W)$ is sent to

$$v \mapsto (x \mapsto \psi(x \cdot v)).$$

□

2.7 Exercises

Chapter 3

Representation theory of the symmetric groups

The goal of this chapter is to construct all irreducible representations and the character table of the symmetric groups $S_n = Bij(\{1, \ldots, n\})$.

3.1 Notations

The symmetric groups are the groups $S_n = Bij(\{1, \ldots, n\})$. An element $\sigma \in S_n$ may be represented by a matrix in the following way ($n = 5$):

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 5 & 2 & 3 \end{pmatrix}$$

meaning that σ is the permutation that sends 1 to 4, 2 to 1, 3 to 5 etc. Of course, since the first line is fixed, we may just abbreviate this by

$$\sigma = \begin{pmatrix} 4 & 1 & 5 & 2 & 3 \end{pmatrix}.$$

Another representation is the so called cycle representation:

$$\sigma = (142)(35)$$

which one reads as 1 is sent to 4, 4 is sent to 2, 2 is sent back to 1; 3 is sent to 5, 5 is sent to 3. We call the ordered subsets of $\{1, \ldots, n\}$ which are cyclically permuted *cycles*. Here the cycles are (142) and (35). Note that the above notation is slightly redundant, since, e.g.

$$\sigma = (142)(35) = (53)(421)$$

i.e., one may cyclically permute the order of numbers in the cycles and change the order the cycles appear. Of course, one may get rid of the redundancy by demanding that each cycle starts with the lowest possible number, and cycles containing a lower number go first.

Note that the cycle decompositionis unique, up to permuting the numbers within cycles cyclically, and changing the order of cycles. In particular, the numbers i_k^σ, $k = 1, 2, \ldots$ counting the cycles of length k are unique. So, to each permutation σ one may associate unique non-negative integers

$$i_k^\sigma$$

$k = 1, 2, \ldots$, such that $\sum_k k i_k^\sigma = n$. In the example above, $i_2 = i_3 = 1$, while all other $i_k = 0$.

Exercise 3.1. Find the cycle decomposition and numbers i_k^σ for

$$\sigma = \begin{pmatrix} 1 & 4 & 5 & 8 & 3 & 2 & 7 & 6 \end{pmatrix}.$$

Note also that the i_k^σ clearly determine a partition of n, namely

$$n = \underbrace{1 + \cdots + 1}_{i_1^\sigma \times} + \underbrace{2 + \cdots + 2}_{i_2^\sigma \times} + \cdots.$$

3.2 Conjugacy classes

As usual, the first step in the construction of the character table is to determine the conjugacy classes.

Lemma 3.1. *Two representations $\sigma, \sigma' \in S_n$ are in the same conjugacy class if and only if $i_k^\sigma = i_k^{\sigma'}$ for all k.*

Proof. Suppose σ has the cycle decomposition

$$\sigma = (i_1 i_2 \cdots i_r)(i_{r+1} \cdots) \cdots (\cdots i_n).$$

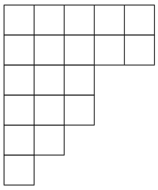
Then $\nu\sigma\nu^{-1}$ for $\nu \in S_n$ has the cycle decomposition (exercise)

$$\nu\sigma\nu^{-1} = (\nu(i_1)\nu(i_2) \cdots \nu(i_r))(\nu(i_{r+1}) \cdots) \cdots (\cdots \nu(i_n)).$$

Hence the cycle length are the same, but different numbers stand in the cycles. Hence the numbers I_k^σ are constant on conjugacy classes, and furthermore two permutations with the same i_k's can be, mapped onto each other by the adjoint action. □

Notation: We will abbreviate the conjugacy class corresponding to numbers $i_1, i_2, \dots$ by $C_{\mathbf{i}}$, where $\mathbf{i}$ is the vector $(i_1, i_2, \dots)$.

Hence the conjugacy classes of S_n are in one to one correspondence with the partitions of n. Each partition may be uniquely encoded by a *Young diagram* λ, for example the following diagram

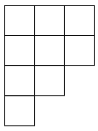


corresponds to the partition

$$19 = 5 + 5 + 3 + 3 + 2 + 1$$

and to the conjugacy class $C_{\mathbf{i}}$ with $\mathbf{i} = (1, 1, 2, 0, 2, 0, \dots, 0)$.

Exercise 3.2. Draw the Young diagrams corresponding to the conjugacy classes of the following elements: $\mathbb{1} \in S_n$, $(12)(3) \in S_3$. Write down a representative of the conjugacy class corresponding to the following diagram.



Convention: We will not distinguish between Young diagrams with n boxes and partitions of n. For a Young diagram λ, we will denote by λ_i the number of boxes in the i-th row. So the associated partition is $n = \lambda_1 + \lambda_2 + \cdots$.

Definition 3.1. *Let λ, λ' be Young diagrams with n boxes. We say that $\lambda \geq \lambda'$ if $\lambda = \lambda'$ or the first nonzero $\lambda_i - \lambda_i'$ is positive. We say $\lambda > \lambda'$ if $\lambda \geq \lambda'$ and in addition $\lambda \neq \lambda'$.*

Lemma 3.2. *The number of partitions $p(n)$ of n (which is also the number of conujugacy classes of S_n) is determined by the following formula*

$$\sum_{n \geq 0} p(n) t^n = \prod_{k \geq 1} \frac{1}{1 - t^k}.$$

Proof. Exercise. □

So our character table will have $p(n)$ columns and $p(n)$ rows.

Exercise 3.3. Show that all characters of S_n are real. (Hint: All group elements are real.) Use the Frobenius-Schur Theorem (exc. 4 of Serie 5) to compute the sum of the dimensions of irreps of S_n.

3.3 Irreducible representations

In fact, it turns out that the irreducible representations are also naturally labelled by Young diagrams λ. In the following we fix the number n as in S_n.

A *Young tableau* is a Young diagram filled with the numbers 1 to n, each occurring once. There is a natural way of filling these numbers into a Young diagram λ, from left to right and top to bottom.

1	2	3	4
5	6	7	8
9			

We call it the default tableau.

Definition 3.2. *For λ a Young diagram (or a Young tableau) we will denote by λ^T the same diagram (tableau) mirrored along the second diagonal.*

To each Young tableau $\hat{\lambda}$ we associate the following subgroup of S_n:

$$G_{\hat{\lambda}} = \{\sigma \in S_n \mid \sigma \text{ leaves the set of numbers in each row invariant}\}.$$

To be clear, this means that σ permutes the numbers within the first row etc., not that every element is fixed, of course.

Example 3.1. Consider the tableau

1	2
3	

.

Then the permutations in $G_{\hat{\lambda}}$ are the identity and (12)(3).

To each tableau $\hat{\lambda}$ we assign two elements of $\mathbb{C}[S_n]$.

$$s_{\hat{\lambda}} := \sum_{\sigma \in G_{\hat{\lambda}}} \sigma \qquad\qquad a_{\hat{\lambda}} := \sum_{\sigma \in G_{\hat{\lambda}^T}} (-1)^\sigma \sigma.$$

Furthermore we set for a a Young diagram λ:

$$s_\lambda := s_{\hat{\lambda}} \qquad\qquad a_\lambda := a_{\hat{\lambda}}$$

where $\hat{\lambda}$ is the left-to-right filled tableau for λ.

Definition 3.3. *The module V_λ corresponding to a Young diagram λ is the representation of S_n on*

$$\mathbb{C}[S_n]s_\lambda a_\lambda$$

by left multiplication.

Theorem 3.1. *The V_λ are irreducible and if $\lambda \neq \lambda'$ are two distinct Young diagrams then V_λ and V'_λ are not isomorphic.*

Note that this settles the question of constructing all irreps of S_n. Set $c_{\hat{\lambda}} = s_{\hat{\lambda}} a_{\hat{\lambda}}$ for a Young tableau $\hat{\lambda}$ and $c_\lambda = s_\lambda a_\lambda$ for a Young diagram λ.

Lemma 3.3. *For all $x \in \mathbb{C}[S_n]$ and λ a Young diagram, $c_\lambda x c_\lambda = \alpha c_\lambda$ for some constant α. Furthermore, if $\lambda' > \lambda$ is a different Young diagram then $c_\lambda x c_{\lambda'} = 0$ for all $x \in \mathbb{C}[S_n]$.*

Proof. It clearly suffices to consider only $x \in G \subset \mathbb{C}[G]$. Consider the second statement first. Note that $x c_{\lambda'} x^{-1} = c_{\tilde{\lambda}'}$ where $\tilde{\lambda}'$ is the tableau obtained from the default tableau by permuting the entries according to x. Hence it suffices to show that $c_\lambda c_{\tilde{\lambda}'} = 0$. Note that $\hat{\lambda}$ necessarily contains two numbers in the same column that are in the same row in $\tilde{\lambda}'$. Let τ be the transposition of those numbers. Then

$$a_\lambda s_{\tilde{\lambda}'} = a_\lambda \tau^2 s_{\tilde{\lambda}'} = -a_\lambda s_{\tilde{\lambda}'}$$

hence $a_\lambda s_{\tilde{\lambda}'} = 0$ and the second statement is shown.

Next, let us try to apply a similar same argument to $\lambda' = \lambda$. Consider $s_\lambda a_{\tilde{\lambda}}$. We know it is $= 0$ if two elements of one row in $\hat{\lambda}$ appear in the same column of $\tilde{\lambda}$. Hence in this case we are done. Otherwise, the elements of the first row of $\hat{\lambda}$ must appear in distinct columns of $\tilde{\lambda}$. Of course $a_{\tilde{\lambda}} = \pm a_{\tilde{\lambda}_1}$, where $\tilde{\lambda}_1$ is obtained from $\tilde{\lambda}$ by reordering the elements in each column so that the elements in the first row of $\hat{\lambda}$ appear in the first row of $\tilde{\lambda}_1$. Proceeding in this manner we may replace $\tilde{\lambda}$ by $\tilde{\lambda}_n$, that differs from $\hat{\lambda}$ only by some permutation $s \in G_{\hat{\lambda}}$. But clearly $s_\lambda s = s$. Summarizing, we find that for $x \in G$, $s_\lambda x a_\lambda$ is either 0 or $s_\lambda a_\lambda$ or $-s_\lambda a_\lambda$. Hence

$$c_\lambda x c_\lambda = s_\lambda (a_\lambda x s_\lambda) a_\lambda = \alpha c_\lambda$$

for some constant α. $\square$

Exercise 3.4. Identify $\mathbb{C}[G]$ with $A = \oplus_j \operatorname{End}(V_j)$ according to Maschke's Theorem. Show that an operator $p \neq 0$ such that for all $x \in A$: $pxp \propto p$ must be a rank one operator in one of the $\operatorname{End}(V_j)$. In particular, from this it follows that V_λ is irreducible.

Show further that if p, p' are such projectors that satisfy $pxp' = 0$, then they must live in different $\operatorname{End}(V_j)$. This shows the Theorem above. (Together with the fact that $c_\lambda \neq 0$, since the coefficient of $\mathbb{1}$ is 1.)

Remark 3.1. One can show in the same manner that the representations $V'_\lambda = \mathbb{C}[S_n]c'_\lambda$ with $c'_\lambda := a_\lambda s_\lambda$ are irreducible, and that V'_λ, V'_μ are not equivalent for $\lambda \neq \mu$. We claim that $V_\lambda \cong V'_\lambda$. By the preceding exercise, it is sufficient to check $c_\lambda c'_\lambda \neq 0$. But for this it is clearly sufficient to check that

$$c_\lambda c_\lambda a_\lambda = s_\lambda a_\lambda^2 s_\lambda a_\lambda = |G_{\lambda^T}|^2 s_\lambda a_\lambda s_\lambda a_\lambda = |G_{\lambda^T}|^2 \alpha c_\lambda.$$

where $\alpha \neq 0$.

From the remark one obtains the following result.

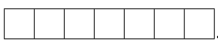
Proposition 3.1. *For any Young diagram λ: $V_\lambda \cong V_{\lambda^T} \otimes sgn$ (as representations of S_n).*

Proof. Consider $\mathbb{C}[S_n]$ as S_n-representation by left multiplication. There is an isomorphism of representations

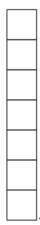
$$\begin{aligned} \mathbb{C}[S_n] &\to \mathbb{C}[S_n] \otimes sgn \\ \sum_\sigma c_\sigma \sigma &\mapsto \sum_\sigma (-1)^\sigma c_\sigma \sigma. \end{aligned}$$

The map $V_\lambda \hookrightarrow \mathbb{C}[S_n] \to \mathbb{C}[S_n] \otimes sgn$ is injective and hence an isomorphism onto its image. The image is $\mathbb{C}[S_n] a_{\lambda^T} s_{\lambda^T} \otimes sgn$. But we saw in the remark that $\mathbb{C}[S_n] a_{\lambda^T} s_{\lambda^T} = V'_{\lambda^T} \cong V_{\lambda^T}$ and hence we are done. $\square$

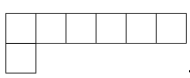
Example 3.2. The trivial representation corresponds to the Young diagrams (n boxes)

.

The alternating (sign) representation corresponds to the Young diagram (n boxes)

.

The representation of S_n on $\mathbb{C}^n$ splits into a sum of the trivial representation and the (standard) representation corresponding to the diagram (n boxes, 2 rows)

.

(Exercise of Serie 7)

3.4 The Frobenius character formula

In principle one may determine the character table from knowledge of the irreducible representations V_λ. However, there is a nice closed form expression for the characters.

Theorem 3.2 (Frobenius Character Formula). *Let λ be a Young diagram (a partition of n) and $C_{\mathbf{i}}$ be a conjugacy class of S_n.*

$$\chi_{V_\lambda}(C_{\mathbf{i}}) = \left(\Delta(x) \prod_k P_k^{i_k}(x) \right)_{x^{\lambda+\rho}}$$

where

- $x = (x_1, \dots, x_n)$, *and for a multiindex* $\lambda = (\lambda_1, \dots, \lambda_n)$,
$$x^\lambda = x_1^{\lambda_1} \cdots x_n^{\lambda_n}$$
as usual.
- $\Delta(x) = \prod_{1\leq i<j\leq n}(x_i - x_j)$ *is the Vandermonde determinant.*
- $P_k(x) = x_1^k + x_2^k + \cdots + x_n^k$.
- *the notation* $(Q)_{x^a}$ *means: take the coefficient of* x^a *in the polynomial* $Q(x)$.
- $\rho = (n-1, n-2, \dots, 1, 0)$.

Recall the Vandermonde formula:

Lemma 3.4.

$$\det \begin{pmatrix} x_1^{n-1} & x_1^{n-2} & \cdots & x_1 & 0 \\ \vdots & & \ddots & & \vdots \\ x_n^{n-1} & x_n^{n-2} & \cdots & x_n & 0 \end{pmatrix} = \prod_{i<j}(x_i - x_j).$$

Proof. Consider both sides as polynomials in x_1. The left hand side has zeros in $x_1 = x_j$ and hence has the form $P(x_2, \dots, x_n)\prod_{j>1}(x_1 - x_j)$. Since the degree of the polynomial on the left is n, th prefactor $P()$ cannot have any further dependence on x_1. Also note that on the left is the one smaller Vandermonde determinant. Hence by induction $P(x_2, \dots, x_n) = \prod_{1<i<j}(x_i - x_j)$ and we are done. □

For completeness, also recall the Cauchy determinant formula.

Lemma 3.5.

$$\det(\frac{1}{x_i - y_j})_{ij} = \det \begin{pmatrix} \frac{1}{x_1-y_1} & \frac{1}{x_1-y_2} & \cdots & \frac{1}{x_1-y_n} \\ \vdots & & \ddots & \vdots \\ \frac{1}{x_n-y_1} & \frac{1}{x_n-y_2} & \cdots & \frac{1}{x_n-y_n} \end{pmatrix}$$
$$= \frac{\prod_{i<j}(x_i - x_j)(y_j - y_i)}{\prod_{i,j}(x_i - y_j)} = (-1)^{n(n-1)/2}\frac{\Delta(x)\Delta(y)}{\prod_{i,j}(x_i - y_j)}.$$

Proof. The proof is similar to the one of the Vandermonde formula. Multiply both sides by the denominator of the right hand side. Then both sides become polynomials in x, y. The left hand side has roots when $x_1 = x_j$ or $y_1 = y_j$. Hence it must have the form

$$P(x_2, \dots, x_n, y_2, \dots, y_n)\prod_j (x_1 - x_j).$$

Again by degree counting the prefactor $P()$ is independent of x_1, y_1. But the prefactor of $x_1^{n-1}y_1^{n-1}$ is easily seen to be $\det(\frac{1}{x_{i+1}-y_{j+1}})_{ij}\prod_{i,j\geq 2}(x_i - y_j)$. □

Remark 3.2. There is also the following alternative form. Replacing x_j by $1/x_j$ and dividing both sides by $x_1 \cdots x_n$ we obtain

$$\det(\frac{1}{1 - x_i y_j})_{ij} = \frac{\Delta(x)\Delta(y)}{\prod_{i,j}(1 - x_i y_j)}.$$

Instead of working with the Specht modules V_λ directly, we rather work with the larger representations

$$U_\lambda := \mathbb{C}[G]s_\lambda = Ind_{G_\lambda}^{S_n}(\mathbb{C}).$$

Exercise 3.5. Show that indeed $\mathbb{C}[G]s_\lambda = Ind_{G_\lambda}^{S_n}(\mathbb{C})$.

