Weyl's character formula

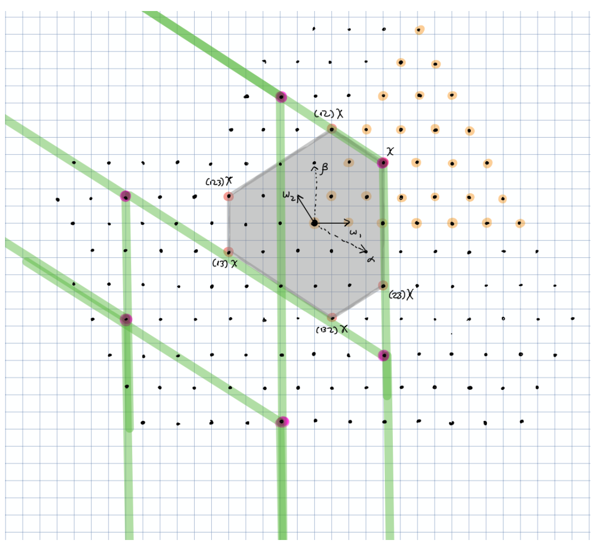
(notes for course taught at HUJI, Fall 2021-2022)

(UNPOLISHED DRAFT)

Alexander Yom Din

July 10, 2024

> Symmetry, as wide or as narrow as you may define its meaning, is one idea by which man through the ages has tried to comprehend and create order, beauty, and perfection.
>
> H. Weyl, *Symmetry*



Contents

1 Sources **3**

2 Topological groups, actions and representations **3**
- 2.1 Topological groups 3
- 2.2 Actions 4
- 2.3 Harmonic analysis 7
- 2.4 Representations 10

3 Basic representation theory of compact groups **14**
- 3.1 Haar measure 14
- 3.2 Complete reducibility, Schur's lemma, multiplicities 16
- 3.3 Character 18
- 3.4 Character in the case of SU(n) 21
- 3.5 Example: Irreps of SU(2) 23
- 3.6 A glimpse at Weyl's integration formula 24
- 3.7 Classification of irreps of SU(n) and Weyl's character formula (without proofs) 27
- 3.8 Some illustrations for Weyl's character formula for SU(3) 30
- 3.9 Some further notes regarding Weyl's character formula 33

4 Manifolds, Lie groups and Lie algebras **35**
- 4.1 Manifolds 35
- 4.2 Tangent spaces and tangent maps 37
- 4.3 Vector fields and flows 39
- 4.4 Lie groups 40
- 4.5 The exponential map 41
- 4.6 The Lie algebra of a Lie group 43
- 4.7 Proof of the Jacobi identity 47
- 4.8 Proof of Theorem 4.32 49
- 4.9 Some of Lie's theorems 51
- 4.10 Representations of Lie groups versus representations of Lie algebras 51
- 4.11 The case of SU(n) 54

5 Representation theory of $\mathfrak{sl}_2$ **55**
- 5.1 Finite-dimensional irreducible modules 56
- 5.2 Detour 1 - tensor products 58
- 5.3 Detour 2 - the Casimir element 59
- 5.4 Complete reducibility of finite-dimensional modules 61

6 The universal enveloping algebra **63**
- 6.1 Algebras and modules 63
- 6.2 The universal eneveloping algebra 64
- 6.3 The universal enveloping algebra as a "deformation" 67

7 Representation theory of $\mathfrak{sl}_n$ **69**
7.1 Weights 69
7.2 Roots 69
7.3 Highest weights 71
7.4 Irreducible highest weight modules 73
7.5 Verma modules 73
7.6 Irreducible highest weight modules - existence 75
7.7 When is L_λ finite-dimensional? 76

8 Formal character **79**
8.1 Convolution 79
8.2 Formal character 80
8.3 Expressing the formal character of an irreducible module in terms of formal characters of Verma modules, given a fact 82
8.4 Proof of Weyl's character formula (given the fact) 84
8.5 The center of the universal enveloping algebra 85

1 Sources

Some of the sources:

- "Complex Semisimple Lie Algebras" by J. P. Serre
- "Lectures on Lie Algebras" by J. Bernstein

2 Topological groups, actions and representations

2.1 Topological groups

Definition 2.1.

- A **topological group** is a set G equipped with both a group structure and a topology, such that the multiplication map $G \times G \to G$ and the inverse map $G \to G$ are continuous.
- Given topological groups G and H, a **morphism of topological groups** from G to H is a map $\phi : G \to H$ which is both continuous and a group homomorphism.

Example 2.2. *Here are some examples of topological groups.*

- *Any group, given the discrete topology, becomes a topological group.*
- $\mathbb{R}$*, with the group operation of addition and its standard topology, is a topological group. So is* $\mathbb{C}$*. Another example is* $\mathbb{Q}_p$*, the additive group of the field of* p*-adic numbers.*

- *We have the topological groups* $\mathbb{R}^\times$, $\mathbb{C}^\times$ *and* $\mathbb{Q}_p^\times$ *- the multiplicative groups of the fields, i.e. the sets of non-zero elements, with group operation being multiplication and the topology on* $F^\times$ *being inherited from* F*, the former being an open subset in the latter.*

- *The group* $\mathrm{GL}_n(\mathbb{R})$ *of invertible matrices over* $\mathbb{R}$ *of order* n*, with the operation of multiplication of matrices and the topology inherited to it as an open subset of the* $\mathbb{R}$*-vector space* $M_n(\mathbb{R})$*. Again, we have also* $\mathrm{GL}_n(\mathbb{C})$ *and* $\mathrm{GL}_n(\mathbb{Q}_p)$*.*

- *The next example is essentially the same as previous one. If we have a finite-dimensional vector space* V *over* $\mathbb{R}$*, we have the topological group* $\mathrm{GL}(V) = \mathrm{GL}_{\mathbb{R}}(V)$ *of invertible* $\mathbb{R}$*-linear transformations from* V *to* V*, with the operation of composition. Similarly, for vector spaces over* $\mathbb{C}$ *or* $\mathbb{Q}_p$ *(the reader is welcome to describe the topology).*