Solution. We have a map $f : Ind_{G_\lambda}^{S_n}(\mathbb{C}) \to \mathbb{C}[G]s_\lambda$ sending $g \otimes 1$ to gs_λ. It is clearly well defined, since $gh \otimes 1$ maps to $ghs_\lambda = gs_\lambda$ for $h \in G_\lambda$. We also have a composition map

$$\bar{f} : \mathbb{C}[G]s_\lambda \to \mathbb{C}[G] \xrightarrow{\cdot \otimes 1} \mathbb{C}[G] \otimes \mathbb{C} \to \mathbb{C}[G] \otimes_{\mathbb{C}[G_\lambda]} \mathbb{C} \to Ind_{G_\lambda}^{S_n}(\mathbb{C}).$$

We claim $f \circ \bar{f} = |G|\mathbb{1}$ and $f \circ \bar{f} = |G|\mathbb{1}$. It is sufficient to check this on s_λ and $\mathbb{1} \otimes 1$, because these elements generate the representations.

$$f \circ \bar{f}(s_\lambda) = f(s_\lambda \otimes 1) = f(|G|\mathbb{1} \otimes 1) = |G|s_\lambda$$
$$\bar{f} \circ \bar{f}(\mathbb{1} \otimes 1) = \bar{f}(s_\lambda) = |G|\mathbb{1} \otimes 1$$

Clearly this representation contains $V_\lambda = \mathbb{C}[G]s_\lambda a_\lambda$, the map $U_\lambda \to V_\lambda$ being right multiplication by a_λ. The character of U_λ can be computed by Corollary 2.5.

Proposition 3.2.

$$\chi_{U_\lambda}(C_{\mathbf{i}}) = \left(\prod_{k=1}^{n} P_k^{i_k}(x) \right)_{x^\lambda}.$$

Proof. By Corollary 2.5 we have the formula

$$\chi_{U_\lambda}(C_{\mathbf{i}}) = \frac{|S_n|}{|G_\lambda||C_{\mathbf{i}}|}|C_{\mathbf{i}} \cap G_\lambda|.$$

Let us compute the quantities involved. First $|S_n| = n!$, $|G_\lambda| = \prod_\beta \lambda_\beta!$ and $|C_{\mathbf{i}}| = \frac{n!}{\prod_\alpha \alpha^{i_\alpha} i_\alpha!}$. Hence

$$\frac{|S_n|}{|G_\lambda||C_{\mathbf{i}}|} = \frac{\prod_\alpha \alpha^{i_\alpha} i_\alpha!}{\prod_\beta \lambda_\beta!}.$$

Consider $|C_{\mathbf{i}} \cap G_\lambda|$. First we have to sum over all ways of distributing i_α cycles of length α ($\alpha = 1, \ldots, n$) into n slots such that the β-th slot receives total cycle length λ_β. Suppose slot β gets j_α^β cycles of length β in some "distribution". Then we have to multiply by the number of elements in $|G_\lambda|$ that belong to that cycle distribution, namely

$$\prod_\beta \frac{\lambda_\beta!}{\prod_\alpha j_\alpha^\beta! \alpha^{j_\alpha^\beta}} = \frac{\prod_\beta \lambda_\beta!}{\prod_\beta \prod_\alpha j_\alpha^\beta! \alpha^{j_\alpha^\beta}} = \frac{\prod_\beta \lambda_\beta!}{\prod_\alpha \alpha^{i_\alpha} \prod_\beta j_\alpha^\beta!}$$

where we used that $\sum_\beta j_\alpha^\beta = i_\alpha$. Summarizing, we want to compute a sum over all distributions, each distribution j_α^β being counted with multiplicity

$$\frac{\prod_\alpha \alpha^{i_\alpha} i_\alpha!}{\prod_\beta \lambda_\beta!} \frac{\prod_\beta \lambda_\beta!}{\prod_\alpha \alpha^{i_\alpha} \prod_\beta j_\alpha^\beta!} = \frac{\prod_\alpha i_\alpha!}{\prod_\alpha \prod_\beta j_\alpha^\beta!}$$

Now let us check what the formula computes. Consider the coefficient of x^λ in $\prod_\alpha P_\alpha^{i_\alpha}$. It is also a sum of distributions of i_α cycles of length α ($\alpha = 1, \ldots, n$) into n slots, so that the β-th slot receives total cycle length λ_α. Each distribution j_α^β is counted with multiplicity

$$\prod_\alpha \frac{i_\alpha!}{\prod_\beta j_\alpha^\beta!} = \frac{\prod_\alpha i_\alpha!}{\prod_\alpha \prod_\beta j_\alpha^\beta!}.$$

So the formula computes the number we want. $\square$

We may use the Vandermonde determinant formula and write

$$\xi_\lambda(C_{\mathbf{i}}) := \left(\Delta(x) \prod_k P_k^{i_k}(x) \right)_{x^{\lambda+\rho}} = \sum_\sigma (-1)^\sigma \left(\prod_k P_k^{i_k}(x) \right)_{x^{\lambda+\rho-\sigma(\rho)}} = \sum_\sigma (-1)^\sigma \chi_{U_{\lambda+\rho-\sigma(\rho)}}(C_{\mathbf{i}}).$$

where we set the summand to zero for σ such that $\lambda + \rho - \sigma(\rho)$ has negative entries. We see two things

- $\xi_\lambda(C_{\mathbf{i}})$ defines a virtual character.
- In the sum only U_μ with $\mu \geq \lambda$ are involved. (Exercise)

Lemma 3.6.

$$\langle \xi_\lambda, \xi_\lambda \rangle = 1.$$

Proof. We have to check that

$$1 = \langle \xi_\lambda, \xi_\lambda \rangle = \frac{1}{n!} \sum_{\substack{\mathbf{i} \\ |\mathbf{i}|=n}} \frac{n!}{\prod_k i_k! k^{i_k}} \left(\Delta(x)\Delta(y) \prod_k P_k^{i_k}(x) P_k^{i_k}(y) \right)_{x^{\lambda+\rho} y^{\lambda+\rho}}.$$

Here we may as well drop the restriction $|\mathbf{i}| = n$ because other terms play no role. Compute

$$\sum_{\mathbf{i}} \frac{1}{\prod_k i_k! k^{i_k}} \prod_k P_k^{i_k}(x) P_k^{i_k}(y) = \prod_k \sum_{i_k} \frac{1}{i_k! k^{i_k}} P_k^{i_k}(x) P_k^{i_k}(y) = \prod_k e^{P_k(x)P_k(y)/k} = e^{\sum_k P_k(x)P_k(y)/k}$$
$$= e^{\sum_k \sum_{\alpha,\beta} x_\alpha^k y_\beta^k / k} = e^{-\sum_{\alpha,\beta} \log(1 - x_\alpha y_\beta)} = \prod_{\alpha,\beta} \frac{1}{1 - x_\alpha y_\beta}.$$

But then, using the Cauchy determinant formula we have to show that

$$1 = \left(\det\left(\frac{1}{1 - x_\alpha y_\beta}\right)_{\alpha\beta} \right)_{x^{\lambda+\rho} y^{\lambda+\rho}} = \left(\sum_{\sigma \in S_n} (-1)^\sigma \prod_j \frac{1}{1 - x_j y_{\sigma(j)}} \right)_{x^{\lambda+\rho} y^{\lambda+\rho}}.$$

Suppose x_1 appars to a power $\lambda_1 + (n-1)$. Then some $y_{\sigma(1)}$ has to appear in the monomial with the same power. But the only possible $\sigma(1)$ the formula would "see" the coefficient of such a monomial is $\sigma(1) = 1$. (All other entries in $\lambda + \rho$ are strictly samller than the first.) Similar one sees that the only contribution comes from the summand $\sigma = \mathbb{1}$, and the contribution is 1. $\square$

Hence the Frobenius Character Theorem follows by an induction and exercise 3.6.

Exercise 3.6. A virtual character is a linear combination

$$\xi = \sum_j n_j \chi_j$$

where $n_j \in \mathbb{Z}$ and the χ_j are characters of representations. Of course, we may assume w.l.o.g. that the χ_j are the characters of the irreducible representations. In this case a virtual charcter is a chracter iff $n_j \geq 0$ for all j. Show that if $(\xi, \xi) = 1$ then ξ or $-\xi$ is an irreducible character.

Proof of Frobenius Character Formula. We already know that ξ_λ is an irreducible character. We will proof by an induction on λ that ξ_λ is actually the character of the irrep V_λ. For λ the trivial partition ($n = n$, it is highest in our ordering) this is easily seen. So suppose we know the result for all $\mu > \lambda$. We saw above that ξ_λ is a linear combination of χ_{U_μ} with $\mu \geq \lambda$, with the coefficient of $\mu = \lambda$ being 1. It follows that one may express χ_{U_λ} as a linear combination of ξ_λ with $\mu \geq \lambda$, and the $\mu = \lambda$ coefficient being 1. So

$$\chi_{U_\lambda} = \xi_\lambda + \sum_{\mu > \lambda} n_\mu \xi_\mu.$$

Furthermore, we know that V_λ is a subrepresentation of U_λ. (So far we have not used the induction hypothesis.)

Now let us decompose U_λ into irreps, and compute the (positive) multiplicty of V_λ

$$1 \leq \langle \chi_\lambda, \chi_{U_\lambda} \rangle = \langle \chi_\lambda, \xi_\lambda \rangle + \sum_{\mu > \lambda} n_\mu \langle \chi_\lambda, \xi_\mu \rangle = \langle \chi_\lambda, \xi_\lambda \rangle.$$

Here we used that $\langle \chi_\lambda, \xi_\mu \rangle$ because $\xi_\mu = \chi_\mu$ by the induction hypothesis (and the first orthogonality Theorem). But since $\xi_\lambda = \chi_{\lambda'}$ for some λ' (and first orthogonality again) we find that $\lambda = \lambda'$ and hence $\xi_\lambda = \chi_\lambda$. $\square$

3.5 The hook lengths formula

In articular the Frobenius character formula allows us to compute the dimension of the representation V_λ.

Theorem 3.3 (Hook lengths formula)**.** *The dimension of the representation V_λ of S_n is*

$$\dim V_\lambda = \frac{n!}{\prod_{\substack{i,j \\ i\leq\lambda_j}} h(i,j)}$$

where $h(i,j)$ is the length of the i,j-hook in the Young diagram λ.

The i,j-hook is the set of boxes i the Young diagram that are to the right or below the box i,j, including box i,j. The hook length is the number of boxes in the hook. For example, here is the 1,2-hook:

$$\begin{array}{|c|c|c|c|c|}
\hline
 & X & X & X & X \\
\hline
 & X & & \\
\cline{1-4}
 & X & & \\
\cline{1-4}
 & X \\
\cline{1-2}
\end{array}$$

The hook length is $h(1,2) = 7$.

Proof. We have to take the $x^{\lambda+\rho}$ coefficient of

$$\Delta(x)(\sum_j x_j)^n.$$

Expanding the determinant Δ this coefficient is

$$\sum_\sigma (-1)^\sigma \frac{n!}{\prod_k (l_k - (n-\sigma(k)))!} = \frac{n!}{\prod_k l_k!} \sum_\sigma (-1)^\sigma \prod_k l_k(l_k-1)\cdots(l_k - n + \sigma(k) + 1)$$

where $l_k = \lambda_k + \rho_k = \lambda_k + n - k$. The term on the right hand side is the determinant

$$\det(l_i(l_i-1)\cdots(l_i - n + j + 1))_{ij}.$$

The i,j-entry of the matrix is a monic polynomial in l_i of degree $n-j$. By elementary column transformations one may put the matrix into the Vandermonde form and compute

$$\det(l_i^{n-j})_{ij} = \prod_{i<j}(l_i - l_j).$$

Next we claim that the expression

$$\frac{\prod_k l_k}{\prod_{i<j}(l_i - l_j)}$$

is just the product over all hook lengths. (This will show the hook length formula.) For example, the factor

$$\frac{l_1!}{\prod_{1<j}(l_1 - l_j)}$$

is the product of the lengths of the hooks in the first row of the Young diagram λ. $\square$

3.6 Induction and restriction

Consider the representation V_λ of S_n, and consider the subgroup $H = S_k \times S_l \subset S_n$. The irreducible representations of H are labelled by pairs of Young diagrams μ, ν, having k and l boxes. We will denote them by

$$V_\mu \boxtimes V_\nu.$$

(exterior tensor product)

The restriction of V_λ decompose into irreducibles

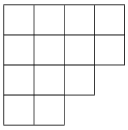
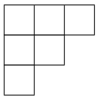
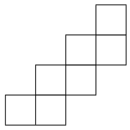
$$Res^{S_n}_{S_k\times S_l} V_\lambda \cong \sum_{\mu,\nu} N^\lambda_{\mu\nu} V_\mu \boxtimes V_\nu$$

for some numbers $N^\lambda_{\mu\nu}$, the *Littlewood-Richardson coefficients.* There is a combinatorial formula, we give without prove.

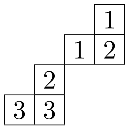
Theorem 3.4. *$N^{\lambda}_{\mu\nu}$ is the number of skew semi-standard tableau of shape λ/μ that contain the number j ν_j times, and satisfy the following additional property. Let S be the sequence of numbers reading the boxes from right to left and top to bottom. Let S_j be the subsequence of the first j characters and n^j_k the number of occurrences of k. Then for each j*

$$n^j_1 \geq n^j_2 \geq n^3_j \geq \cdots .$$

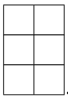
Let us explain the notation. A skew diagram λ/μ is the boxes that remain if we remove μ from λ. For example:

$\lambda =$  $\mu =$  $\lambda/\mu =$ 

A semi-standard tableau is a filling of a Young diagram by numbers $1, 2, \ldots$ (possibly with repetitions) in which each row is non-decreasing and in which each column is increasing. For example



would be allowed. It corresponds to

$\nu =$ .

There is a simpler special case.

Proposition 3.3 (Pieri's rule)**.**

$$Res^{S_n}_{S_{n-1}} V_\lambda \cong \sum_{\mu} V_\mu$$

where the sum is over the set of all diagrams μ that can be obtained by deleting one box from λ.

Exercise 3.7 (Pieri's rule)**.** Show that more generally

$$Res^{S_n}_{S_k} V_\lambda \cong \sum_{\mu} N_\mu V_\mu$$

where N_μ is the number of ways the skew diagram λ/μ can be filled by numbers $1, \ldots, n-k$, such that the numbers in rows and columns are strictly increasing.

3.7 Schur-Weyl duality

We will see later that the irreducible representations of $GL(n)$ are also labelled by Young tableau, so there is one, say W_λ for each tableau. The link (or one link) to the representation theory of the symmetric groups is provided by the following theorem.

Theorem 3.5 (Schur-Weyl duality)**.** *Let U be the defining representation of $GL(n)$ on $\mathbb{C}^n$. Then, as representations of $GL(n) \times S_k$*

$$U^{\otimes k} \cong \oplus_\lambda W_\lambda \boxtimes V_\lambda$$

whre the sum is over Young diagrams with n boxes, and S_n acts on the left by permutation of the factors.

We will discuss this in detail later.

Chapter 4

Lie groups, Lie algebras and their representations

4.1 Overview

Definition 4.1. *A* Lie group *is a group G endowed with the structure of a smooth real manifold, such that the product map $G \times G \to G$ and the inverse map $G \to G$ are smooth.*

A Lie group homomorphism $f : G \to H$ is a group homomorphism that is in addition smooth. A representation of a Lie group G on the finite dimensional vector space V is a Lie group homomorphism $G \to GL(V)$.

In practice (almost) all the Lie groups we are interested in will be closed subgroups of $GL(n)$, determined by algebraic conditions $F_1 = 0, \dots, F_r = 0$. So in particular elements of G are matrices and the product is matrix multiplication.

We will call a closed Lie subgroup of $GL(n)$ a matrix Lie group.

Example 4.1. Some examples are the following.

- $SL(n, \mathbb{K})$ is the group of $n \times n$ matrices of determinant one. (For us $\mathbb{K} = \mathbb{R}, \mathbb{C}$.) They form a group since if A, B have determinant 1, then $\det(AB) = \det A \det B = 1$, and since the inverse of matrix with non-zero determinant exists. Furthermore, the multiplication and inverse maps are clearly smooth.
- The orthogonal group $O(n)$ is the group of real $n \times n$ matrices satisfying $R^T R = \mathbb{1}$. More generally, $O(p,q)$ is the group of matrices R such that $R^T I_{p,q} R = M$ where $I_{p,q} = diag(1, .., 1, -1, .., -1)$. $SO(n)$, $SO(p,q)$ are obtained by demanding in addition that $\det R = 1$.
- Similarly, one obtains the unitary group $U(n)$ of complex matrices satisfying $U^\dagger U = \mathbb{1}$. There are also the indefinite versions $U(p,q)$ of matrices satisfying $U^\dagger I_{p,q} U = I_{p,q}$, and one may demand the determinant to be one to obtain $SU(n)$, $SU(p,q)$.
- The groups of upper and strictly upper triangular matrices are Lie groups.
- $\mathbb{R}$ (with addition) is a Lie group.

Policy: The theory of Lie groups naturally needs some results and notions of differential geometry. In order to not break the promise that the prerequisites for this course are only linear algebra we adopt the following policy:

- We will avoid use of differential geometry where possible.
- If not, I will give proofs for the case of matrix groups, and the students are only expected to know the results in this case. The (usually easy) generalization to general Lie groups will be an exercise for those students familiar with differential geometry.
- We will work with Lie algebras instead of Lie groups whenever possible.

Note that not every Lie group may be embedded into a matrix group.

Exercise 4.1. (For advanced students.) Show that the universal covering group $\widetilde{SL}(2,\mathbb{R})$ of $SL(2,\mathbb{R})$ cannot be realized as a group of matrices.

Solution. Let $\rho : \widetilde{SL}(2,\mathbb{R}) \to GL(V)$ be a faithful representation. We obtain a representation of $\mathfrak{sl}(2,\mathbb{R})$. By complexification we obtain a representation of $\mathfrak{sl}(2,\mathbb{C})$. $SL(2,\mathbb{C})$ is simply connected and hence we obtain a representation of $SL(2,\mathbb{C})$ on $V \otimes \mathbb{C}$. Now $SL(2,\mathbb{R}) \subset SL(2,\mathbb{C})$, hence we obtain another representation of $\widetilde{SL}(2,\mathbb{R})$ via the map $\widetilde{SL}(2,\mathbb{R}) \to SL(2,\mathbb{R})$. It is the complexification of the original representation (check it on the Lie algebra). Hence it is not faithful since it factors through $SL(2,\mathbb{R})$.

Definition 4.2. *A Lie algebra is a vector space* $\mathfrak{g}$ *together with a bilinear map*

$$[,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

such that:

- *For all* $x \in \mathfrak{g} : [x,x] = 0$.
- *For all* $x, y, z \in \mathfrak{g}$ *the* Jacobi identity

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

  holds.

Let $\mathfrak{g}$, $\mathfrak{g}'$ *be Lie algebras. A Lie algebras homomorphism* $f : \mathfrak{g} \to \mathfrak{g}'$ *is a linear map* $f : \mathfrak{g} \to \mathfrak{g}'$ *that commutes with the Lie brackets, i. e.,*

$$f([x,y]) = [f(x), f(y)].$$

A Lie algebra isomorphism is an invertible homomorphism.

Remark 4.1. The first condition implies that

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

and hence the bracket is antisymmetric. Vice versa, antisymmetry implies that $[x,x] = 0$ in characteristic > 2. We will only consider complex or real Lie algebras, i. e., the underlying vector space is over $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

Example 4.2. Let V be a vector space. Then the endomorphisms $\mathfrak{gl}(V) := End(V)$ form a Lie algebra with the Lie bracket being the commutator.

Definition 4.3. *A representation of the Lie algebra* $\mathfrak{g}$ *on the vector space* V *is a Lie algebra homomorphism* $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$.

Example 4.3.

- Let $\mathfrak{g}$ be a Lie algebra and V be any vector space. Then the trivial representation of $\mathfrak{g}$ on V is the zero map $\mathfrak{g} \to \mathfrak{gl}(V)$.
- The adjoint representation of $\mathfrak{g}$ on itself ($V = \mathfrak{g}$) is defined such that

$$\rho(x) \mapsto \mathrm{ad}_x := (y \in \mathfrak{g} \mapsto [x,y]).$$

  To see that it is a representation, check that

$$[\mathrm{ad}_x, \mathrm{ad}_y](z) = [x,[y,z]] - [y,[x,z]] = [[x,y],z] = \mathrm{ad}_{[x,y]}(z).$$

The key point is that to any Lie group G we may associate a Lie algebra $\mathfrak{g}$, and to a large extend the representation theory and structure of G is determined by that of $\mathfrak{g}$. Concretely, $\mathfrak{g} = T_{\mathbb{1}}G$ is the tangent space at the identity of G. The Lie brackt is constructed as follows. Note that if $\phi : G \to G$ is an autmorphism of Lie groups, then $D\phi_{\mathbb{1}}$ is an automorphism of the tangent spaces. For any $g \in G$, the adjoint action $Ad_g : G \to G$, $h \mapsto ghg^{-1}$ defines an automorphism. The corresponding G-action on the tangent space is also denoted by $Ad_g : \mathfrak{g} \to \mathfrak{g}$ by abuse of notation. In particular it follows that for a curve $\gamma(t)$ passing through $\mathbb{1}$ at $t = 0$, and for $y \in \mathfrak{g}$, each element of

$$Ad_{\gamma(t)} y$$

is in $\mathfrak{g}$. The derivative at $t = 0$ is by definition the Lie bracket of $x = \dot{\gamma}(t = 0) \in \mathfrak{g}$ and y.

Proposition 4.1. *Let G be a Lie group. Then the tangent space at the identity $\mathfrak{g} = T_{\mathbb{1}}G$ is a Lie algebra with bracket*

$$[x, y] = \frac{d}{dt}\bigg|_{t=0} Ad_{\gamma(t)} y$$

where $x, y \in \mathfrak{g}$, and γ is a curve representing the tangent vectors x, y.

If $F : G \to H$ is a Lie group homomorphism than the derivative at the identity defines a Lie algebra homomorphism between the Lie algebras of G and H, $f = D_{\mathbb{1}}F : \mathfrak{g} \to \mathfrak{h}$.

For matrix groups the proof is essentially contained in the following example.

Example 4.4. If G is a subgroup of $GL(n)$ defined by relations F_j, then $\mathfrak{g} \subset \mathfrak{gl}(n)$ is the Lie algebra of matrices h that satisfy

$$D_{\mathbb{1}}F_j \cdot h = 0$$

for all j. The Lie bracket is the commutator of matrices, as can be seen as follows. Suppose γ is a curve on $GL(n)$ such that $\gamma(0) = \mathbb{1}$ and $\dot\gamma(0) = x$. Then taking the derivative of $\mathbb{1} \equiv \gamma(t)\gamma(t)^{-1}$ at $t = 0$ we find that

$$0 = x + \frac{d\gamma(t)^{-1}}{dt}\bigg|_{t=0}$$

and hence $\frac{d\gamma(t)^{-1}}{dt}\Big|_{t=0} = -x$. Thus

$$[x, y] := xy - yx.$$

This clearly is antisymmetric. It satisfies the Jacobi identity since

$$[x, [y, z]] + (cyclic) = xyz - xzy - yzx + zyx + (cyclic) = 0.$$

Hence one obtains a Lie algebra.

To see the second statement in the proposition, note that $F(Ad_g h) = F(ghg^{-1}) = F(g)F(h)F(g)^{-1} = Ad_{F(g)}F(h)$. Hence, taking the derivative in h at $\mathbb{1}$, $f(Ad_g y) = Ad_F(g)f(y)$. Setting $g = \gamma(t)$, $\dot\gamma(0) =: x$ and taking the derivative at $t = 0$ one obtains the result:

$$\begin{aligned} f([x, y]) &= \frac{d}{dt} f(Ad_{\gamma(t)}y)\bigg|_{t=0} = \frac{d}{dt} Ad_{F(\gamma(t))} f(y)\bigg|_{t=0} \\ &= \frac{d}{dt} F(\gamma(t))\bigg|_{t=0} f(y) + f(y)\,\frac{d}{dt} F(\gamma(t))^{-1}\bigg|_{t=0} = f(x)f(y) - f(y)f(x). \end{aligned}$$

Exercise 4.2. Show the above proposition in the general case, i. e., without assuming that G is a group of matrices.

In particular, any representation of G determines a representation of $\mathfrak{g}$. The converse is not always true. As we shall see later, however, “not too false” either.

There is a map $\exp : \mathfrak{g} \to G$, the exponential map. For a matrix group this is just the exponential of matrices.

Finally, we note (without proof) that Lie algebras (-in contrast to Lie groups-) may always assumed to be matrix algebras.

Theorem 4.1 (Ado's Theorem)**.** *Let $\mathfrak{g}$ be a finite dimensional Lie algebra. Then there is an n such that $\mathfrak{g}$ is isomorphic to a Lie subalgebra of $\mathfrak{gl}(n)$.*

Proof. For a proof, see []. □

The treatment in this chapter will mostly follow [**?**]. (Knapp)

4.2 General definitions and facts about Lie algebras

Definition 4.4. *A Lie algebra $\mathfrak{g}$ is abelian if its Lie bracket vanishes identically.*

A Lie algebra $\mathfrak{g}$ is nilpotent if for some large enough n

$$\mathfrak{g}_n := \underbrace{[\mathfrak{g}, [\mathfrak{g}, [\cdots \mathfrak{g}]]]}_{n\times} = 0.$$

The series $\mathfrak{g} = \mathfrak{g}_1 \supset \mathfrak{g}_2 \supset \mathfrak{g}_3 \supset \cdots$ is called the lower central series.

A Lie algebra $\mathfrak{g}$ is solvable if for some large enough n, $\mathfrak{g}^{(n)} = 0$ where $\mathfrak{g}^{(0)} := \mathfrak{g}$ and recursively $\mathfrak{g}^{(n)} = [\mathfrak{g}^{(n-1)}, \mathfrak{g}^{(n-1)}]$. The series $\mathfrak{g}^{(0)} \supset \mathfrak{g}^{(1)} \supset \cdots$ is called the derived series.

In particular, any abelian Lie algebra is nilpotent, and any nilpotent Lie algebra is solvable.

Example 4.5. The standard examples are as follows: The Lie algebra of diagonal matrices is abelian, the Lie algebra of strictly upper triangular matrices is nilpotent and the Lie algebra of upper triangular matrices is solvable. In fact, we will see below that any solvable Lie algebra can be realized as a Lie subalgebra of the Lie algebra of upper triangular matrices. Similarly, one can show that any nilpotent Lie algebra can be realized as a subalgebra of the strictly upper triangular matrices.

Definition 4.5. *A Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$ is a linear subspace closed under the Lie bracket, i.e., $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$. An ideal $\mathfrak{l} \subset \mathfrak{g}$ is a linear subspace such that for all $l \in \mathfrak{l}, x \in \mathfrak{g}$, $[l, x] \in \mathfrak{l}$.*

In the latter case the Lie bracket naturally descends to a Lie bracket on the quotient space $\mathfrak{g}/\mathfrak{l}$ by defining

$$[x + \mathfrak{l}, y + \mathfrak{l}] = [x, y] + \mathfrak{l}.$$

One calls $\mathfrak{g}/\mathfrak{l}$ the quotient Lie algebra. The projection $\mathfrak{g} \to \mathfrak{g}/\mathfrak{l}$ is clearly a Lie algebra homomorphism.

Exercise 4.3. Check that if $\mathfrak{l}_1, \mathfrak{l}_2 \subset \mathfrak{g}$ are ideals, then the sum $\mathfrak{l}_1 + \mathfrak{l}_2$ and intersection $\mathfrak{l}_1 \cap \mathfrak{l}_2$ are ideals as well. Furthermore, if $f : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism and $\mathfrak{l} \subset \mathfrak{h}$ is an ideal, then $f^{-1}(\mathfrak{l}) \subset \mathfrak{g}$ is an ideal as well.

Finally, if $\mathfrak{l} \subset \mathfrak{g}$ is an ideal and $\mathfrak{h} \subset \mathfrak{g}$ is subalgebra containing $\mathfrak{l}$, then show that $\mathfrak{l} \subset \mathfrak{h}$ is an ideal.

The notion of representation of Lie algebras was defined above. Consider two relevant generic examples: (i) the trivial representation $x \to 0$ on any vector space V. (ii) The adjoint representation of $\mathfrak{g}$ on $\mathfrak{g}$, given by $x \mapsto ad_x$, where $\mathrm{ad}_x : y \mapsto [x, y]$.

One cannot not give a general classification of Lie algebras, or their representations, not even in finite dimensions. However, there is a partial result.

Proposition 4.2. *Every finite dimensional Lie algebra $\mathfrak{g}$ has a unique maximal solvable ideal $\mathfrak{r} \subset \mathfrak{g}$, called the radical of $\mathfrak{g}$. The quotient $\mathfrak{g}/\mathfrak{r}$ has no non-trivial (i.e., non-zero) solvable ideals.*

Proof. The sum of all solvable ideals is a solvable ideal (exercise), which is automatically maximal. Suppose $\mathfrak{g}/\mathfrak{r}$ has a solvable ideal $\mathfrak{r}'$. Then $f^{-1}(\mathfrak{r}') \subset \mathfrak{g}$ is a solvable ideal, where $f : \mathfrak{g} \to \mathfrak{g}/\mathfrak{r}$ is the projection (exercise). Hence it is contained in $\mathfrak{r}$, hence $\mathfrak{r}' = 0$. □

Definition 4.6. *A Lie algebra $\mathfrak{g}$ is called simple if it does not have a non-trivial (i. e., not 0 of $\mathfrak{g}$) ideal. A Lie algebra $\mathfrak{g}$ is semi-simple if it has no non-zero solvable ideals.*

Hence to any finite dimensional Lie algebra we may associate a solvable and a semisimple Lie algebra. The good news is that semi-simple Lie algebras can be completely classified, and one has a complete understanding of their finite dimensional representations. The bad news is that there is currently for solvable (or nilpotent) Lie algebras.

4.3 The theorems of Lie and Engel

In the following, we will focus mostly on semi-simple Lie algebras. However, let us state the two most important Theorems about nilpotent and solvable Lie algebras.

Theorem 4.2 (Lie's Theorem)**.** *Let $\mathfrak{g}$ be a solvable Lie algebra, and let $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation, where V is a* complex *vector space. Then there is a simultaneous eigenvector $v \neq 0 \in V$ of all $\rho(x)$, for $x \in \mathfrak{g}$.*

Sketch of proof. One performs an induction on $\dim \mathfrak{g}$, the case $\dim \mathfrak{g} = 1$ being trivial. Otherwise one may pick a codimension 1 ideal $\mathfrak{h} \subset \mathfrak{g}$. (It exists since $[\mathfrak{g}, \mathfrak{g}] \neq \mathfrak{g}$, so taking for $\mathfrak{h}$ any codimension 1 subspace containing $[\mathfrak{g}, \mathfrak{g}]$ will do.) Pick a vector X in the complement. Let e be a joint eigenvector of $\mathfrak{h}$, i.e., $\rho(y)e = \lambda(y)e$ for all $y \in \mathfrak{h}$. Define $e_j = \rho(X)^j e$ and let W be the span of these vectors.

Claim A: $\rho(y)e_j = \lambda(y)e_j$ for all $y \in \mathfrak{h}$. Assuming the claim, the Theorem easily follows, since we may pick for v any eigenvector of $\rho(X)\mid_W$.

Claim B: $\rho(y)e_j = \lambda(y)e_j + span(e_0, \ldots, e_{n-1})$ for all $y \in \mathfrak{h}$. Proof of Claim B by induction on j:

$$\begin{aligned} \rho(y)e_{j+1} &= \rho(y)\rho(X)e_j \\ &= \rho([y, X])e_j + \rho(X)\rho(y)e_j \\ &= \lambda([y, X])e_j + \lambda(y)e_{j+1} + \cdots . \end{aligned}$$

Note that the same proof shows Claim A, if we can show that $\lambda([y, X]) = 0$. But $\lambda(z) = \frac{1}{\dim W} tr\rho(z)$ by Claim B. Since the trace of a commutator vanishes, we have shown Claim A.

□

Corollary 4.1 (Lie's Theorem)**.** *Under the assumptions above, there is a basis of V such that all matrices $\rho(x)$, for $x \in \mathfrak{g}$, are upper triangular in this basis.*

Proof. We do an induction on $n = \dim V$. For $\dim V = 1$ the statement is obvious. Otherwise, let $e_n := v \in V$ be a simultaneous eigenvector as in the Theorem. Let U be the subspace spanned by v. There is a natural representation ρ' of $\mathfrak{g}$ on V/U. Using the induction hypothesis, pick a basis $e'_1, \dots, e'_{n-1}$ of V/U, such that all representation matrices are upper triangular in this basis. Let $e_1, \dots, e_{n-1}$ be some pre-images in V. Then we claim that in the basis $e_1, \dots, e_n$ all matrices $\rho(x)$ are upper triangular. Indeed $\rho(x)e_j$ lift $\rho'(x)e'_j$ and hence $\rho(x)e_j = \sum_{i\geq j} c_i e_i$ for some constants c_i. □

The second important Theorem is Engel's Theorem, which provides an alternative definition of nilpotence. An element $x \in \mathrm{End}(V)$, V a vector space, is called nilpotent if $x^n = 0$ for some n.