- *Various closed subgroups*[1] *of* $\mathrm{GL}_n(\mathbb{R})$, $\mathrm{GL}_n(\mathbb{C})$ *and* $\mathrm{GL}_n(\mathbb{Q}_p)$*. For example, the subgroups* $\mathrm{SL}_n(F) \subset \mathrm{GL}_n(F)$ *consisting of matrices of determinant* 1*. Or, the subgroups* $\mathrm{O}(n) \subset \mathrm{GL}_n(\mathbb{R})$ *and* $\mathrm{U}(n) \subset \mathrm{GL}_n(\mathbb{C})$ *consisting of orthogonal, respectively unitary, matrices. We also have* $\mathrm{SO}(n) := \mathrm{O}(n) \cap \mathrm{SL}_n(\mathbb{R})$ *and* $\mathrm{SU}(n) := \mathrm{U}(n) \cap \mathrm{SL}_n(\mathbb{C})$*.*

- *There are also natural topological groups which are not locally compact. For example, Given a topological group* G *we can consider the topological space of continuous maps* $\mathrm{Map}(S^1, G)$ *from the circle* S^1 *to* G*, equipped with the compact-open topology, and define the group operation pointwise. This is a "loop group".*

Remark 2.3. Except those of the last items, all the groups in Example 2.2 are locally compact.

Exercise 2.1. *Show that the topological groups* $\mathrm{O}(n)$ *and* $\mathrm{U}(n)$ *are compact.*

2.2 Actions

Definition 2.4. Let G be a topological group.

- Let X be a set. An **abstract G-action** on X is a map (with no requirement of continuity what-so-ever) $a : G \times X \to X$ satisfying
  (1) $a(1_G, x) = x$ for all $x \in X$.
  (2) $a(g_1, a(g_2, x)) = a(g_1 g_2, x)$ for all $g_1, g_2 \in G$ and $x \in X$.

- Let X be a topological space. A **G-action** on X is an abstract action $a : G \times X \to X$ which is continuous.

[1] Although any subgroup of a topological group becomes itself a topological group with the subspace topology, it is most natural to look at closed subgroups, because those are the subgroups for which the quotient G/H will be a T_1-space.

- A G**-space** is a topological space X equipped with a G-action.

  Given a G-space X, we almost always keep the action map a implicit, writing gx or $g \cdot x$ instead of $a(g,x)$[2].

- Given G-spaces X and Y, a **morphism of** G**-spaces** from X to Y is a map $\phi : X \to Y$ satisfying $\phi(gx) = g\phi(x)$ for all $g \in G$ and $x \in X$.

Example 2.5.

1. *Let G be a topological group. There are three strandard actions of G on itself. The **left regular action** is given by $a(g,g') := gg'$. The **right regular action** is given by $a(g,g') := g'g^{-1}$. The **conjugation action** is given by $a(g,g') := gg'g^{-1}$.*

2. *Let G be a topological group and let $H \subset G$ be a closed subgroup. We have a canonical surjective map $\pi : G \to G/H$ (sending g to gH) and we give G/H the corresponding quotient topology, i.e. $U \subset G/H$ is defined to be open if $\pi^{-1}(U) \subset G$ is open. Then we make G/H a G-space by setting $a(g, g'H) := gg'H$.*

3. *We have the **standard action** of $\mathrm{GL}_n(\mathbb{R})$ on $\mathbb{R}^n$ given by multiplying a vector by a matrix.*

4. *Consider $S^{n-1} \subset \mathbb{R}^n$, the closed subspace consisting of vectors of length 1 with respect to the standard inner product (the "unit sphere"). Then $\mathrm{SO}(n)$ acts on S^{n-1} by multiplying a vector by a matrix.*

5. *Let $\mathbb{H} \subset \mathbb{C}$ consist of complex numbers z for which $\mathrm{Im}(z) > 0$ (the "upper half plane"). Let $G := SL_2(\mathbb{R})$. We have an action of G on $\mathbb{H}$ ("by Möbius transformations"), given by setting*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z := \frac{az+b}{cz+d}.$$

6. *Consider the symmetric $\mathbb{R}$-bilinear form on $\mathbb{R}^4$ given by*

$$Q((x_1,x_2,x_3,t),(x_1',x_2',x_3',t')) := x_1x_1' + x_2x_2' + x_3x_3' - tt'$$

   (appearing in special relativity). Consider the "light cone" $X := \{v \in \mathbb{R}^4 \mid Q(v,v) = 0\}$. Consider the closed subgroup $\mathrm{SO}(3,1) \subset \mathrm{GL}_4(\mathbb{R})$ consisting of matrices A which preserve Q, i.e. which satisfy $Q(Av,Aw) = Q(v,w)$ for all $v,w \in \mathbb{R}^4$. Then we have an action of $\mathrm{SO}(3,1)$ on X by multiplying a vector by a matrix.

Definition 2.6. Let G be a topological group and let X be a G-space.

[2] In the same way as when given a group G, we keep the multiplication map, say m, implicit, don't give it a name, and write g_1g_2 instead of $m(g_1,g_2)$.

- X is said to be **transitive** if X is non-empty and for every $x_1, x_2 \in X$ there exists $g \in G$ such that $gx_1 = x_2$.
- X is said to be **homogeneous** if it is isomorphic[3] to the G-space G/H for some closed subgroup $H \subset G$.

Clearly, a homogeneous G-space is transitive. To check whether the converse holds, let X be a transitive G-space. Choose some $x_0 \in X$. Denote

$$G_{x_0} := \{g \in G \mid gx_0 = x_0\} \subset G.$$

Then G_{x_0} is a closed subgroup of G, called the **stabilizer** of x_0. We have a map

$$\phi : G/G_{x_0} \to X, \quad gG_{x_0} \mapsto gx_0.$$

Check, that ϕ is a morphism of G-spaces. Check, that ϕ is bijective. Thus, the only problem that might be is that ϕ is not a homeomorphism, i.e. that the inverse of ϕ is not continuous. This is equivalent to the map $G \to X$ given by $g \mapsto gx_0$ not being an open map.

Lemma 2.7. *Let G be a topological group and let X be a G-space. Suppose that G and X are locally compact and that G is separable[4]. Then if X is a transitive G-space it is also homogeneous.*

Proof. Fix $x_0 \in X$. As just explained, we want to check that the map $\psi : G \to X$ given by $g \mapsto gx_0$ is open. Let $U \subset G$ be a non-empty open subset. We want to see that $\psi(U) \subset X$ is an open subset. To that end, fix $u_0 \in U$, and we want to see that $\psi(u_0)$ is an interior point of $\psi(U)$. Translating everything by u_0^{-1}, we can assume without loss of generality that $1_G \in U$ and $u_0 = 1_G$. Let us pick a compact neighbourhood of 1_G lying in U, call it $1_G \in V \subset U$, such that $V^{-1} \cdot V \subset U$. It is enough to show that $\psi(V)$ contains some interior point. Indeed, if $v \in V$ is such that $\psi(v)$ is an interior point of $\psi(V)$, $\psi(1_G) = v^{-1}\psi(v)$ will be an interior point of $v^{-1}\psi(V) = \psi(v^{-1}V)$, and since $v^{-1}V \subset U$, $\psi(1_G)$ will also be an interior point of $\psi(U)$, as desired. Thus, we want to see that $\psi(V)$ contains some interior point. Since G is separable, we can find a countable subset $\{g_i\} \subset G$ which is dense in G. Then it is immediate to see that $\cup_i g_i V = G$. Hence $\cup_i g_i \psi(V) = X$. By Baire's category theorem[5], for some i the subset $g_i\psi(V)$ of X has an interior point. Translating, the subset $\psi(V)$ of X has an interior point, as desired. $\square$

In our practice we will only deal with second countable[6] locally compact spaces, and hence the last lemma shows that there is no difference between transitive and homogeneous G-spaces.

[3]An **isomorphism** is a morphism which admits an inverse morphism - so we can speak of an isomorphism of topological groups, an isomorphism of G-spaces, etc.

[4]A topological space is **separable** if it contains a countable subset which is dense in it.

[5]**Baire's category theorem** says, in particular, that if a locally compact space is presented as a countable union of closed subsets, then one of these closed subsets has an interior point.

[6]A topological space is **second countable** if it has a countable base for the topology. Second countable topological spaces are separable.

Exercise 2.2. *Check which of the examples in Example 2.5 are homogeneous.*

2.3 Harmonic analysis

Let X be a topological space. We can ask a basic question in harmonic analysis: How to study a, say continuous, function $f : X \to \mathbb{C}$? The basic idea is that we want a systematic way of writing such an f as some (infinite) sum of "simple" functions which we can understand. When having a G-action on X, the basic idea is that those "simple" functions should be functions that "transform simply under the G-action". What does it mean more precisely?

First, denoting by $C(X)$ the $\mathbb{C}$-vector space of continuous functions from X to $\mathbb{C}$, let us notice that we have an abstract[7] action of G on $C(X)$: Given $g \in G$ and $f \in C(X)$, we set gf to be the function sending x to $f(g^{-1}x)$, i.e. we set $(gf)(x) := f(g^{-1}x)$.

The simplest behaviour is of being G**-invariant**: A function $f \in C(X)$ is G-invariant if $gf = f$ for all $g \in G$, i.e. $f(g^{-1}x) = f(x)$ for all $g \in G$ and $x \in X$. In other words, given $g \in G$ denote by $T_g : C(X) \to C(X)$ the linear operator given by $f \mapsto gf$. Then f is G-invariant if it is an eigenvector of all the operators T_g, with eigenvalue 1.

A generalization is as follows. Let $\chi : G \to \mathbb{C}^\times$ be a function. A function $f \in C(X)$ is χ**-equivariant**, or a G**-eigenfunction with eigencharacter** χ, if $gf = \chi(g) \cdot f$ for all $g \in G$, i.e. $f(g^{-1}x) = \chi(g) \cdot f(x)$ for all $g \in G$ and $x \in X$. In other words, f is χ-equivariant if, for every $g \in G$, f is an eigenvector of T_g with eigenvalue $\chi(g)$. We have the following exercise:

Exercise 2.3. *If $f \neq 0$, then χ is in fact a morphism of topological groups.*

Because of the exercise, we only consider χ's which are morphisms of topological groups.

Definition 2.8. Let G be a topological group. A **quasi-character** of G is a morphism of topological groups $G \to \mathbb{C}^\times$. A **character** of G is a morphism of topological groups $G \to \mathbb{C}^\times_{|-|=1}$. Let us denote by $\mathrm{qCh}(G)$ (resp. $\mathrm{Ch}(G)$) the abelian group of quasi-characters of G (resp. characters of G), where the group operation is pointwise multiplication. So $\mathrm{Ch}(G)$ is a subgroup of $\mathrm{qCh}(G)$.

Exercise 2.4. *Let G be a topological group. Show that if G is compact, then every quasi-character of G is a character of G.*

Remark 2.9. Often in harmonic analysis one is interested in χ-eigenfunctions only when χ is a character, rather than merely a quasi-character. But, this is not always to case, for example the Laplace transform involves also quasi-characters. Anyway, our focus in this course will be compact groups, for which, in view of Exercise 2.4, there is no difference.

[7]One can give $C(X)$ a topology so that this abstract action will be an action, i.e. it will be continuous, but we don't want to discuss this now.

Example 2.10. *Let us illustrate. Consider the unit circle $S^1 \subset \mathbb{R}^2$, and the action of $SO(2)$ on it, multiplying a vector by a matrix, as in one of the examples above. Here, it is convenient to identify both S^1 and $SO(2)$ with $\mathbb{R}/\mathbb{Z}$ as follows. Notice that $\mathbb{R}/\mathbb{Z}$ is a topological group naturally ($\mathbb{R}$ is a topological group with respect to addition, and the quotient by the closed normal subgroup $\mathbb{Z}$ is again a topological group naturally). Let us abbreviate $[x] := x + \mathbb{Z}$. We have an isomorphism of topological groups*

$$\mathbb{R}/\mathbb{Z} \xrightarrow{\sim} SO(2)$$

given by

$$[x] \mapsto \begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix}.$$

We have an isomorphism of topological spaces

$$\mathbb{R}/\mathbb{Z} \xrightarrow{\sim} S^1$$

given by

$$[x] \mapsto \begin{pmatrix} \cos x \\ \sin x \end{pmatrix}.$$

Under these isomorphisms, our action becomes the action of $\mathbb{R}/\mathbb{Z}$ on $\mathbb{R}/\mathbb{Z}$ given by $a([x],[y]) := [x+y]$, i.e. simply translation (or what we called the regular (left or right) action). What characters of $\mathbb{R}/\mathbb{Z}$ do we have? Given $n \in \mathbb{Z}$, we have $\chi_n \in \mathrm{Ch}(\mathbb{R}/\mathbb{Z})$ given by $\chi_n([x]) := e^{2\pi i n x}$. One checks that $\mathbb{Z} \to \mathrm{Ch}(\mathbb{R}/\mathbb{Z})$ given by $n \mapsto \chi_n$ is an isomorphism of abelian groups. Now, what are the eigenfunctions? One immediately sees that, setting $f_n := \chi_{-n}$, we have that f_n is χ_n-equivariant, i.e.

$$f_n([x] - [y]) = e^{2\pi i n y} f_n([x]), \quad \forall x, y \in \mathbb{R},$$

and all χ_n-equivariant functions are scalar multiples of f_n.