Theorem 4.3 (Engel's Theorem)**.** *Let $V \neq 0$ be a vector space and $\mathfrak{g} \subset \mathfrak{gl}(V)$ a Lie algebra, all of whose members are nilpotent endomorphisms. Then there is a basis of V such that all $x \in \mathfrak{g}$ have strictly upper triangular matrices in this basis. In particular, $\mathfrak{g}$ is nilpotent.*

Corollary 4.2 (Engel's Theorem, nilpotence test)**.** *If $\mathfrak{g}$ is a Lie algebra such that ad_x is a nilpotent endomorphism for each $x \in \mathfrak{g}$, then $\mathfrak{g}$ is nilpotent.*

Proof. Denote by $\mathrm{ad}\mathfrak{g} \subset \mathfrak{gl}(V)$ the image of ad. By the Theorem $\mathrm{ad}\mathfrak{g}$ is nilpotent. But note that $\mathrm{ad}\mathfrak{g} \cong \mathfrak{g}/\mathfrak{c}$, where $\mathfrak{c} \subset \mathfrak{g}$ is the center of $\mathfrak{g}$, i.e., the kernel of ad. By nilpotence of this latter Lie algebra may be rephrased as saying that for some n and all $x_1, \dots, x_n \in \mathfrak{g}$

$$[x_1, [x_2, [\cdots [x_{n-1}, x_n] \cdots] \in \mathfrak{c}.$$

But then for all $x_1, \dots, x_{n+1} \in \mathfrak{g}$

$$[x_1, [x_2, [\cdots [x_n, x_{n+1}] \cdots] = 0$$

and hence $\mathfrak{g}$ is nilpotent. □

4.4 The Killing form and Cartan's criteria

On any (finite dimensional) Lie algebra we may define the following symmetric blinear form

$$K(x, y) := tr(ad_x ad_y)$$

the *Killing form.* It is invarant, i.e., for each $x, y, z \in \mathfrak{g}$

$$K([x,z], y) = tr(ad_{[x,z]} ad_y) = tr(ad_x ad_z ad_y) - tr(ad_z ad_x ad_y) = tr(ad_x ad_z ad_y) - tr(ad_x ad_y ad_z) = K(x, [z, y]).$$

The Killing form may be used to test for solvability or semi-simplicity.

Theorem 4.4 (Cartan's criterion for solvability)**.** *A finite dimensional complex or real Lie algebra $\mathfrak{g}$ is solvable iff for all $X \in \mathfrak{g}$ and $Y \in [\mathfrak{g}, \mathfrak{g}]$: $K(X, Y) = 0$, where K is the Killing form.*

The proof is developed in Serie 8.

Theorem 4.5 (Cartan's criterion for semi-simplicity)**.** *A real or complex Lie algebra is semi-simple iff its Killing form is non-degenerate.*

Proof. We denote by $\mathrm{rad}K$ the set of vectors annihilated by the Killing form, i.e.,

$$\mathrm{rad}K = \{x \in \mathfrak{g} \mid K(x, y) = 0 \forall y \in \mathfrak{g}\}.$$

It is an ideal since if $x \in \mathrm{rad}K$, $y \in \mathfrak{g}$

$$K([x, y], \cdot) = K(x, [y, \cdot]) = 0.$$

We claim that rad$K$ is solvable. We use Cartan's criterion to test this. Let K' be the killing form of rad$K$. Note that for $x \in \mathrm{rad}K$
$$\mathrm{ad}_x = \begin{pmatrix} * & * \\ 0 & 0 \end{pmatrix}.$$
Hence for $x, y \in \mathrm{rad}K$
$$K'(x,y) = tr_{\mathrm{rad}K}(\mathrm{ad}_x \mathrm{ad}_y \mid_{\mathrm{rad}K}) = tr_{\mathfrak{g}}(\mathrm{ad}_x \mathrm{ad}_y) = K(x,y) = 0.$$
Hence rad$K$ is indeed solvable by the Theorem above. Hence if rad$K \neq 0$, then $\mathfrak{g}$ is not semisimple.

Conversely, suppose $\mathfrak{g}$ is not semi-simple, with (non-zero) radical $\mathfrak{r} \subset \mathfrak{g}$. Since $\mathfrak{r}$ is solvable it has an abelian ideal $\mathfrak{a}$. This is also an ideal in $\mathfrak{g}$. The matrices of $\mathrm{ad}_a, \mathrm{ad}_a$, $x \in \mathfrak{g}$, $a \in \mathfrak{a}$ have the form
$$\mathrm{ad}_a = \begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix} \qquad \mathrm{ad}_x = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix}.$$
Thus $\mathrm{ad}_a \mathrm{ad}_x$ is strictly upper triangular and hence $K(a,x) = tr(\mathrm{ad}_a \mathrm{ad}_x) = 0$. □

Remark 4.2. Note that in the proof we also saw that $\mathfrak{g}$ is semisimple iff it does not have an abelian ideal ($\neq 0$).

For $\mathfrak{g}_1, \mathfrak{g}_2$ Lie algebras, their direct sum $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ is the Lie algebra with underlying vector space $\mathfrak{g}_1 \oplus \mathfrak{g}_2$, with bracket such that $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are ideals and commute.

Corollary 4.3. *A finite dimensional Lie algebra $\mathfrak{g}$ is semi-simple iff it is isomorphic to a direct sum of simple Lie algebras.*

Hence in order to classify all semisimple Lie algebras it suffices to classify all simple Lie algebras.

Proof. $\Leftarrow$: Exercise.

$\Rightarrow$: Let $\mathfrak{h} \subset \mathfrak{g}$ be a minimal non-zero ideal. (If no non-trivial ideal exists, we are done.) Let $\mathfrak{h}^\perp$ be the orthogonal complement with respect to the Killing form. Check that $\mathfrak{h}^\perp$ is an ideal, for $x \in \mathfrak{g}, y \in \mathfrak{h}^\perp, z \in \mathfrak{h}$:
$$K([x,y],z) = -K(y,[x,z]) = 0$$
since $[x,z] \in \mathfrak{h}$. Since the intersection of ideals is an ideal, $\mathfrak{h} \cap \mathfrak{h}^\perp$ is an ideal, hence either 0 or $\mathfrak{h}$ by minimality of $\mathfrak{h}$. If it is h, then (as we saw above) the Killing form of $\mathfrak{h}$ vanishes and $\mathfrak{h}$ is solvable, a contradiction. Hence $\mathfrak{h} \cap \mathfrak{h}^\perp = 0$, and we can write
$$\mathfrak{g} \cong \mathfrak{h} \oplus \mathfrak{h}^\perp.$$
$\mathfrak{h}$ is simple, since it cannot contain a non-trivial ideal by minimality. The Killing form of $\mathfrak{h}^\perp$ is the restriction of the Killing form on $\mathfrak{g}$ (proof as above). Hence it must be nondegenerate, for otherwise there were a non-zero $y \in \mathfrak{h}^\perp$ such that $y \perp \mathfrak{h}^\perp$, hence $y \in \mathrm{rad}K$, and hence $\mathfrak{g}$ not semi-simple by the Theorem. □

Example 4.6. Let us compute the Killing form of $\mathfrak{gl}_n$ or $\mathfrak{sl}_n$. For a matrix x, the adjoint action is described by the matrix $x \otimes \mathbb{1} - \mathbb{1} \otimes x$. Hence we compute
$$tr(ad_x ad_y) = tr(xy \otimes \mathbb{1} - x \otimes y - y \otimes x + \mathbb{1} \otimes xy) = 2n tr(xy) - 2tr(x)tr(y).$$
This form is degenerate, since it vanishes for $x = \mathbb{1}$.

If we restrict to $\mathfrak{sl}_\mathfrak{n}$, the Killing form remains the same, but we may simplify by using $tr(x) = tr(y) = 0$. Then the Killing form becomes $2ntr(xy)$, and is non-degenerate. Hence $\mathfrak{sl}_n$ is semi-simple. (In fact, it is simple.)

Example 4.7. Let us compute the Killing form on $\mathfrak{so}_3 \cong \mathbb{R}^3$. The bracket is the wedge product. Hence
$$[x,[y,z]] = y(x \cdot z) - z(x \cdot y).$$
Hence F
$$tr(ad_x ad_y) = tr(yx^T - (x \cdot y)\mathbb{1}) = -2(x \cdot y).$$
This is nondegenerate and hence $\mathfrak{so}_3$ is semi-simple.

4.5 Classification of complex simple Lie algebras

Theorem 4.6. *Any complex simple Lie algebra is isomorphic to one in the following list:*

- $\mathfrak{sl}(n+1,\mathbb{C}) = \{x \in \mathrm{Mat}_{n+1\times n+1}(\mathbb{C}) \mid trx = 0\}$. *($A_n$, special linear Lie algebra)*
- $\mathfrak{so}(2n+1,\mathbb{C}) = \{x \in \mathrm{Mat}_{2n+1\times 2n+1}(\mathbb{C}) \mid x + x^T = 0\}$. *($B_n$, special orthogonal Lie algebra)*
- $\mathfrak{sp}(2n,\mathbb{C}) = \{x \in \mathrm{Mat}_{2n\times 2n}(\mathbb{C}) \mid x^T\Omega + \Omega x = 0\}$, *where* $\Omega = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}$. *($C_n$, symplectic Lie algebra)*
- $\mathfrak{so}(2n,\mathbb{C}) = \{x \in \mathrm{Mat}_{2n\times 2n}(\mathbb{C}) \mid x + x^T = 0\}$. *($D_n$, special orthogonal Lie algebra)*
- $\mathfrak{e}_6, \mathfrak{e}_7, \mathfrak{e}_8$ *of dimensions 78, 133, 248.*
- $\mathfrak{f}_4$ *of dimension 52 are the (complexifications of the) derivations of the algebra* $\mathfrak{h}_3(\mathbb{O})$ *of Hermitian* 3×3 *matrices with octonionic entries, with product* $A \circ B = \frac{1}{2}(AB + BA)$.
- $\mathfrak{g}_2$ *of dimenson 14 are the (complexifications of the) derivations of the octonion (non-associative) algebra.*

Notation: We write $\mathfrak{so}(n) = \mathfrak{so}(n,\mathbb{R})$, $\mathfrak{sl}(n) = \mathfrak{sl}(n,\mathbb{R})$ etc. for the real versions of the above Lie algebras.

The first four families are called the classical simple lie algebras. The remainder are called the exceptional Lie algebras. They can be defined using the Freudenthal magic square construction. We give a version due to Vinberg. Also, we only construct the real versions, the Lie algebras above are obtained by tensoring with $\mathbb{C}$.

In any case one needs the non-associative algebra of octonions. The octonions $\mathbb{O}$ are the non-associative $\mathbb{R}$-algebra of pairs (a,b), where $a,b \in \mathbb{H}$ are quaternions, with product

$$(a,b)(c,d) = (ab - \bar{d}b, da + b\bar{c}).$$

The unit is $1 = (1,0)$. The conjugation is defined for $u = (a,b) \in \mathbb{O}$ as

$$\bar{u} := (\bar{a}, -b).$$

The real part is $\Re u := \frac{1}{2}(u + \bar{u}) = \Re(a)$. There is a norm $|u| = \sqrt{u\bar{u}}$ satisfying $|uv| = |u||v|$. It comes from the inner product $\langle u,v\rangle = \Re u\bar{v}$. $\mathbb{O}$ has the structure of a non-associative division algebra, i. e., every non-zero element is invertible ($u^{-1} = \bar{u}/|u|^2$).

Given suitable (possibly non-associative) division algebras $\mathbb{K}, \mathbb{K}'$ we define a Lie algebra

$$M(\mathbb{K},\mathbb{K}') = \mathfrak{der}(\mathbb{K}) \oplus \mathfrak{der}(\mathbb{K}) \oplus \mathfrak{sa}_3(\mathbb{K}\otimes\mathbb{K})$$

where the first two summands are commuting subalgebras, that act on the last summand by acting on the matrix entries. The last summand are anti-Hermitian traceless matrices with bracket

$$[A,B] = (AB - BA)_0 + \frac{1}{3}\sum_{i,j=1}^{3}(D_{A_{ij},B_{ij}})$$

where

$$D_{a\otimes a', b\otimes b'} = \langle a,b\rangle\, d_{a',b'} + \langle a',b'\rangle\, d_{a,b}$$

where $d_{a,b}$ is the derivation of $\mathbb{K}$ defined by

$$d_{a,b}x = [[a,b],x] - 3((ab)x - a(bx)).$$

Exercise 4.4. Check that $d_{a,b}$ is indeed a derivation and that $M(\mathbb{K},\mathbb{K}')$ is a Lie algebra.

We obtain the following table (the magic square)

	$\mathbb{R}$	$\mathbb{C}$	$\mathbb{H}$	$\mathbb{O}$
$\mathbb{R}$	$\mathfrak{so}(3)$	$\mathfrak{su}(3)$	$\mathfrak{sp}(3)$	$\mathfrak{f}_4$
$\mathbb{C}$	$\mathfrak{su}(3)$	$\mathfrak{su}(3)\oplus\mathfrak{su}(3)$	$\mathfrak{so}(12)$	$\mathfrak{e}_6$
$\mathbb{H}$	$\mathfrak{sp}(3)$	$\mathfrak{su}(6)$	$\mathfrak{so}(12)$	$\mathfrak{e}_7$
$\mathbb{O}$	$\mathfrak{f}_4$	$\mathfrak{e}_6$	$\mathfrak{e}_7$	$\mathfrak{e}_8$

Exercise 4.5. Check that the dimensions are indeed the ones provided.

4.6 Classification of real simple Lie algebras

For $\mathfrak{g}$ a real simple Lie algebra, $\mathfrak{g}_{\mathbb{C}} := \mathfrak{g} \otimes \mathbb{C}$ is a complex Lie algebra, but not necessarily simple. However, one can show that either $\mathfrak{g}$ is already complex, or not, in which case $\mathfrak{g}_{\mathbb{C}}$ is simple. Hence the classification of real simple Lie algebras will refine that of complex simple Lie algebras. For $\mathfrak{g}$ complex, we will denote by $\mathfrak{g}_{\mathbb{R}}$ the same Lie algebra, but considered as real Lie algebra.

Theorem 4.7. *Any real simple Lie algebra is isomorphic to one in the following list:*

- $\mathfrak{g}_{\mathbb{R}}$*, for* $\mathfrak{g}$ *a complex simple Lie algebra from the list above.*
- $\mathfrak{so}(p,q) = \{x \in \mathrm{Mat}_{p+q\times p+q}(\mathbb{R}) \mid x^T I_{p,q} + I_{p,q} x = 0\}$*, where* $I_{p,q}$ *is a diagonal matrix with* p 1*s and* q -1*s on the diagonal. In particular,* $\mathfrak{so}_n = \mathfrak{so}_{n,0}$*.*
- $\mathfrak{su}(p,q) = \{x \in \mathrm{Mat}_{p+q\times p+q}(\mathbb{C}) \mid x^\dagger I_{p,q} + I_{p,q} x = 0, trx = 0\}$*. In particular,* $\mathfrak{su}(n) = \mathfrak{su}(n,0)$*.*
- $\mathfrak{sp}(p,q) = \{x \in \mathrm{Mat}_{p+q\times p+q}(\mathbb{H}) \mid x^* I_{p,q} + I_{p,q} x = 0\}$*. In particular* $\mathfrak{sp}(n) = \mathfrak{sp}(n,0)$*.*
- $\mathfrak{sp}(n,\mathbb{R}) = \{x \in \mathrm{Mat}_{2n\times 2n}(\mathbb{R}) \mid x^T \Omega + \Omega x = 0\}$*.*
- $\mathfrak{sl}(n) = \mathfrak{sl}(n,\mathbb{R}) = \{x \in \mathrm{Mat}_{n\times n}(\mathbb{R}) \mid trx = 0\}$*.*
- $\mathfrak{sl}(n,\mathbb{H}) = \{x \in \mathrm{Mat}_{n\times n}(\mathbb{H}) \mid trx = 0\}$*.*
- $\mathfrak{so}^*(2n) = \{x \in \mathfrak{su}(n,n) \mid x^T J_n + J_n x = 0\}$*, where* $J = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$*.*
- *17 real Lie algebras whose complexifications are one of the exceptional complex Lie algebras. (There are 5 for* $\mathfrak{e}_6$*, 4 for* $\mathfrak{e}_7$*,* 3 *for* $\mathfrak{e}_8$*, 3 for* $\mathfrak{f}_4$ *and* 2 *for* $\mathfrak{g}_2$*.)*

4.7 Generalities on representations of Lie algebras

Definition 4.7. *Let* $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$*,* $\rho' : \mathfrak{g} \to \mathfrak{gl}(V')$ *be representations. A morphism betwenn these representations (or intertwiner) is a linear map* $f : V \to V'$ *such that for all* $x \in \mathfrak{g}$

$$f \circ \rho(x) = \rho'(x) \circ f.$$

f *is an isomorphism if it is an isomorphism of the underlying vector spaces.*

Definition 4.8. *Let* ρ, ρ' *be representations of* $\mathfrak{g}$ *on the bvector spaces* V, V'*.*

- *The direct sum of the two representation is the representation* $\rho_\oplus$ *on* $V \oplus V'$ *such that*

$$\rho_\oplus(x)(v + v') = \rho(x)v + \rho'(x)v'.$$

- *The tensor product representation is the representation* $\rho_\otimes$ *on* $V \otimes V'$ *such that*

$$\rho_\oplus(x)(v \otimes v') = \rho(x)v \otimes v' + v \otimes \rho'(x)v'.$$

- *The linear maps* $\mathrm{Hom}(V, V')$ *can be endowed with the structure of a representation of* $\mathfrak{g}$ *such that*

$$\rho_{\mathrm{Hom}}(f) = \rho'(x)f + f\rho(x).$$

Exercise 4.6. Check that the formulas above indeed define representations.