Let us continue with this example. As we said, the idea is that we want to write any function $f \in C(\mathbb{R}/\mathbb{Z})$ as an infinite sum of functions which behave simply under the translation action. So, we want to write any function $f \in C(\mathbb{R}/\mathbb{Z})$ as an infinite $\mathbb{C}$-linear combination of the f_n's. More explicitly, we can think of continuous functions on $\mathbb{R}/\mathbb{Z}$ as continuous functions on $\mathbb{R}$ which are periodic, with period 1. So given $f \in C(\mathbb{R})$ which is 1-periodic, we want to write (yet non-formally, heuristically)

$$f(x) = \sum_{n \in \mathbb{Z}} c_n \cdot e^{-2\pi i n x}.$$

This is the subject of classical Fourier theory - the subject of **Fourier series**. First, one needs to guess what should be the coefficients c_n. For this, we integrate (yet non-formally, heuristically):

$$\int_0^1 f(x) \cdot e^{2\pi i m x} \cdot dx = \sum_{n \in \mathbb{Z}} c_n \cdot \int_0^1 e^{-2\pi i n x} e^{2\pi i m x} \cdot dx = c_m.$$

Now, one can formulate various formal claims, for example:

Theorem 2.11. *Given $n \in \mathbb{Z}$, denote by $f_n \in C(\mathbb{R})$ the 1-periodic function $x \mapsto e^{-2\pi i n x}$. Let $f \in C(\mathbb{R})$ be 1-periodic and smooth. Define*

$$c_n := \int_0^1 f(x) \cdot f_{-n}(x) \cdot dx.$$

Then

$$f = \sum_{n \in \mathbb{Z}} c_n f_n$$

absolutely and uniformly.

Proof. Omitted. □

Remark 2.12. In this course, we concentrate on compact groups. Already now we can see how non-compact groups provide more complication. Namely, consider the group $\mathbb{R}$ instead of $\mathbb{R}/\mathbb{Z}$. Explicitly this has the meaning that we now consider continuous functions in $C(\mathbb{R})$ which are not necessarily 1-periodic. This time, we have $\mathbb{R} \xrightarrow{\sim} \mathrm{Ch}(\mathbb{R})$, given by $t \mapsto \chi_t$, where $\chi_t(x) := e^{2\pi i t x}$. So the "space of parameters" is now not discrete. Therefore, we will expect a general function to not decompose as an infinite sum of simple functions, but rather as an integral of simple functions. Namely, we have $f_t(x) := e^{-2\pi i t x}$ as before, but now we will want to write

$$f(x) = \int_{-\infty}^{+\infty} c_t \cdot e^{2\pi i t x} \cdot dt.$$

Remark 2.13. Complete harmonic analysis of functions on X in terms of G is, generally speaking, impossible, unless X is a homogeneous G-space. For example, imagine $\mathbb{R}$ acting on $\mathbb{R}^2$ by $x' \cdot (x, y) := (x + x', y)$. Functions which "transform simply" under the action in that case will be functions of the form $(x, y) \mapsto h(y) \cdot e^{-2\pi i u x}$ for $u \in \mathbb{R}$ and $g \in C(\mathbb{R})$, i.e. "in the y-direction" we are completely unrestrained. Thus, the action has not helped us to gain any simplification "in the y-direction". To gain simplification in "all directions", the action needs to be homogeneous[8].

Finally, let us illustrate how such harmonic analysis can be used.

Theorem 2.14. *Let $\alpha \in \mathbb{R}$ be an irrational number. Let $(a, b) \subset [0, 1]$ be a subinterval. Then*

$$\lim_{N \to \infty} \frac{1}{N} \cdot (\text{number of } 0 \le n \le N-1 \text{ for which } n\alpha \in (a, b) + \mathbb{Z}) = b - a.$$

In words, $\{[n\alpha]\}_{n \ge 0}$ is equidistributed in $\mathbb{R}/\mathbb{Z}$.

Proof. Notice that if consider a function f on $\mathbb{R}/\mathbb{Z}$ to be the characteristic function of (a, b), then the statement of the theorem is formulated as

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} f([n\alpha]) = \int_0^1 f([x]) dx. \tag{2.1}$$

[8] Or, maybe, just having a dense orbit also sometimes allows for "complete analysis".

Next, we would like to understand that it is enough to estbalish (2.1) for all smooth functions f on $\mathbb{R}/\mathbb{Z}$ - this is a small exercise, approximating the characteristic function by smooth functions from below and above. Now, notice that the equality (2.1) is stable under linear combinations and under passage to a limit of a uniformly convergent sequence. Hence, in view of Theorem 2.11, in order to establish (2.1) for all smooth functions f it is enough to establish it for the functions $f := f_m$, for $m \in \mathbb{Z}$. To that end, let us calculate:

$$\frac{1}{N}\sum_{n=0}^{N-1} f_m([n\alpha]) = \frac{1}{N}\sum_{n=0}^{N-1} e^{2\pi i m n\alpha} = \begin{cases} 1 & m = 0 \\ \frac{1}{N}\frac{e^{2\pi i m N\alpha}-1}{e^{2\pi i m\alpha}-1} & m \neq 0 \end{cases} \xrightarrow{N\to\infty}$$

$$\xrightarrow{N\to\infty} \begin{cases} 1 & m = 0 \\ 0 & m \neq 0 \end{cases}$$

and since we also have

$$\int_0^1 f_m(x)dx = \begin{cases} 1 & m = 0 \\ 0 & m \neq 0 \end{cases}$$

we are done. □

2.4 Representations

Representations naturally arise when we consider more complicated examples than the one above. Namely, let us consider SO(3) acting on S^2 as we had somewhere above. So, attempting to do as above, we again first ask about Ch(SO(3)). However, it turns out that $\mathrm{Ch}(\mathrm{SO}(3)) = \{1\}$ (this is an exercise whose solution we omit currently). The SO(3)-eigenfunctions with trivial eigencharacter are the constant functions. So, we need to somehow extend our understanding of "functions which transform simply under the action". A simple thing to notice is that if we consider functions in $C(S^2)$ which are of the form

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto ax + by + cz$$

for some $a, b, c \in \mathbb{C}$, those form a 3-dimensional $\mathbb{C}$-linear subspace of $C(S^2)$, call this subspace L, and we notice that L is SO(3)-invariant, i.e. $gf \in L$ whenever $g \in \mathrm{SO}(3)$ and $f \in L$. So we have now a hint regarding what to "transform simply" could mean more generally - a (non-zero) SO(3)-eigenfunction is simply a function spanning a SO(3)-invariant 1-dimensional subspace, and we can generalize, and search for SO(3)-invariant f.d. subspaces.

Thus, let us try to formulate abstractly what we are aiming at currently. Let G be a topological group and let X be a G-space. Let $L \subset C(X)$ be a f.d. $\mathbb{C}$-linear subspace which is ***G*-invariant**, i.e. $gf \in L$ whenever $g \in G$ and $f \in L$. We have an abstract action $a : G \times L \to L$ inherited from the abstract action of G on $C(X)$.

Exercise 2.5. *Let M be a f.d. $\mathbb{C}$-vector space. Choose an isomorphism of $\mathbb{C}$-vector spaces $M \cong \mathbb{C}^n$, and using it transport the standard topology of $\mathbb{C}^n$ to M. Show that the resulting topology on M does not depend on the choice of isomorphism. Thus a f.d. $\mathbb{C}$-vector space has a well-defined topology - we will always consider them with that topology.*

Exercise 2.6. *Show that the abstract action a above is, in fact, an action, i.e. it is continuous (where we have explained in Exercise 2.5 what is the topology to be taken on L).*

Furthermore, clearly this action a is $\mathbb{C}$-linear in the second variable, i.e. for any $g \in G$ the map $L \to L$ given by $v \mapsto a(g, v)$ is $\mathbb{C}$-linear. What replaces the eigenvalue prescription $\chi \in \mathrm{Ch}(G)$ in our current generalization is an "abstract model" for our L, i.e. a f.d. $\mathbb{C}$-vector space M, equipped with a G-action which is $\mathbb{C}$-linear in the second variable (and M has nothing to do with X - that is the meaning of the adjective "abstract"). So we define:

Definition 2.15. Let G be a topological group.

- Let V be a $\mathbb{C}$-vector space. A **$\mathbb{C}$-linear abstract G-action** on V is an abstract action $G \times V \to V$ which is $\mathbb{C}$-linear in the second variable.
- An **abstract G-representation** is a $\mathbb{C}$-vector space V equipped with a $\mathbb{C}$-linear abstract G-action.
- Let V be a f.d. $\mathbb{C}$-vector space. A **$\mathbb{C}$-linear G-action** on V is a G-action on V which is $\mathbb{C}$-linear as an abstract action, i.e. the abstract action $G \times V \to V$ should be continuous and $\mathbb{C}$-linear in the second variable.
- A **f.d. G-representation** is a f.d. $\mathbb{C}$-vector space M equipped with a $\mathbb{C}$-linear G-action.
- Let V_1 and V_2 be abstract G-representations. A **morphism of G-representations** from V_1 to V_2 is a $\mathbb{C}$-linear map $T : V_1 \to V_2$ satisfying $T(gv) = gT(v)$ for all $g \in G$ and $v \in V_1$. We denote by $\mathrm{Hom}_G(V_1, V_2)$ the $\mathbb{C}$-vector space of morphisms of G-representations from V_1 to V_2 (the structure of $\mathbb{C}$-vector space on this set is just by it being a $\mathbb{C}$-vector subspace of $\mathrm{Hom}_{\mathbb{C}}(V_1, V_2)$).

Thus, our L of before is a 3-dimensional $\mathrm{SO}(3)$-representation, which we have found inside $C(S^2)$, which itself is an abstract $\mathrm{SO}(3)$-representation.

Exercise 2.7. *Here is a basic reformulation of what a f.d. representation is. Let G be a topological group and let M be a f.d. $\mathbb{C}$-vector space. Show that the set of $\mathbb{C}$-linear G-actions on M is in bijection with the set of topological group morphisms $G \to \mathrm{GL}(M)$, by sending $a : G \times M \to M$ to the morphism $\rho : G \to \mathrm{GL}(M)$ defined by $\rho(g)(v) := a(g, v)$. We will swap freely between these two equivalent formulations.*

However, a f.d. G-representation is still not the precise generalization of a quasi-character of G. To explain this, let us consider the following simple notions:

Remark 2.16.

- Let V be a f.d. G-representation. Let $W \subset V$ be a G-invariant $\mathbb{C}$-linear subspace (i.e. for $g \in G$ and $v \in W$ we have $gv \in W$). Then W itself, with the G-action gotten by restriction of that on V, is a f.d. G-representation. For that reason, a G-invariant $\mathbb{C}$-linear subspace is also called a **G-subrepresentation**.

- Recall first some notions for vector spaces. There is the notion of an **(external) direct sum**: Let W and U be $\mathbb{C}$-vector spaces. Then we construct a new $\mathbb{C}$-vector space $W \oplus U$ as the Cartesian product of W and U, with addition and multiplication by scalar performed element-wise. There is also the notion of an **internal direct sum**: If V is a $\mathbb{C}$-vector space and $W, U \subset V$ are $\mathbb{C}$-vector subspaces, then V is said to be the direct sum of W and U if the $\mathbb{C}$-linear map $W \oplus U \to V$ given by $(w, u) \mapsto w + u$ is an isomorphism of $\mathbb{C}$-vector spaces. Equivalently, if $V = W + U$ and $W \cap U = \{0\}$. One then also writes $V = W \oplus U$ (causing a very slight abuse of notation).