Definition 4.9. *Let* $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ *be a representation of a Lie algebra* $\mathfrak{g}$ *on a vector space* V*. Let* $W \subset V$ *be a subspace. We will say that* W *is* $\mathfrak{g}$*-invariant if* $\rho(x)W \subset W$ *for all* $x \in \mathfrak{g}$*. We will say that* ρ *is irreducible if it does not have an invariant subspace other than* $W = 0$*,* $W = V$*.*

We will say that ρ *is completely reducible if it is isomorphic to a direct sum of irreducible representations, i.e.,*

$$V \cong \bigoplus_{j \in J} V_j$$

with each V_j *irreducible.*

Not every representation of every Lie algebra is completely reducible. However, those of the semisimple Lie algebras are, as we shall see later.

Exercise 4.7. Show that if V is a representation of a Lie algebra $\mathfrak{g}$ and $U \subset V$ is an invariant subspace, then there is a natural representation on V/U such that the projection $V \to V/U$ is an intertwiner.

Suppose V, V' are representations, $f : V \to V'$ is an intertwiner, and $U \subset V'$ is an invariant subspace. Show that $f^{-1}(U) \subset V$ is an invariant subspace.

There is also a version of Schur's lemma, with the same proof.

Lemma 4.1 (Schur's Lemma)**.** *Let $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$, $\rho' : \mathfrak{g} \to \mathfrak{gl}(V')$ be irreducible representations and $f : V \to V'$ an intertwiner. Then either $f \equiv 0$ or f is invertible.*

If the base field $\mathbb{K} = \mathbb{C}$ (or algebraically closed) and $V' = V$ then in addition $f = \lambda\mathbb{1}$ for some $\lambda \in \mathbb{C}$.

Proof. First statement: The image and kernel of f are invariant subspaces.

Second statement: Let λ be an eigenvalue of f. Then $f - \lambda\mathbb{1}$ is not invertible, hence 0 by the first statement. □

Definition 4.10. *Let $\mathfrak{g}$ be a real Lie algebra and V be a real (or complex) vector space with inner product, carrying a representation ρ of $\mathfrak{g}$. Then we say that the representation is orthogonal (or unitary) if the operators $\rho(x)$ are self-adjoint for all $x \in \mathfrak{g}$.*

Unfortunately, in the case of Lie algebras not every representation may be made unitary in general. In the following we will only discuss real or complex Lie algebras, with emphasis on the complex case.

4.8 Representation theory of $\mathfrak{sl}(2,\mathbb{C})$

In order to understand the representation theory of arbitrary semi-simple Lie algebras, one first needs to understand the representation theory of $\mathfrak{sl}(2,\mathbb{C})$ well.

Recall that $\mathfrak{sl}(2,\mathbb{C}) = \{x \in \mathrm{Mat}_{2\times 2}(\mathbb{C}) \mid trx = 0\}$. We will often use the following basis

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

One easily checks that

$$[h,e] = 2e \qquad [h,f] = -2f \qquad [e,f] = h.$$

Let $n \geq 0$ be an integer. We will construct the following $n+1$-dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$. Let $\{e_0, \ldots, e_n\}$ be some basis vectors and set

$$\begin{aligned} &\rho(e)e_0 = 0 \\ &\rho(f)e_j = e_{j+1} \quad \text{where } e_{n+1} := 0 \\ &\rho(h)e_j = (n-2j)e_j \\ &\rho(e)e_j = j(n-j+1)e_{j-1}. \end{aligned}$$

We claim that these formulas define a representation. We need to check the three relations above.

$$\begin{aligned} &[\rho(h),\rho(f)]e_j = (n-2(j+1)-(n-2j))e_{j+1} = -2\rho(f)e_j \\ &[\rho(h),\rho(f)]e_n = 0 - 0 = 0 \\ &[\rho(h),\rho(e)]e_j = j(n-j+1)(n-2(j-1)-(n-2j))e_{j-1} = 2\rho(e)e_j \\ &[\rho(e),\rho(f)]e_j = (j+1)(n-j)e_j - j(n-j+1)e_j = (n-2j)e_j = \rho(h)e_j \\ &[\rho(e),\rho(f)]e_n = 0 - n(n-n+1)e_n = (n-2n)e_n = \rho(h)e_j. \end{aligned}$$

We call the representation above ρ_n, with underlying vector space V_n.

Exercise 4.8. We can endow V_n with an inner product such that

$$\langle e_i, e_j \rangle = \delta_{ij} \frac{j!}{(n-j)!}.$$

Show that $\rho(h)^\dagger = \rho(h)$ and $\rho(e)^\dagger = \rho(f)$. Show that the representation of $\mathfrak{su}_2 \subset \mathfrak{sl}(2,\mathbb{C})^{\mathbb{R}}$ obtained by restriction is unitary.

Theorem 4.8. *For each integer $n \geq 1$ the representation V_n above is the unique $n+1$-dimensional irreducible representation of $\mathfrak{sl}(2,\mathbb{C})$.*

Proof. Let V be some irreducible $n+1$-dimensional representation. Let $v \in V$ be some eigenvector of $\rho(h)$ for the eigenvalue $\lambda \in \mathbb{C}$. Then

$$\rho(h)\rho(e)v = \rho(e)\rho(h)v + \rho([h,e])v = (\lambda+2)\rho(e)v.$$

Hence $\rho(e)v$ is also an eigenvector of $\rho(h)$ for a different eigenvalue. Similarly, one checks that

$$\rho(h)\rho(f)v = (\lambda-2)\rho(f)v.$$

By finite dimensionality the eigenvector v may be picked such that $\rho(e)v = 0$. Define $f_0, f_1, \dots$ such that $f_j = \rho(f)^j v$. Then clearly $\rho(h)f_j = (\lambda - 2j)f_j$ and $\rho(f)f_j = f_{j+1}$. We show by induction that $\rho(e)f_j = j(\lambda - j + 1)f_{j-1}$. For $j = 0$ this is true by definition of v. Assume the claim up to some j. Compute

$$\rho(e)f_{j+1} = \rho(e)\rho(f)f_j = \rho(f)\rho(e)f_j + \rho([e,f])f_j = (j(\lambda-j+1)+\lambda-2j)f_j = ((j+1)\lambda - j(j+1))f_j.$$

Let W be the span of all f_j. It is a non-zero invariant subspace and hence $W = V$. It also follows that $f_{n+1} = f_{n+2} = \dots = 0$, while $f_n \neq 0$. The only remaining fact to check is that $\lambda = n$. Compute

$$0 = tr([\rho(e),\rho(f)]) = tr(\rho(h)) = \sum_{j=0}^{n}(\lambda - 2j) = (n+1)\lambda - 2\frac{n(n+1)}{2} = (n+1)(\lambda - n).$$

Hence $\lambda = n$ since $n+1 > 0$. □

Exercise 4.9. Find an infinite dimensional irreducible representation of $\mathfrak{sl}(2,\mathbb{C})$.

For a representation ρ of $\mathfrak{sl}(2,\mathbb{C})$ we define the Casimir element

$$Z = \frac{1}{2}\rho(h)^2 + \rho(h) + 2\rho(f)\rho(e).$$

Lemma 4.2. *For any representation ρ, the Casimir element Z commutes with all $\rho(x)$, $x \in \mathfrak{sl}(2,\mathbb{C})$. For the irreducible representation V_n it acts by multiplication with $\frac{1}{2}n^2 + n$.*

Proof. Let us check

$$[\rho(e), Z] = -\rho(e)\rho(h) - \rho(h)\rho(e) + 2\rho(e) + 2\rho(h)\rho(e) = 0.$$

The other calculations are similar. On V_n we compute

$$Ze_0 = (\frac{1}{2}n^2 + n)e_0.$$

By Schur's Lemma Z must act as a constant and hence the result follows. □

Theorem 4.9. *Every finite dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$ is completely reducible, i. e., isomorphic to a direct sum of representations of the form V_n above.*

Let V be the representation in question. We will do an induction on $\dim V$. For $\dim V = 1$ the statement follows from the previous Theorem since V is irreducible. Otherwise, proceeding as in the proof of the previous Theorem, we may identify an invariant subspace $W \subset V$, such that $W \cong V_n$ for some n.

Claim 1: W has an invariant complement, i.e., $V \cong W \oplus W'$ for some invariant subspace.

Assuming the claim, the Theorem follows by invoking the induction hypothesis on W' (note: $\dim W' < \dim V$). To show the claim we will construct a projector p onto W that commutes with the $\mathfrak{sl}(2,\mathbb{C})$-action. Concretely, we want that $p \in \mathrm{Hom}(V,W) \subset \mathrm{Hom}(V,V)$, $p\mid_W = \mathbb{1}_W$ and $[\rho(x), p] = 0$ for all $x \in \mathfrak{sl}(2,\mathbb{C})$. Then we may take $W' = (\mathbb{1} - p)V \subset V$.

We consider $\mathrm{Hom}(V,V)$ as a representation of $\mathfrak{sl}(2,\mathbb{C})$. It is easy to check that $\mathrm{Hom}(V,W)$ is an invariant subspace. Consider the subspace $X \subset \mathrm{Hom}(V,W)$ consisting of morphisms F such that $F\mid_W = \lambda\mathbb{1}$ for some $\lambda \in \mathbb{C}$, and let $X_0 \subset X$ be the space of morphisms such that $F\mid_W = 0$. Clearly X and X_0 are invariant subspaces, since for $w \in W$, $F \in X$, $x \in \mathfrak{sl}(2,\mathbb{C})$

$$[\rho(x), F]w = \lambda\rho(x)w - \lambda\rho(x)w = 0.$$

Furthermore $X_0 \subset X$ is of codimension 1.

Lemma 4.3. *Suppose that U is a finite dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$ and $U_0 \subset U$ is an invariant subspace of codimension 1. Then there is an invariant complement $U_1 \subset U$, i.e., $U \cong U_0 \oplus U_1$.*

Note that U_1 is one-dimensional and hence $\mathfrak{sl}(2,\mathbb{C})$ acts trivially on U_1. Assuming the Lemma, let us finish the proof of the Theorem. Let p be a non-zero element of an invariant complement of X_0 in X. We may assume that $p\,|_W = \mathbb{1}_W$ by rescaling. Since as we saw above $\mathfrak{sl}(2,\mathbb{C})$ sends X to X_0 and hence p to zero. So p is the sought after invariant projector onto W.

Proof of the Lemma. Note that picking a not necessarily invariant complement of U_0 all reprresentation matrices $\rho(x)$ have the form

$$\rho(x) \sim \begin{pmatrix} a & 0 \\ b & A \end{pmatrix}$$

where a is a scalar, b is a column vector and A is a square matrix. Since all elements of $\mathfrak{sl}(2,\mathbb{C})$ are commutators, we can conclude that $a = 0$. Furthermore if $A = 0$ for all x, then also $b = 0$. This shows that in particular in the case $\dim U_0 = 1$, the action is trivial and any complement to U_0 is invariant.

Suppose next that U/U_0 is irreducible, of dimension > 1. Then the Casimir Z acts as a non-zero constant on U and its image is of codimension 1, since the Casimir element of U/U_0 is zero. Hence $\ker Z$ is of dimension 1 and is an invariant complement, since the kernel of an intertwiner is an invariant subspace.

The remaining cases are handled by induction on $\dim U$. The cases $\dim U \le 2$ are trivial or handled above. Suppose U_0 is not irreducible. Then find an irreducible subspace $A \subset U_0$. By the induction hypothesis we may find an invariant complement Y' of $U_0/A \subset U/A$. Denote the preimage of Y' under the projection $U \to U/A$ by Y. It is invariant, has dimension $\dim A + 1$ and contains A. By the induction hypothesis there is an invariant complement $U_1 \subset Y \subset U$ such that $Y = U_1 \oplus A$. In particular, U_1 is a one-dimensional invariant subspace not contained in $Y + U_0$, hence the invariant complement we looked for. $\square$

Corollary 4.4. *Let ρ be a representation of $\mathfrak{sl}(2,\mathbb{C})$ on the finite dimensional complex vector space V.*

- *All operators $\rho(h)$ are diagonalizable, the eigenvalues are integers and the multiplicities of the eigenvalues k and $-k$ are equal.*
- *There is an inner product on V such that $\rho(h)$ is self-adjoint, $\rho(e)^\dagger = \rho(f)$ and such that the representation of $\mathfrak{su}_2 \subset \mathfrak{sl}(2,\mathbb{C})$ is unitary. (I.e., all $\rho(x)$, $x \in \mathfrak{su}_2$ are anti-self-adjoint.)*

Remark 4.3. There is also another proof of the Theorem using "averaging" similar to the proof of the analogous statement in the finite group case. Concretely the key step of finding an invariant complement to $W \subset V$ is done as follows:

1. From the representation of $\mathfrak{sl}(2,\mathbb{C})$ we obtain a complex representation of the real Lie algebra $\mathfrak{su}_2$.
2. Suppose $\mathfrak{g}$ is the Lie algebra of a simply connected Lie group G, $\mathfrak{h}$ is the Lie algebra of some other Lie group H and $f : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism. Then there is a homomorphism of Lie groups $F : G \to H$ inducing f. In particular, any representation of $\mathfrak{g}$ lifts to a representation of G.
3. On every Lie group there is an invariant measure, the Haar measure.
4. Let $\langle , \rangle'$ be some inner product on V. Define an inner product

$$\langle v, w\rangle := \int_{SU(2)} \langle \rho(g)v, \rho(g)w\rangle' \, d\mu(g)$$

   where μ is the Haar measure. One easily checks that $\langle , \rangle$ is invariant, i.e., all representation matrices $\rho(g)$ are unitary w.r.t. this inner product.
5. The orthogonal complement to an invariant subspace is invariant and hence we may pick $W' = W^\perp$.

Exercise 4.10.
- Show that if $\mathfrak{g}$ is real or complex semisimple, then $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$.
- Let e_j be a basis of $\mathfrak{g}$ and let e^j be the dual basis w.r.t. the Killing form. Show that for any representation ρ the Casimir element $Z = \sum_j \rho(e_j)\rho(e^j)$ does not depend on the choice of basis and commutes with $\rho(\mathfrak{g})$.
- Suppose the representation is irreducible. Check that Z acts trivially iff the representation is trivial.
- Mimick the above proof to show that any finite dimensional representation of a semi-simple Lie algebra is completely reducible.

4.9 General structure theory of semi-simple Lie algebras

Definition 4.11. *Let $\mathfrak{g}$ be a complex semi-simple Lie algebra. A subalgebra $\mathfrak{h} \subset \mathfrak{g}$ is called toral if it is abelian and all ad_x, $x \in \mathfrak{h}$ are diagonalizable. It is called maximal toral or Cartan subalgebra (or maximal torus) if it is maximal among the toral subalgebras.*

Without proof we state the following results:

Proposition 4.3. *Let $\mathfrak{h} \subset \mathfrak{g}$ be a Cartan subalgebra and suppose $x \in \mathfrak{g}$ is such that $[x, h] = 0$ for all $h \in \mathfrak{h}$. Then $x \in \mathfrak{h}$.*

Proof. Serie 11, exercise 2. □

Theorem 4.10. *Let $\mathfrak{h}_1, \mathfrak{h}_2 \subset \mathfrak{g}$ be two Cartan subalgebras of the semi-simple Lie algebra $\mathfrak{g}$. Then there exists an automorphism ϕ of $\mathfrak{g}$ such that $\phi(\mathfrak{h}_1) = \mathfrak{h}_2$.*

In particular, all Cartan subalgebras have the same dimension, which is called the *rank* of the Lie algebra $\mathfrak{g}$.

Example 4.8. Consider $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C})$. Then the diagonal (traceless) matrices form an abelian subalgebra $\mathfrak{h}$, and all ad_x are clearly diagonalizable. This subalgebra is maximal, since a matrix commuting with all traceless diagonal matrices must itself be diagonal. Hence $\mathfrak{h}$ is Cartan, and the rank of $\mathfrak{sl}(n, \mathbb{C})$ is $n - 1$.