- Let V and W be two f.d. G-representations. We construct a f.d. G-representation $V \oplus W$, called the **(external) direct sum** of V and W, as follows. As a $\mathbb{C}$-vector space it is the direct sum of V and W. The G-action is given by $g(v, w) := (gv, gw)$.

- Let V be a f.d. G-representation. Let $W, U \subset V$ be two G-subrepresentations such that $V = W \oplus U$ (internal direct sum). Then the (external) direct sum $W \oplus U$ of W and U is isomorphic as a G-representation to V via $(w, u) \mapsto w + u$.

We don't want to look for things isomorphic to a direct sum $M_1 \oplus M_2$ inside $C(X)$ - it is inefficient once we have already looked for things which are isomorphic to M_1 and things which are isomorphic to M_2. Thus, we in some sense want to only consider "smallest possible" G-representations. One arrives to the following definition:

Definition 2.17. Let G be a topological group and let M be a f.d. G-representation. We say that M is **irreducible** if $M \neq 0$ and the only G-subrepresentations of M are 0 and M. The term "irreducible representation" is often abbreviated as "irrep".

Remark 2.18. One can also define M to be **indecomposable** if given G-subrepresentations $M_1, M_2 \subset M$ such that $M = M_1 \oplus M_2$, one has either $M_1 = 0$ or $M_2 = 0$. Then clearly an irreducible representation is indecomposable. We will see later that if G is compact then the converse also holds.

We can now state a theorem for the action of $\mathrm{SO}(3)$ on S^2:

Theorem 2.19. *For every $n \in 2\mathbb{Z}_{\geq 0}+1$ there exists a unique $\mathrm{SO}(3)$-invariant n-dimensional subspace $L_n \subset C(S^2)$ which is irreducible as an $\mathrm{SO}(3)$-representation.*

Given a smooth[9] $f \in C(S^2)$ *there exists a unique collection* $(f_n)_{n \in 2\mathbb{Z}_{\geq 0}+1}$ *with* $f_n \in L_n$ *such that*

$$f = \sum_{n \in 2\mathbb{Z}_{\geq 0}+1} f_n$$

absolutely and uniformly.

Proof. Omitted. □

Definition 2.20. Let G be a topological group. We denote by $\mathrm{Irr}(G)$ the set of isomorphism classes of irreducible f.d. G-representations. Given an irreducible f.d. G-representation M, we denote by $[M] \in \mathrm{Irr}(G)$ the corresponding isomorphism class.

$\mathrm{Irr}(G)$ is our generalization of $\mathrm{qCh}(G)$ - this is what replaces eigencharacters. Given $[M] \in \mathrm{Irr}(G)$, one looks for G-invariant f.d. $\mathbb{C}$-linear subspaces $L \subset C(X)$ which are isomorphic, as G-representations, to M - this is what replaces eigenfunctions.

How is this related to the previous search for G-eigenfunctions? A quasi-character $\chi \in \mathrm{qCh}(G)$ gives rise to an irreducible 1-dimensional G-representation which we denote by $\mathbb{C}_\chi$. It is constructed as follows. As a $\mathbb{C}$-vector space, it is simply $\mathbb{C}$ itself. The G-action is given by $g \cdot c := \chi(g)c$. Now, a G-invariant $\mathbb{C}$-linear subspace $L \subset C(X)$ which is isomorphic to $\mathbb{C}_\chi$ as a G-representation is simply a 1-dimensional subspace all of whose vectors are G-eigenfunctions with eigencharacter χ (check this!).

Exercise 2.8. *Show that every* 1*-dimensional* G*-representation is irreducible, and is isomorphic to* $\mathbb{C}_\chi$ *for some quasi-character* $\chi \in \mathrm{qCh}(G)$*. Show also that given two quasi-characters* $\chi_1, \chi_2 \in \mathrm{qCh}(G)$ *such that* $\chi_1 \neq \chi_2$*, the* 1*-dimensional* G*-representations* $\mathbb{C}_{\chi_1}$ *and* $\mathbb{C}_{\chi_2}$ *are non-isomorphic. In other words, we have an injection* $\mathrm{qCh}(G) \hookrightarrow \mathrm{Irr}(G)$ *given by* $\chi \mapsto [\mathbb{C}_\chi]$*, whose image is the set of isomorphism classes of* 1*-dimensional* G*-representations.*

Example 2.21. *One can show that every irreducible f.d.* $\mathrm{SO}(3)$*-representation is isomorphic to* L_n *for some* $n \in 2\mathbb{Z}_{\geq 0} + 1$ *(in the notation of Theorem 2.19). Of course, this is a very nice situation - in general, there can be irreducible f.d.* G*-representations which do not appear in a certain* $C(X)$*, and there can be irreducible f.d.* G*-representations which appear in a certain* $C(X)$ *more than once.*

[9] One of the characterizations of a function $f \in C(S^2)$ being smooth is that for every $p \in S^2$ there exists an open $p \in U \subset \mathbb{R}^3$ and a smooth function $\widetilde{f} \in C(U)$ such that $\widetilde{f}|_{U \cap S^2} = f$.

3 Basic representation theory of compact groups

3.1 Haar measure

By a locally compact space we will always mean a second countable locally compact space[10]. By $C_c(X) \subset C(X)$ we denote the subspace of functions with compact support[11].

Definition 3.1. Let X be a locally compact topological space. A **signed Radon measure** on X is a functional $\int : C_c(X) \to \mathbb{C}$ with the following property:

- (continuity) Given a sequence $\{f_n\} \subset C_c(X)$ converging uniformly to $f \in C_c(X)$, such that there exists a compact subset $K \subset X$ with the property that $f_n|_{X \smallsetminus K} = 0$ for all n, the sequence $\{\int f_n\}$ converges to $\int f$.

The set of signed Radon measures is naturally a $\mathbb{C}$-vector space, and we denote it by $\mathcal{M}(X)$. We say that a signed Radon measure $\int : C_c(X) \to \mathbb{C}$ is a **Radon measure** if it satisfies in addition the following property:

- (positivity) Given $f \in C_c(X)$ such that $f(x) \geq 0$ for all $x \in X$, we have $\int f \geq 0$.

Example 3.2.

- *On $\mathbb{R}$ we have the Radon measure sending $f \in C_c(\mathbb{R})$ to the usual Riemann integral $\int_{-\infty}^{+\infty} f(x) \cdot dx$.*
- *Generalizing the previous example, given a continuous function $g \in C(\mathbb{R})$, we have on $\mathbb{R}$ the signed Radon measure sending $f \in C_c(\mathbb{R})$ to $\int_{-\infty}^{\infty} g(x) f(x) \cdot dx$. It is a Radon measure if and only if $g(x) \geq 0$ for all $x \in \mathbb{R}$.*
- *On $\mathbb{R}$ we have the Radon measure δ_0 (the **Dirac delta**) sending $f \in C_c(\mathbb{R})$ to $f(0)$.*
- *Given a (countable) set X, considering X as a discrete topological space we have the **counting** Radon measure on X given by $\int f := \sum_{x \in X} f(x)$.*

Exercise 3.1. *Let X be a locally compact space and V a f.d. $\mathbb{C}$-vector space. We denote by $C_c(X, V)$ the $\mathbb{C}$-vector space of continuous functions from X to V which have compact support. Let $\int \in \mathcal{M}(X)$. Then we can naturally define an "extension" of $\int$ to a $\mathbb{C}$-linear map $C_c(X, V) \to V$, which by abuse of notation we also call $\int$. It is the unique $\mathbb{C}$-linear map which sends any expression $\sum_{1 \leq i \leq n} f_i v_i$ (where $f_i \in C_c(X)$ and $v_i \in V$) to $\sum_{1 \leq i \leq n} \left(\int f_i\right) v_i$. We have some natural properties. For example, given a $\mathbb{C}$-linear operator $T : V \to W$ between two f.d. $\mathbb{C}$-vector spaces, we have $T\left(\int F\right) = \int (T \circ F)$ where $F \in C_c(X, V)$.*

[10] I assume second countability to be on the safe side and not think about technicalities (and, since most spaces in practice are second countable, this is not very restrictive).

[11] The support of a function $f \in C(X)$ is defined as the closure in X of the subset of X consisting of $x \in X$ for which $f(x) \neq 0$.

Let G be a topological group and let X be a G-space, which is locally compact. Recall that we have on the $\mathbb{C}$-vector space $C(X)$ the structure of an abstract G-representation - given $g \in G$ and $f \in C(X)$ we define $gf \in C(X)$ by $(gf)(x) := f(g^{-1}x)$. It is clear that $C_c(X) \subset C(X)$ is a G-subrepresentation. Now, on the $\mathbb{C}$-vector space $\mathcal{M}(X)$ of signed Radon measures we also have the structure of an abstract G-representation - given $g \in G$ and $\int \in \mathcal{M}(X)$ we define $g\int \in \mathcal{M}(X)$ by $(g\int)(f) := \int g^{-1}f$. In general, given an abstract G-representation V, a vector $v \in V$ is said to be **G-invariant** if $gv = v$ for all $g \in G$. So, we can speak about G-invariant signed Radon measures on X.

Remark-Notation 3.3. Let G be a topological group. Recall the **left regular action** of G on G given by $a(g, g') := gg'$ and the **right regular action** of G on G given by $a(g, g') := g'g^{-1}$. We correspondingly get two abstract $\mathbb{C}$-linear actions of G on $C_c(G)$ and on $\mathcal{M}(G)$, as described above. Given $g \in G$, we denote by $L_g : C_c(G) \to C_c(G)$ and $R_g : C_c(G) \to C_c(G)$ the corresponding $\mathbb{C}$-linear operators of acting by g, so concretely $(L_g f)(g') = f(g^{-1}g')$ and $(R_g f)(g') = f(gg')$. Also, we denote (slightly abusing notation) by $L_g : \mathcal{M}(G) \to \mathcal{M}(G)$ and $R_g : \mathcal{M}(G) \to \mathcal{M}(G)$ the corresponding $\mathbb{C}$-linear operators of acting by g, so $\left(L_g \int\right)(f) = \int L_{g^{-1}} f$ and $\left(R_g \int\right)(f) = \int R_{g^{-1}} f$.

Theorem-Definition 3.4 (Haar)**.** *Let G be a locally compact group.*

- *(existence) There exists a non-zero Radon measure $\int \in \mathcal{M}(X)$ which is right G-invariant (i.e. G-invariant w.r.t. the right regular action), i.e. satisfying $R_g \int = \int$ for all $g \in G$.*
- *(uniqueness) Any two non-zero right G-invariant signed Radon measures on X differ by a scalar.*

A right G-invariant non-zero Radon measure on X is called a ***right Haar measure****. Thus, any two right Haar measures differ by a scalar in $\mathbb{R}^{\times}_{>0}$.*

Proof. Omitted. □

Remark 3.5. Of course, by considering the left regular action, we analogously obtain the notion of a **left Haar measure**.

Exercise 3.2. *Let G be a discrete group. Show that the counting measure on G is both a left and a right Haar measure. In particular, for that class of groups there is a canonical choice for a right Haar measure (which in general is only defined up to a positive scalar).*

Example 3.6. *A Haar measure on $\mathbb{R}$ (clearly on an abelian group there is no difference between right and left Haar measures) is given by the usual Riemann integral $f \mapsto \int_{-\infty}^{+\infty} f(x) \cdot dx$.*

Exercise 3.3. *Let G be a locally compact group and let $\int$ be a right Haar measure on G. Let $f \in C_c(G)$ satisfy $f(g) \geq 0$ for all $g \in G$ and suppose that $f \neq 0$. Then $\int f > 0$.*

Remark 3.7. Let us see that right Haar measures coincide with left Haar measures for compact groups as well. Let $\int \in \mathcal{M}(G)$ be a right Haar measure. Let $g \in G$. One immediately sees that $L_g \int$ is again a right Haar measure. By the uniqueness of a right Haar measure, there exists $c \in \mathbb{R}_{>0}$ such that $L_g \int = c \cdot \int$. Notice that $(L_g \int)(1) = \int(L_{g^{-1}}(1)) = \int 1$ (where $1 \in C_c(G)$ is the function which is equal to 1 everywhere - it has compact support since G is compact!) and on the other hand $(L_g \int)(1) = (c \int)(1) = c \cdot \int 1$. Comparing, we obtain $c = 1$, and so $L_g \int = \int$. Since this holds for every $g \in G$, by definition $\int$ is a left Haar measure.