Exercise 4.11. Verify the Theorem above for $\mathfrak{sl}(n, \mathbb{C})$. In other words: Let $\mathfrak{h}$, $\mathfrak{h}'$ be two Cartan subalgebras. Show that there is an invertible matrix A such that $\mathfrak{h}' = A\mathfrak{h}A^{-1}$. *Hint: A matrix x is diagonalizable iff ad_x is.*

Let us simultaneously diagonalize all ad_x, $x \in \mathfrak{h}$. This yields a decomposition (as $\mathfrak{h}$ representations)

$$\mathfrak{g} \cong \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha$$

where $\Delta \subset \mathfrak{h}^*$ is a finite subset of the dual space of $\mathfrak{h}$, the set of non-zero simultaneous eigenvalues, and $\mathfrak{g}_\alpha \subset \mathfrak{g}$ is the eigenspace w.r.t. α. In other words ad_x acts as $\lambda(x)$ on $\mathfrak{g}_\alpha$. Note also that $\mathfrak{h}$ is the eigenspace corresponding to the simultaneous eigenvalue 0. Indeed, $[\mathfrak{h}, \mathfrak{h}] = 0$ and if X commutes with $\mathfrak{h}$ then $X \in \mathfrak{h}$ by the proposition above. The elements of Δ are called the *roots* of $\mathfrak{g}$ (with respect to $\mathfrak{h}$). Elements of $\mathfrak{g}_\alpha$ are called root vectors for the root α.

Let us state some immediate properties of the roots. Recall that $K(x, y) := tr(\mathrm{ad}_x \mathrm{ad}_y)$ is the Killing form.

Proposition 4.4. *Let $\mathfrak{g}$ be a finite dimensional complex semi-simple Lie algebra and fix a Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$ as above. Then:*

- $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$.
- *If $\alpha, \beta \in \{0\} \cup \Delta$ and $\alpha + \beta \neq 0$, then $K(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$.*
- *If $\alpha \in \{0\} \cup \Delta$ then K is non-degenerate on $\mathfrak{g}_\alpha \times \mathfrak{g}_{-\alpha}$. In particular K is non-degenerate on $\mathfrak{h}$. Hence for each root α there is an $H_\alpha \in \mathfrak{h}$ such that $K(H_\alpha, h) = \alpha(h)$ for all $h \in \mathfrak{h}$.*
- $\alpha \in \Delta \Rightarrow -\alpha \in \Delta$.
- *The Killing form on $\mathfrak{h}$ is $K(h, h') = \sum_{\alpha \in \Delta} \alpha(h)\alpha(h')$.*
- *Δ spans $\mathfrak{h}^*$.*

Proof. Note that for $h \in \mathfrak{h}, a \in \mathfrak{g}_\alpha, b \in \mathfrak{g}_\beta$,

$$[h, [a, b]] = [[h, a], b] + [a, [h, b]] = (\alpha(h) + \beta(h))[a, b].$$

Hence follows the first statement. The second statement is immediate since $\mathrm{ad}_a \mathrm{ad}_b$ has zero (block-)diagonal elements. For the third statement: Note that K is non-degenerate and hence for each non-zero a as above there must be a a' such that $K(a, a') \neq 0$. By the second statement $a' \in \mathfrak{g}_{-\alpha}$. From this the fourth statement follows immediately.

The last but one statement is obvious.

Last statement: Suppose $h \in \mathfrak{h}$ is such that $\alpha(h) = 0$ for all $\alpha \in \Delta$. Then $\mathrm{ad}_h = 0$ contradicting the assumption of non-degeneracy of the Killing form.

□

Example 4.9. In $\mathfrak{sl}(n,\mathbb{C})$ with the Cartan subalgebra above the root vectors are the matrices $E_{i,j}$ (with $i \neq j$) with entry 1 in the i,j slot and all other entries 0. The corresponding root sends the diagonal matrix D to $D_{ii} - D_{jj}$. Those linear maps indeed span the dual space to the diagonal traceless matrices. The Killing form is (up to a constant multiple) $tr(xy)$. Hence it is clear that $E_{i,j}$ is orthogonal to $E_{k,l}$ for $(i,j) \neq (k,l)$.

The link to the representation theory of $\mathfrak{sl}(2,\mathbb{C})$ is as follows.

Lemma 4.4. *Pick a nonzero $E_\alpha \in \mathfrak{g}_\alpha$ and $E_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $K(E_\alpha, E_{-\alpha}) = 1$. (It is possible by the proposition above.) Then $[E_\alpha, E_{-\alpha}] = H_\alpha$. Furthermore $\alpha(H_\alpha) \neq 0$ and*

$$h = \frac{2}{\alpha(H_\alpha)} H_\alpha \qquad e = \frac{2}{\alpha(H_\alpha)} E_\alpha \qquad f = E_{-\alpha}$$

span a Lie subalgebra isomorphic to $\mathfrak{sl}(2,\mathbb{C})$.

Proof. For all $h \in \mathfrak{h}$

$$K([E_\alpha, E_{-\alpha}], h) = K(E_\alpha, [E_{-\alpha}, h]) = \alpha(h) K(E_\alpha, E_{-\alpha}) = \alpha(h).$$

Now suppose $\alpha(H_\alpha) = 0$ and pick some root β such that $\beta(H_\alpha) \neq 0$. Let $V := \bigoplus_{n\in\mathbb{Z}} \mathfrak{g}_{\beta+n\alpha}$. Since H_α is a commutator,

$$0 = tr_V \mathrm{ad}_{H_\alpha} = \sum_n d_n(\beta(H_\alpha) + n\alpha(H_\alpha)) = \beta(H_\alpha) \sum_n d_n \neq 0.$$

where d_n are the dimensions of $\mathfrak{g}_{\beta+n\alpha}$. This is a contradiction and hence $\alpha(H_\alpha) \neq 0$.

Finally the relations of $\mathfrak{sl}(2,\mathbb{C})$ are clearly satisfied. □

This Lemma, together with the results about the representation theory of $\mathfrak{sl}(2,\mathbb{C})$ allows one to deduce a remarkable amount of information about $\mathfrak{g}$.

Let us define a nondegenerate bilinear form $\langle , \rangle$ on $\mathfrak{h}^*$ by dualizing the Killing form, in particular,

$$\langle \alpha, \beta \rangle := K(H_\alpha, H_\beta).$$

Let $\alpha \in \Delta, \beta \in \{0\} \cup \Delta$. The α-string through β is defined to be the subset

$$\oplus_{n\in\mathbb{Z}} \mathfrak{g}_{n\alpha+\beta}.$$

Proposition 4.5. *For α, β as above the α-string through β has nonzero terms for $-p \leq n \leq q$ with $p, q \geq 0$, with no gaps. Furthermore*

$$p - q = \frac{2\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} \in \mathbb{Z}.$$

Proof. Let h, e, f be as above. Then ad_h acts as

$$\frac{2}{\langle \alpha, \alpha \rangle}(\langle \alpha, \beta \rangle + n\langle \alpha, \alpha \rangle) = \frac{2\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} + 2n$$

on the n-th piece of the string. The result follows from Theorem **??** above. Indeed we must have that the possible values of n satisfy

$$-k \leq \frac{2\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} + 2n \leq k$$

for some k. In other words

$$2p - 2q = k + \frac{2\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} - (k - \frac{2\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}) = 2\frac{2\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}.$$

□

We denote by $\mathfrak{h}_0$ the real span of the elements H_α, $\alpha \in \Delta$.

Corollary 4.5. *$\mathfrak{h}_0$ is a real form for $\mathfrak{h}$, i.e., the natural map $\mathfrak{h}_0 \otimes \mathbb{C} \to \mathfrak{h}$ is an isomorphism. The real span $\mathbb{R}\Delta$ of Δ is a real form for $\mathfrak{h}^*$, naturally identified with $\mathfrak{h}_0^*$. The bilinear form $\langle , \rangle$ above is positive definite on $\mathbb{R}\Delta \cong \mathfrak{h}_0^*$. The Killing form is positive definite on $\mathfrak{h}_0$.*

Proof. First compute

$$\langle\alpha,\alpha\rangle = K(H_\alpha,H_\alpha) = \sum_{\beta\in\Delta}\langle\alpha,\beta\rangle^2 = \langle\alpha,\alpha\rangle^2\sum_{\beta\in\Delta}(\frac{1}{2}(p_\beta-q_\beta))^2)$$

using the previous proposition. Solving this equation for $\langle\alpha,\alpha\rangle$ (using $\langle\alpha,\alpha\rangle\neq 0$), we see that $\langle\alpha,\alpha\rangle\in\mathbb{Q}\subset\mathbb{R}$. Hence, using the previous proposition again, we find that $\langle\alpha,\beta\rangle\in\mathbb{Q}\subset\mathbb{R}$ for all $\alpha,\beta\in\Delta$. It follows that for any $\lambda\in\mathbb{R}\Delta$

$$\langle\lambda,\lambda\rangle = \sum_{\beta\in\Delta}\langle\lambda,\beta\rangle^2\geq 0$$

with equality iff λ is orthogonal to all roots. But that means $\lambda=0$ by non-degeneracy of $K(,)$ and since Δ spans $\mathfrak{h}^*$ over $\mathbb{C}$. Similarly, one shows that $K(,)$ is positive definite on $\mathfrak{h}_0$.

Next note that the maps $\mathfrak{h}_0\otimes\mathbb{C}\to\mathfrak{h}$, $\mathbb{R}\Delta\otimes\mathbb{C}\to\mathfrak{h}^*$ are onto. It will suffice to show that $dim\mathfrak{h}_0 = dim\mathbb{R}\Delta = r := dim_\mathbb{C}\mathfrak{h}$, and we know that $dim\mathfrak{h}_0, dim\mathbb{R}\Delta\geq r$. Since (as we saw) elements of $\mathbb{R}\Delta$ are real on $\mathfrak{h}_0$, $dim\mathbb{R}\Delta\leq 2r - dim\mathfrak{h}_0\leq r$ (exercise). Hence $dim\mathbb{R}\Delta = r = dim\mathfrak{h}_0$. □

Next we want to split the set of roots into positive and negative roots. Pick some indices $\alpha_1,\dots,\alpha_r$ such that $H_{\alpha_1},\dots,H_{\alpha_r}$ is a basis of $\mathfrak{h}_0$. We say that $0\neq\lambda\in\mathfrak{h}_0^*$ is positive, and write $\lambda>0$, if the first non-zero number in the list $\lambda(H_{\alpha_1}),\lambda(H_{\alpha_2}),\dots$ is positive. We say that λ is negative otherwise. In general, we write $\lambda>\lambda'$ iff $\lambda-\lambda'>0$ and $\lambda\geq\lambda'$ iff $\lambda>\lambda'$ or $\lambda=\lambda'$. This induces a total ordering on the elements of $\mathfrak{h}_0^*$.

Exercise 4.12. If $\lambda>0$ then $-\lambda$ is negative. The sum of positive elements is positive. Positive multiples of positive elements are positive.

The notion of positivity is not canonical, but depends on a choice we have to make.

Definition 4.12. *A root $\alpha\in\Delta$ is called simple if $\alpha>0$ and α cannot be written as the sum of two positive roots. The set of simple roots is denoted by* Π.

Example 4.10. (Exercise) Consider the case $\mathfrak{sl}(n,\mathbb{C})$ with the Cartan subalgebra $\mathfrak{h}$ the diagonal traceless matrices. The dual space $\mathfrak{h}^*$ may be identified with diagonal matrices, modulo the identity, or also with traceless matrices. The roots are (i.e., can be identified with) matrices $\alpha_{i,j}$ ($i\neq j$) with entries 1 and -1 in positions i and j along the diagonal. Under this identification $H_{\alpha_{i,j}}$ is the same matrix. Also note that all roots have length 2. $\mathfrak{h}_0$ are the real diagonal traceless matrices. If we pick a basis $H_{\alpha_{1,2}}, H_{\alpha_{2,3}},\dots,H_{\alpha_{n-1,n}}$, then the ordering on elements of $\mathfrak{h}_0^*$ is the lexicographic ordering with respect to the entries on the diagonal. The roots $\alpha_{i,i+1}$ are the simple roots (exercise).

Proposition 4.6. *The simple roots form a basis of $\mathfrak{h}_0^*$. Any positive root α may be written as*

$$\alpha = \sum_{j=1}^r n_j\alpha_j$$

where $\alpha_1,\dots,\alpha_r$ are the simple roots and $n_1,\dots,n_r$ are non-negative integers.

Proof. Exercise. □

Exercise 4.13. Show that for $\alpha,\beta\in\Delta$ roots, $\frac{2\langle\alpha,\beta\rangle}{\langle\alpha,\alpha\rangle}\in\{0\pm 1,\pm 2,\pm 3\}$. If α,β are distinct and simple, then only the non-positive numbers can occur. The Dynkin diagram of a semisimple Lie algebra $\mathfrak{g}$ is a graph with r nodes corresponding to the simple roots, and with $k=\frac{4\langle\alpha,\beta\rangle^2}{\langle\alpha,\alpha\rangle\langle\beta,\beta\rangle}$ edges connecting the nodes corresponding to α and β. Draw the Dynkin diagrams for $\mathfrak{sl}(n,\mathbb{C})$, $\mathfrak{sp}(2n,\mathbb{C})$ and $\mathfrak{so}(2n(+1),\mathbb{C})$.

4.10 Representation theory of complex semi-simple Lie algebras

Let $\mathfrak{g}$ be a complex semi-simple Lie algebra and fix a Cartan subalgebra $\mathfrak{h}\subset\mathfrak{g}$. Let Δ be the set of roots. We want to study finite dimensional representations of $\mathfrak{g}$. Let ρ be such a representation on the (finite dimensional complex) vector space V.

Lemma 4.5. *Let ρ be a finite dimensional representation of the complex semi-simple Lie algebra $\mathfrak{g}$. Fix a Cartan subalgebra $\mathfrak{h}\subset\mathfrak{g}$ and define the elements H_α as above. Then each $\rho(H_\alpha)$ is diagonalizable. The eigenvalues are integer multiples of $\frac{\langle\alpha,\alpha\rangle}{2}$.*

Proof. We may restrict the representation of $\mathfrak{g}$ to a representation of the subalgebra spanned by $H_\alpha, E_\alpha, E_{-\alpha}$. The latter subalgebra is isomorphic to $\mathfrak{sl}(2,\mathbb{C})$, with $2H_\alpha/|\alpha|^2$ playing the role of h. Hence, by the classification of representations of $\mathfrak{sl}(2,\mathbb{C})$ (Theorem **??** above), the result follows. □

By the Lemma we can simultaneously diagonalize all $\rho(h), h \in \mathfrak{h}$. The corresponding decomposition of V is

$$V \cong \bigoplus_{\lambda \in W_V} V_\lambda$$

where for $\lambda \in \mathfrak{h}^*$

$$V_\lambda = \{v \in V \mid \rho(h)v = \lambda(h)v \forall h \in \mathfrak{h}\}$$

and the sum is over some finite subset $W_V \subset \mathfrak{h}^*$,

$$W_V = \{\lambda \in \mathfrak{h}^* \mid V_\lambda \neq 0\}.$$

W_V is called the set of *weights* of the representation V.

It is clear that $\rho(\mathfrak{g}_\alpha)V_\lambda \subset V_{\lambda+\alpha}$. Furthermore, by the Lemma above

$$\frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle} = \frac{2\lambda(H_\alpha)}{\langle\alpha,\alpha\rangle} \in \mathbb{Z}$$

for all $\alpha \in \Delta$.

Definition 4.13. *A linear form $\lambda \in \mathfrak{h}^*$ is called (algebraically) integral if $\frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle} \in \mathbb{Z}$ for all $\alpha \in \Delta$. The algebraically integral weights form a discrete subset of $\mathfrak{h}^*$ called the weight lattice.*

Example 4.11. For $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{C})$ with standard Cartan subalgebra $\mathbb{C}h$, the roots are $\pm\alpha$, where $\alpha(h) = 2$, and the weight lattice is $\mathbb{Z}\alpha/2$.

Let us fix a total ordering on the elements of $\mathfrak{h}_0^*$ as above, and let Π be the corresponding set of simple roots. In particular we may talk about the heighest weight of a representation.

Definition 4.14. *A linear form $\lambda \in \mathfrak{h}_0^*$ is called* dominant *if $\langle\lambda,\alpha\rangle \geq 0$ for all $\alpha \in \Pi$ (or equivalently for all positive roots α).*

Lemma 4.6. *The heighest weight of any finite dimensional representation V is dominant.*

Proof. Let λ be the heighest weight and consider a positive root α. Let $E_{\pm\alpha}$, H_α be as before, and denote by $\mathfrak{k}$ the subalgebra isomorphic to $\mathfrak{sl}(2,\mathbb{C})$ they span. Since $\alpha + \lambda > \lambda$ since $\alpha > 0$, we must have $V_{\lambda+\alpha} = 0$. Hence $\rho(E_\alpha)V_\lambda \subset V_{\lambda+\alpha} = 0$. Consider V as a representation of $\mathfrak{k}$. Then by the classification of representations of $\mathfrak{sl}(2,\mathbb{C})$ we know that $2H_\alpha/|\alpha|^2$ acts by a non-negative constant. Hence $\lambda(H_\alpha) = \langle\lambda,\alpha\rangle \geq 0$. □

Theorem 4.11. *Let $\mathfrak{g}$ be semi-simple, fix a Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$, and a notion of positivity on the roots as above. Then there is a bijection between the isomorphism classes of finite dimensional irreducible representations of $\mathfrak{g}$ and the dominant algebraically integral linear forms $\lambda \in \mathfrak{g}_0^*$. Concretely, the bijection sends a representation to its heighest weight.*

Remark 4.4. Pick a dual basis α_j^* to the basis $\frac{2\alpha_j}{|\alpha_j|^2}$, i.e., $\langle\alpha_j^*, \alpha_i\rangle = \delta_{ij}|\alpha_j|^2/2$. Then any dominant algebraically integral weight can be written uniquely in the form

$$\sum_j n_j \alpha_j^*$$

where the n_j are non-negative integers.

Example 4.12. Consider $\mathfrak{g} = \mathfrak{sl}(n,\mathbb{C})$ with Cartan subalgebra and positive roots as before. As before, we may identify elements of $\mathfrak{h}_0^*$ with diagonal matrices, modulo the identity. The condition of dominance means that the entries along the diagonal are non-decreasing. The condition of integrality means that the differences of any pair of entries along the diagonal are integers. Hence the dominant integral weights are in 1-1 correspondence to sequences of integers

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n = 0.$$

Note that these numbers naturally determine a Young diagram with $n-1$ rows.

To prove the Theorem above, we have to show that the map between isomorphism classes of irreducible representations and dominant integral linear forms is surjective and injective. We are done if we can show the following statement:

Claim: For every dominant integral weight λ there is a finite dimensional irreducible representation U_λ with heighest weight λ with the following universal property: Let W be another irreducible representation with heighest weight λ and fix heighest weight vectors $u \in U_\lambda, w \in W$. Then there is a unique intertwiner $U_\lambda \to W$ sending u to w.

If the claim is true than the map of the Theorem is clearly surjective. Furthermore, if W, W' are two irreps with heighest weight λ, then there are maps $W \leftarrow U_\lambda \to W'$. By Schur's Lemma, both maps are isomorphisms.

To construct U_λ above, let us construct another (infinite dimensional) representation V_λ with heighest weight λ, generated by a highest weight vector $v \in V_\lambda$ and the following universal property.

Let W be a representation with heighest weight λ generated by a heighest weight vector $w \in W$. Then there is a unique intertwiner $V_\lambda \to W$ sending v to w.

Such a V_λ will be called *Verma module.*

Let

$$U\mathfrak{g} = \left(\bigoplus_{n=0}^{\infty} \mathfrak{g}^{\otimes n}\right) / \langle x \otimes y - y \otimes x - [x, y]\rangle$$

be the universal enveloping algebra. It is a representation of $\mathfrak{g}$ by left multiplication. Let $\mathfrak{n}_+ = \oplus_{\alpha\in\Delta^+}\mathfrak{g}_\alpha \subset \mathfrak{g}$ be the space spanned by all positive root vectors. Then clearly $\mathfrak{b} := \mathfrak{h} \oplus \mathfrak{n}_+$ is a subalgebra (the Borel subalgebra).

Example 4.13. For $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C})$ the Borel subalgebra is the subalgebra of (traceless) upper triangular matrices.

Denote by $\mathbb{C}_\lambda$ the representation of $\mathfrak{b}$ on $\mathbb{C}$ such that $h \in \mathfrak{h}$ acts by $\lambda(h)$ and all E_α act as zero.

Proposition 4.7. *The action of $\mathfrak{g}$ on $U\mathfrak{g}$ descends to*

$$V_\lambda := U\mathfrak{g} \otimes_{U\mathfrak{b}} \mathbb{C}_\lambda = U\mathfrak{g} \otimes \mathbb{C}_\lambda / span(\{ab \otimes 1 - \lambda(b)a \otimes 1 \mid a \in U\mathfrak{g}, b \in U\mathfrak{b}\}).$$

Furthermore, V_λ satisfies the universal property above.

V_λ is in general infinite dimensional and not irreducible. However, define

$$U_\lambda := V_\lambda / \sum_X X$$

where the sum is over all proper irreducible subspaces.

Exercise 4.14. Show that U_λ is irreducible.

Theorem 4.12. *U_λ is irreducible, has heighest weight λ and is generated by the image of $1 \otimes 1 \in V_\lambda$. If $\lambda \in \mathfrak{h}^*$ is real valued on $\mathfrak{h}_0$, dominant and integral then U_λ is finite dimensional. In fact the Weyl dimension formula*

$$\dim U_\lambda = \frac{\prod_{\alpha\in\Delta^+}\langle\lambda+\rho,\alpha\rangle}{\prod_{\alpha\in\Delta^+}\langle\rho,\alpha\rangle}$$

holds, where $\rho := \frac{1}{2}\sum_{\alpha\in\Delta^+}\alpha$.

There is also an explicit formula for the multiplicity of the weights, i.e., for the dimension of the space $U_{\lambda,\mu} \subset U_\lambda$ on which each $h \in \mathfrak{h}$ acts as $\mu(h)$. (The Kostant multiplicity formula.)

Example 4.14. For $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{C})$ the dominant integral λ correspond to integers n, such that $\lambda(h) = n$. The Verma module is infinite dimensional and spanned by vectors $e_0, e_1, \ldots$ such that $\rho(h)e_j = (n-2j)e_j$. Concretely, $e_j = f^j \otimes 1$. One shows easily that $\rho(e)e_j = j(n-j+1)e_{j-1}$.

Consider the vector e_{-n-1}. We have $\rho(e)e_{-n-1} = 0$ by the above formula. Hence the subspace spanned by $e_{-n-1}, e_{-n-2}, \ldots$ is invariant. Quotienting by this subspace we find the finite dimensional representation of $\mathfrak{sl}(2, \mathbb{C})$ we constructed before.

Chapter 5

Space groups

5.1 Introduction

The goal of this chapter is to study "periodic patterns", patterns that have a translational symmtery. (see handout).

In general, consider the Euclidean group in d dimensions $\mathrm{Euc}(d)$:

Definition 5.1. *The Euclidean group* $\mathrm{Euc}(d)$ *is the group of pairs*

$$(t, R) \in \mathbb{R}^d \times O(d)$$

of a d-dimensional vector and an orthogonal matrix. The group composition is

$$(t, R)(t', R') = (t + Rt', RR')$$

the unit is $(0, \mathbb{1})$ *and the inverse is*

$$(t, R)^{-1} = (-R^T t, R^T).$$

Elements of the Euclidean group are transformations of the d dimensional affine space, (t, R) acting on some $x \in \mathbb{R}^d$ as $(t, R)x = t + Rx$. There is a short exacty sequence of groups (i. e., the image of each morphism is equal to the kernel of the next)

$$1 \to \mathbb{R}^d \to \mathrm{Euc}(d) \to O(d) \to 1.$$

Here the second map is $t \mapsto (t, \mathbb{1})$ and the third is $(t, R) \mapsto R$. The symmetry groups of the patterns we consider are subgroups $G \subset \mathrm{Euc}(d)$. To each such subgroup we may associate a group of pure translations

$$G_t = \{t \in \mathbb{R}^d | (t, \mathbb{1}) \in G\} \subset \mathbb{R}^d \subset \mathrm{Euc}(d)$$

and a group of orthogonal transformations

$$G_o = \{R \mid \exists t : (t, R) \in G\} \subset O(d) \subset \mathrm{Euc}(d).$$

They fit into a short exact sequence

$$1 \to G_t \to G \to G_o \to 1.$$

We want to study groups for which both G_t and G_o are discrete.

Definition 5.2. *(Non-standard) We say a subgroup* $G \subset \mathrm{Euc}(d)$ *is a* (d, k)*-space group if*

- G_0 *is finite.*
- G_t *is a* k *dimensional lattice, i. e. there are linearly independent vectors* $e_1, \ldots, e_k$ *such that* $G_t = \mathbb{Z}\langle e_1, \ldots, e_k \rangle$.

More standard names for these groups are as follows

- $(2, 1)$-space groups are called frieze groups. They are the isomorphism groups of (infinite) friezes as in the examples.

- $(2,2)$-space groups are called wallpaper groups. They are the isomorphism groups of (infinite) floor tilings.
- $(3,3)$-space groups are called space groups. They are the isomorphism groups of (infinite) crystall lattices.

We want to classify these types of groups (up to group automorphism) in the most relevant cases. One may proceed as follows:

1. Classify all finite subgroups $H \subset O(d)$. These are also called *point groups*. Note that we will classify the group together with the embedding into $O(d)$, up to conugation in $O(d)$. In particular, we will count two isomorphic groups with non-isomorphic embeddings into $O(d)$ as two distinct cases.

2. Note that G_o acts on G_t by conjugation: If $(s,R) \in G$, $(t,\mathbb{1}) \in G_t$ then

$$(s,R)(t,\mathbb{1})(s,R)^{-1} = (s,R)(t,\mathbb{1})(-R^T s, R^T) = (Rt,\mathbb{1}) \in G_t.$$

   In particular, the lattice G_t must be left invariant under the G_o action. So we should classify all pairs (H,Λ) such that Λ is a lattice on which H acts. For a given H two pairs (H,Λ), (H,Λ') are identified if the groups $H \ltimes \Lambda$ and $H \ltimes \Lambda'$ are isomorphic. The lattices accurring here are called *Bravais lattices*.

3. Once we know G_t, G_o and the action of G_o on G_t we may classify all groups G that fit inside the sequence

$$1 \to G_t \to G \to G_o \to 1.$$

   In practice, G_o will have a small number of generators $R_1, R_2, \dots$ and relations. There are (to be found) $\tau_1, \tau_2, \dots \in \mathbb{R}^d$, such that G is generated by G_t and the $A_j := (\tau_j, R_j)$. For every relation $X(R_1, R_2, \dots) = \mathbb{1}$ of G_o, we have to solve equations of the form

$$X(A_1, A_2, \dots) \in G_t.$$

   Furthermore, composition A_j with some $t_j \in G_t$ changes τ_j by t_j. Hence the τ_j differing by lattice vectors in G_t may be identified. Futhermore, we may conjugate the whole group G by a translation $t \in R^d$. This changes each τ_j to $\tau_j + (\mathbb{1} - R_j)t$. In practice (i.e., in the cases we consider), this leaves only finitely many cases to be checked.

5.2 Toy example: (1,1)-space groups

Let us consider the 1-dimensional (toy) case. Since $O(1) = \mathbb{Z}_2$ is finite, there are only two possible choices for G_o, namely $G_o = 1$, $G_o = O(2)$.

In the first case ($G_o = 1$), any lattice is G_o invariant, and we obtain the symmetry group $\mathbb{Z}$. A 1-d pattern that has this symmetry is the following (for example)

$$\cdots \to\to\to\to \cdots .$$

If $G_o = O(2) = \{\pm 1\}$, then again any lattice is G_o invariant. We may assume $G_t = \mathbb{Z}$ by rescaling. G_o has one generator (-1) and one relation $(-1)^2 = 1$. So we need to find $\tau \in \mathbb{R}$, such that

$$(\tau, -1)$$

and $\mathbb{Z}$ generate G. In fact τ's which differ by elements in $G_t \cong \mathbb{Z}$ may be identified, so we may in fact assume that $0 \leq \tau < 1$. The single relation leaves us with the equation

$$(\tau, -1)^2 = (\tau - \tau, 1) \in G_t$$

which is clearly satisfied for all τ. Conjugation by a global translation $t = -\tau/2$ will change τ to $(1-(-1))t + \tau = 0$. Hence we have only one possible case, realized e.g., by this pattern.

$$\cdots \to\leftarrow\to\leftarrow\to\leftarrow\to\leftarrow \cdots .$$

5.3 Frieze groups

We first have to classify all finite subgroups of $O(2)$.

Exercise 5.1. Show that each finite subgroup of $O(2)$ is conjugate to one of the following (warning: bad notation):

- C_n, $n = 1, 2, \dots$, in particular $C_1 = 1$ is the trivial subgroup and $C_2 = \{\pm 1\}$.

- $C_{nv} \cong D_n$ generated by an n-fold rotation and a reflection at the x axis. In particular, C_{1v} has two elements.

Any 1d lattice in $\mathbb{R}^2$ has the form $\mathbb{Z}v$. In fact we may assume v is a standard basis vector pointing along the x-axis by a rotation and scaling. We want v to be left invariant by G_o. It is easy to see that this excludes all cases except those for $n = 1, 2$ above, i.e., C_1, C_2, D_1, D_2. However, mind that in case D_1 there are two possibilities: the reflection may be orthogonal to the line through v or along the line through v, leaving 5 cases to consider.

- C_1: Trivial case, only translational symmetry, $G \cong \mathbb{Z}$.

- C_2: By considerations as in the (1,1) case, there is only one case, G being generated by $(0, -\mathbb{1})$ and v.

- D_1, case (a): The additional generator of G_o is a reflection σ at the line through v. We want to find $\tau \in \mathbb{R}^2$ such that G is generated by v and (τ, σ). By conjugating with an overall translation t we may change $\tau \to \tau + (\mathbb{1} - \sigma)t$. This allows one to set $\tau = \lambda v$. Since τ is important only modulo G_t we may assume $0 \leq \lambda < 1$. The condition

$$(\tau, \sigma)^2 = (\tau + \sigma\tau, \mathbb{1}) = (2\lambda v, \mathbb{1}) \in G_t$$

  leaves us with 2 cases: $\lambda = 0$ and $\lambda = \frac{1}{2}$.

- D_1, case (b): Same as before, but now v is orthogonal to the reflection axis of σ. Overall translation allows us to change $\tau \to \tau + (\mathbb{1} - \sigma)t$. Hence we may assume $\tau \perp v$. But then our equation reads:

$$(\tau, \sigma)^2 = (\tau + \sigma\tau, \mathbb{1}) = (2\tau, \mathbb{1}) \in G_t$$

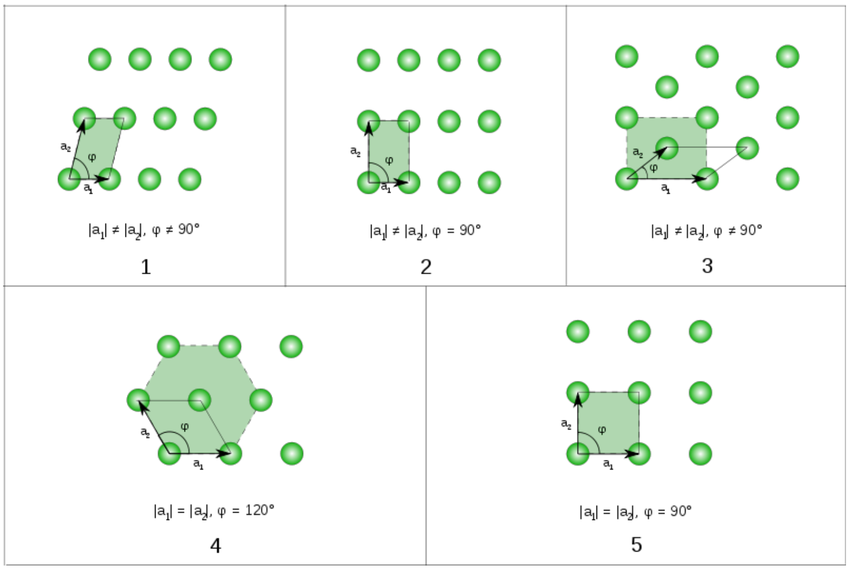
  which can only be satisfied for $\tau = 0$.

- D_2: D_2 is generated by two reflections σ_1, σ_2 orthogonal to and along v. We need to find τ_1, τ_2. Overall translation by t changes them to $\tau_j + (\mathbb{1} - \sigma_j)t$. Hence we may assume that $\tau_2 = \lambda v$ and $\tau_1 \perp v$. As before, we find that $\tau_1 = 0$ and $\lambda \in \{0, \frac{1}{2}\}$. This yields two cases in total.

So there are 7 frieze groups in total.

5.4 Wallpaper groups

One can consider the $(2, 2)$ case along the same lines. The classification of point groups has been done above. The classification of possible lattices (i.e., possible G_t with a G_o action) yields 5 possible cases, the Bravais lattices in two dimensions:



1. Square lattice.
2. Hexagonal lattice.
3. Rectangular lattice.
4. Centered rectangular lattice.
5. Obligue lattice.

By a similar but more tedious analysis one can show that there are 17 wallpaper groups.

5.5 The list of finite subgroups of $SO(3)$ and $O(3)$

Let us introduce several finite subgroups of $SO(3)$. We shall see below that this list is complete.

- The cyclic group C_n is generated by one rotation by $2\pi/n$. It has n elements.
- The dihedral group D_n is generated by one rotation by $2\pi/n$, and a π-rotation around an orthogonal axis. It has $2n$ elements.
- The rotations that leave fixed one of the five platonic solids form a finite subgroup of $SO(3)$ each. Since the cube and octahedron are dual to each other, as are the icosahedron and the dodecahedron, one obtains 3 distinct groups: T, O, I, with 12, 24 and 60 elements. (Polyhedral groups)

In addition to the above, there are a few more subgroups of $O(3)$.

- Groups C_{nv} obtained from C_n by adding an additional reflection plane parallel to the rotation axis.
- Groups C_{nh} obtained from C_n by adding an additional reflection plane orthogonal to the rotation axis.
- Groups D_{nh} obtained from D_n by adding an additional reflection plane orthogonal to the n-fold rotation axis.
- Groups D_{nd} that contain an additional rotation by π/n, followed by a reflection at the orthogonal plane to the rotation axis.
- Groups S_{2n}, that are generated by the an additional rotation by π/n, followed by a reflection at the orthogonal plane to the rotation axis.
- The polyhedral groups T_d (24), T_h (24), O_h (48), I_h (120). Here T_d, O_h, I_h are the full symmetry groups of the tetrahedron, octahedron and icosahedron. T_h (pyritohedral symmetry group) is the symmetry group of a volleyball.

Example 5.1. Note that some of the above groups are isomorphic as groups, but not as subgroups of $(S)O(3)$. Show in particular that except for the polyhedral groups all groups above are isomorphic to C_n, D_n or products thereof with C_2.

Remark 5.1. To emphasize whether a group is considered as a subgroup of $O(3)$ or just as a group crystallographers confusingly speak about "abstract groups" to emphasize they do not consider the group as subgroup of $O(3)$ but just as a group by itself.

5.6 Finite subgroups of $SO(3)$ and $O(3)$

Theorem 5.1. *Let $G \subset SO(3)$ be a non-trivial finite subgroup. Then G is isomorphic to one of C_n, D_n, T, O, I.*

Proof. Consider the set F of points on the sphere S^2 fixed by some non-trivial rotation in $SO(3)$. Each non-trivial rotation fixes exactly two points. Furthermore any given f is fixed by the n_f rotations around the axis described by f, $n_f - 1$ of which are non-trivial. Hence

$$2(|G| - 1) = \sum_{f \in F} (n_f - 1).$$

G acts on F. The stabilizer of any $f \in F$ is composed of the n_f rotations around this axis. Hence, by the Orbit-Stabilizer Theorem:

$$2(|G|-1) = \sum_{f \in F}(n_f - 1) = \sum_{[f] \in F/G} |G \cdot f|(n_f - 1) = \sum_{[f] \in F/G} \frac{|G|}{|\mathrm{Stab}_f|}(n_f - 1) = \sum_{[f] \in F/G} |G|(1 - \frac{1}{n_f}).$$

Dividing by $|G|$ yields

$$\sum_{[f] \in F/G} (1 - \frac{1}{n_f}) = 1(1 - \frac{1}{|G|}) < 2.$$

Since $1 - \frac{1}{n_f} \geq \frac{1}{2}$, we learn that there can be at most 3 orbits, $|G/F| \leq 3$. Consider these three cases:

- 1 orbit: We would have $\frac{2}{|G|} - \frac{1}{n_f} = 1$, which is impossible since $|G| \geq 2$.
- 2 orbits, with n_1 and n_2-fold rotations: We find $\frac{2}{|G|} = \frac{1}{n_1} + \frac{1}{n_2}$. Since $n_1, n_2 \leq |G|$, this is only possible for $n_1 = n_2 = |G|$, so that all elements in G are rotations around the same axis. Hence $G \cong C_n$, with $n = |G|$.
- 3 orbits, with n_1, n_2 and n_3-fold rotations: We find

$$\frac{2}{|G|} + 1 = \frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3}. \tag{5.1}$$

  Assume w.l.o.g. $n_1 \leq n_2 \leq n_3$.

  - If $n_1 = n_2 = 2$, then $n_3 = |G|/2$. By the orbit stabilizer Theorem the orbit of the n_3-fold axis fixed point must have size 2, i.e., correspond to a single axis, say the z-axis. Hence the other rotations must leave that axis fixed, and hence be π-rotations around orthogonal axis. Let α be an angle between two such orthogonal rotation axis, and let τ_1, τ_2 be the corresponding π-rotations. Then $\tau_1\tau_2$ is a rotation around the z-axis by an angle 2α. Hence 2α must be a multiple of $2\pi/n_3$. The only possibility is then to have $2n$ orthogonal axis, at angles π/n. Hence we get the dihedral group D_n.
  - If $n_1 = 2$, $n_2 = 3$, then n_3 is restricted to either of $3, 4, 5$. Consider $n_3 = 3$ first. Here $|G| = 12$, and the orbits have sizes 6, 4, 4. In particular, there are 4 three-fold rotation axis in one orbit. The orbit has to be preserved under all rotations around axis in this orbit. The only possible configuration is that the axis form the vertices of a tetrahedron (exercise). The tetrahedron has exactly 6 two-fold symmetry axis, hence the location of the two-fold rotation axis is fixed. But the rotations found thus far generate the tetrahedral group T, of order 12, and hence $G = T$.

- $n_1 = 2$, $n_2 = 3$, $n_3 = 4$, $|G| = 24$: The sizes of the orbits are $12, 8, 6$. The only possible locations of the six four-fold rotation (half-)axis are on the faces of a cube (exercise). The cube has exactly an additional 12 two-fold half-axis and 4 three-fold half axis, hence these must be the locations of our 2- and 3-fold axes, and $G = O$
- $n_1 = 2$, $n_2 = 3$, $n_3 = 5$, $|G| = 60$: The orbits have sizes $30, 20, 12$. The only possibility to arrange 12 5-fold axis is on the vertices of an icosahedron (exercise). By the same argument as before, $G = I$.
- If $n_1 = 2$, $n_2 \geq 4$, the equation (5.1) cannot be satisfied.
- If $n_1 \geq 3$, the equation (5.1) cannot be satisfied.

□

Theorem 5.2. *Let G be a finite subgroup of $O(3)$. Then G is isomorphic to one group in the following list: C_n, C_{nh}, C_{nv}, D_n, D_{nh}, D_{nd}, S_{2n}, T, T_d, T_h, O, O_h, I, I_h.*

Proof. Consider the normal subgroup $G_0 := G \cap SO(3) \subset G$. It is finite and hence must be one of the groups listed in the previous Theorem. We treat each case separately:

- If G_0 is trivial, then either G is trivial or $G = C_{1h}$ or $G = S_2$ (exercise).
- $G_0 = C_n$ $(n \geq 2)$: Additional rotoreflection must fix the single rotation axis and hence must reflect at a plane either containing or orthogonal to the axis. If containing, it must be a pure reflection. If only containing, one finds C_{nv}. If only orthogonal, the rotoreflections can have angle α such that $2\alpha = kn/2\pi$. Either all $\alpha = kn/2\pi$ (we get C_{nh} or $\alpha = kn/4\pi$ (we get S_{2n}). If there are both kinds of rotoreflections (containing and orthogonal), we can make another pure rotation with different axis, a contradiction.
- $G_0 = D_n$: One has the same two possibilities as before for rotoreflections (containing/orthogonal). If only containing, we get D_{nv}. If there are orthogonal rotoreflections, we have two options: The rotoreflections all have angles $\alpha = kn/2\pi$ (we get D_{nh}) or $\alpha = kn/4\pi$. In the latter case we can also build containing reflections and get D_{nd}.
- $G_0 = T$: 4 axis in tetrahedral configuration have 6 pure reflection symmetry planes, each containing 2 axis, and 3 rotoreflection planes (the rotation angle being π). In either case, one of the (roto-) reflections generates the others, together with the tetrahedral rotations. This leaves 4 possibilities: (i) no reflections (we get T), (ii) only reflections (T_d), (iii) only rotoreflections (pyrotahedral symmetry T_h) or (iv) both. If both, combining a rotoreflection with a reflection yields a new rotation, i.e., a contradiction. (check)
- $G_0 = O$: 3 axis in octahedral configuration have 3 reflection symmetry planes each containing 2 axes, 6 planes containing 1 axis each, and four $\pi/6$-rotoreflection planes containing none of the axis. Each class generates the other two together with the O operations. (check) Hence we can only get one new group, O_h.
- $G_0 = I$: check that each rotoreflection leaving invariant the configuration of 12 five-fold axis generates, together with I, the full icosahedral group I_h.

□

Exercise 5.2. Among these groups, list all that do not have rotation axes other than $2, 3, 4, 6$-fold. You should find 32 groups, the *cristallographic point groups*.

Exercise 5.3. Show that $T_d \cong O \cong S_4$.

Exercise 5.4. List all inclusions between the crystallographic point groups. I.e., draw a graph with 32 vertices, and an arrow from G to H if $G \supset H$ and there is no other point group between G and H.

5.7 The 14 Bravais lattices in 3D

Theorem 5.3 (Crystallographic exclusion principle). *Suppose a 3 dimensional lattice* Λ *is preserved by* $R \in SO(3)$. *Then the angle of rotation of* R *is one of* $0, \pi, \frac{2\pi}{3}, \frac{2\pi}{4}, \frac{2\pi}{6}$.

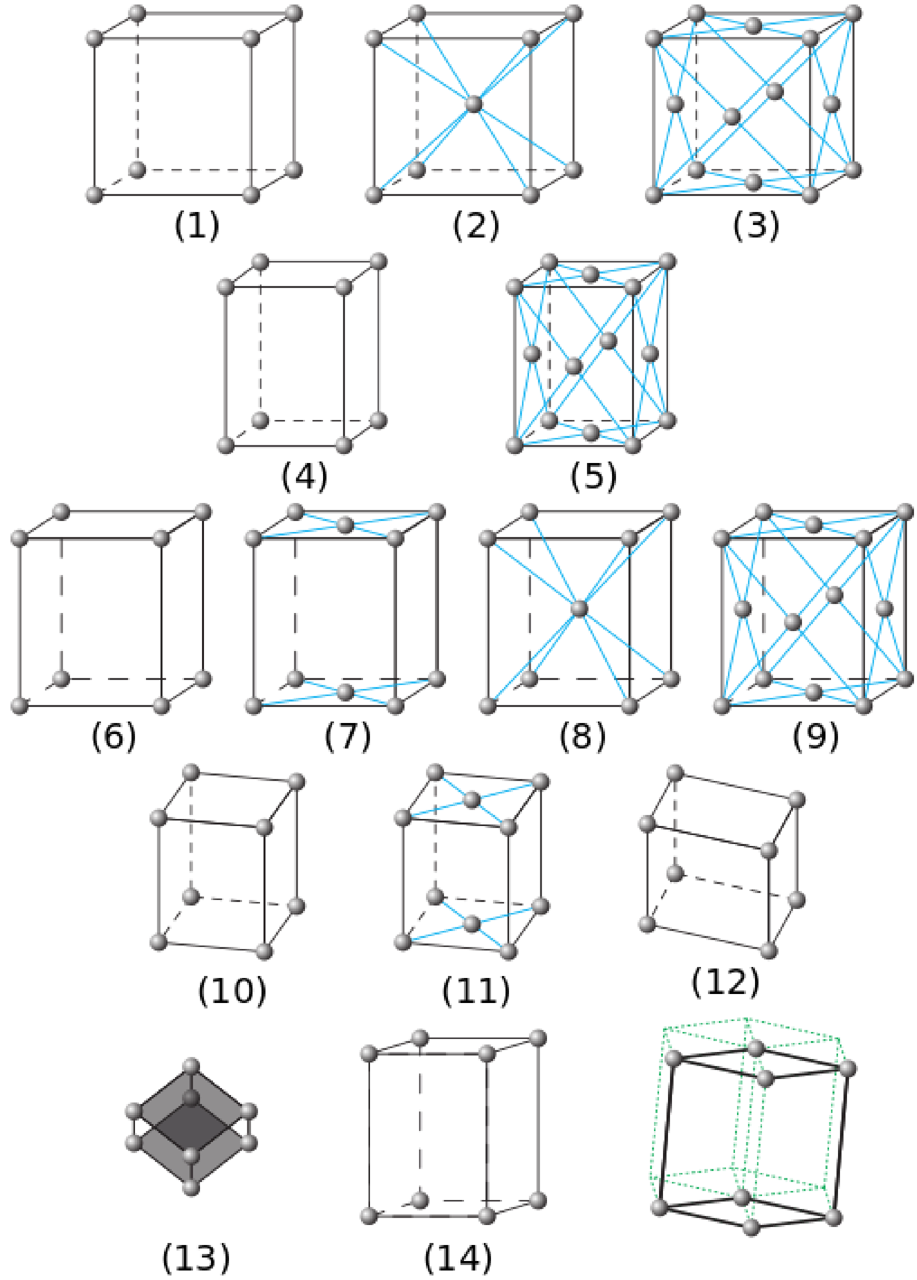
Proof. If $\mathbb{R}$ has rotation angle 0 we are done. Otherwise denote the rotation axis n. There is a lattice vector $v \in \Lambda$ not on the rotation axis. Hence $v \pm Rv$ is in the lattice and orthogonal to n. So we may assume that v is orthogonal to n. We may also assume that v is of smallest norm among such (non-zero) orthogonal vectors. If in the subgroup $\{R^k \mid k \in \mathbb{Z}\}$ there are non-trivial rotations R' of smaller angle than $\frac{2\pi}{6}$, then $v - R'v$ is orthogonal to n and of smaller norm than v, a contradiction. This handles all cases except the one that the smallest angle of rotation is $\frac{2\pi}{5}$, so assume R has that rotation angle. Then $R^2v - Rv + v \in \Lambda$ is not equal to zero, orthogonal to n and of smaller norm than v (see picture), a contradiction. $\square$

In particular, the icosahedral group is excluded.

Exercise 5.5. Let $G \subset SO(3)$ ($G \subset O(3)$) be a subgroup fixing a 3 dimensional lattice Λ. Show that G is finite.

Solution. $G \subset SO(3)$: Either there are infinite(-fold) rotation axis, or there are infinitely many rotation axis. In the first case we have a contradiction to the Theorem. In the second case, we may pick two rotation axis arbitrarily close to each other by compactness of the sphere. If one of the rotation axis is not $2, 3, 4$ or 6-fold we have a contradiction. Otherwise the composition of the two rotations is not a $2, 3, 4$ or 6-fold rotation if the axes are picked close enough. (Contradiction)

One may show that the possible lattices (that carry some G_o action) fall in one of the following 14 classes, the Bravais lattices.



They support the following symmetries:

Cubic	O, O_h, T, T_H, T_d
Tetragonal	$D_{4h}, D_4, D_{2d}, C_{4v}, S_4, C_4$
Orthorhombic	D_{2h}, D_2, C_{2v}
Hexagonal	$D_{6h}, D_6, D_{3h}, C_{6v}, C_{3h}, C_6$
Trigonal	$D_{3d}, D_3, C_{3v}, S_6, C_3$
Monoclinic	C_{2h}, C_2, C_{1h}
Triclinic	C_1, S_2

5.8 Computing extensions, and the classification of space groups

The final problem is to compute all $G \subset \mathrm{Euc}(\mathbb{R}^3)$ that fit into the exact sequence.

$$1 \to \Lambda \to G \to G_r \to 1$$

We will assume here that an explicit presentation of G_r is known, with generators $g_1, \ldots g_r$ and relations $R_1, \ldots R_k$. In fact we may have $r \leq 3$, cf. table **??**. Now each g_j has a unique pre-image in G such that the corresponding translation vector t_j is in the unit cell (wrt. Λ). Conversely, G is generated by the $g_j + t_j$, together with Λ. Hence our task is to enumerate all possible tuples $(t_1, \ldots, t_r)$, with all t_j's in the unit cell. Let R_i be a relation and for fixed t_j's let R'_i be the element of G obtained by replacing each g_j by $g_j + t_j$. Then R'_i projects to the unit in $O(3)$ and hence we must have $R'_i \in \Lambda$. Conversely, if $R'_i \in \Lambda$ for all i, then the group G generated by the $g_j + t_j$ really fits into the above exact sequence (exercise). So have reduced the task to solving the set of equations

$$R'_i \in \Lambda, \quad i = 1, 2, \ldots, k \tag{5.2}$$

for $t_1, \ldots, t_r$ in the unit cell. Note also that R'_i is linear in the $t_1, \ldots, t_r$. The equations may be solved either by hand, or on the computer.

Exercise 5.6. Devise an algorithm to compute all possible solutions $t_1, \ldots, t_r$ in the unit cell.

One obtains 230 cases, listed in table **??**.

5.8.1 Example on how to find the table entries

Let us consider the example of .... and solve the equations (5.2).

Exercise 5.7. Do the same for the ... entry of the table.

5.9 Special case: wallpaper groups

Bibliography