Let us here also notice that the right Haar measure $\int \in \mathcal{M}(G)$ for our compact group G can be always normalized so that $\int 1 = 1$. We can say that such a right Haar measure has **total mass** 1.

3.2 Complete reducibility, Schur's lemma, multiplicities

Definition 3.8. Let G be a topological group. Let V be a f.d. G-representation. An inner product $\langle -, - \rangle$ on V is said to be **G-invariant** if

$$\langle gv_1, gv_2 \rangle = \langle v_1, v_2 \rangle, \quad \forall g \in G, \ v_1, v_2 \in V.$$

Lemma 3.9. *Let G be a compact group. Let V be a f.d. G-representation. Then there exists a G-invariant inner product on V.*

Proof. Let $\langle -, - \rangle_0$ be any inner product on V. Denoting by $\int$ a Haar measure on G, let us define a function $\langle -, - \rangle : V \times V \to \mathbb{C}$ by

$$\langle v_1, v_2 \rangle := \int (g \mapsto \langle gv_1, gv_2 \rangle).$$

Then clearly $\langle -, - \rangle$ is an inner product on V (the strict positivity is a consequence of Exercise 3.3), and it is G-invariant, since given $h \in G$ we have

$$\langle hv_1, hv_2 \rangle = \int (g \mapsto \langle ghv_1, ghv_2 \rangle) = \int (g \mapsto \langle gv_1, gv_2 \rangle) = \langle v_1, v_2 \rangle.$$

□

Claim 3.10. *Let G be a compact group. Let V be a f.d. G-representation. Let $W \subset V$ be a G-invariant $\mathbb{C}$-linear subspace. Then there exists a G-invariant $\mathbb{C}$-linear subspace $U \subset V$ such that $V = W \oplus U$.*

Proof. Let $\langle -, - \rangle$ be a G-invariant inner product on V (which exists by Lemma 3.9). Consider $W^\perp \subset V$ - the orthogonal complement to W w.r.t. $\langle -, - \rangle$. We have $V = W \oplus W^\perp$, so it is enough to check that $W^\perp$ is G-invariant. Thus, given $v \in W^\perp$ and $g \in G$, we want to check that $gv \in W^\perp$. For this, we need to check that $\langle w, gv \rangle = 0$ for all $w \in W$. But we have

$$\langle w, gv \rangle = \langle g^{-1}w, g^{-1}(gv) \rangle = \langle g^{-1}w, v \rangle$$

and since $g^{-1}w \in W$, this is 0, as desired. □

Corollary 3.11. *Let G be a compact group. Every f.d. G-representation is the direct sum of irreducible f.d. G-representations.*

Proof. We continue breaking the representation into direct sum of smaller ones using Claim 3.10, until we hit irreducible representations. $\square$

Example 3.12. *Let us consider the representation of S_3 on $V := \mathbb{C}^3$, given by*

$$\sigma \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} := \begin{pmatrix} x_{\sigma^{-1}(1)} \\ x_{\sigma^{-1}(2)} \\ x_{\sigma^{-1}(3)} \end{pmatrix}.$$

The $\mathbb{C}$-span of $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, denote it by L, is an S_3-subrepresentation of V. Notice that the standard inner product on $V = \mathbb{C}^3$ is in fact S_3-invariant. Hence $L^\perp$ is an S_3-subrepresentation of V as well. It is an exercise to see that $L^\perp$ is an irreducible S_3-representation, which simply means in this case that there are no non-zero vectors in $L^\perp$ which are eigenvectors for all operators from the S_3-action.

Exercise 3.4. *Let V and W be abstract G-representations and let $T : V \to W$ be a morphism of G-representations. Show that* $\mathrm{Ker}(T)$ *is a G-subrepresentation of V and* $\mathrm{Im}(T)$ *is a G-subrepresentation of W.*

Claim 3.13 (Schur's lemma)**.** *Let G be a topological group. Let E and F be two irreducible f.d. G-representations. Then $Hom_G(E, F) = 0$ if E and F are not isomorphic and* $\dim_{\mathbb{C}} \mathrm{Hom}_G(E, F) = 1$ *if E and F are isomorphic.*

Proof. Let $T : E \to F$ be a non-zero morphism of G-representations. Consider $\mathrm{Ker}(T)$. Since it is a G-subrepresentation of E and since E is irreducible we have either $\mathrm{Ker}(T) = 0$ or $\mathrm{Ker}(T) = E$. In the latter case we have $T = 0$, so we must be in the former case, i.e. T is injective. Similarly, $\mathrm{Im}(T)$ is a G-subrepresentation of F and therefore $\mathrm{Im}(T) = 0$ or $\mathrm{Im}(T) = F$. In the former case $T = 0$ and so we must be in the latter case, i.e. T is surjective. Thus T is bijective, and hence an isomorphism of G-representations[12].

We have shown that if E and F are non-isomorphic then $\mathrm{Hom}_G(E, F) = 0$. Now assume that E and F are isomorphic, and we want to see that $\mathrm{Hom}_G(E, F)$ is 1-dimensional. It is enough[13] to check that $\mathrm{Hom}_G(E, E)$ is 1-dimensional. Of course, the 1-dimensional subspace of scalar operators lies in $\mathrm{Hom}_G(E, E)$, so we need to check that given $T \in \mathrm{Hom}_G(E, E)$ in fact T is a scalar operator. Let $\lambda \in \mathbb{C}$ be an eigenvalue of T. Since $\mathrm{Ker}(T - \lambda \cdot \mathrm{Id}_E)$ is a non-zero G-invariant $\mathbb{C}$-vector subspace of E, we must have $\mathrm{Ker}(T - \lambda \cdot \mathrm{Id}_E) = E$. So $T = \lambda \cdot \mathrm{Id}_E$, as desired. $\square$

[12] This is a very small exercise - a bijective morphism of G-representations is an isomoprhism of G-representations, i.e. its inverse is also a morphism of G-representations.

[13] That it is enough is a very small exercise.

Corollary-Definition 3.14. *Let G be a compact group. Let V be a f.d. G-representation and let E be an irreducible f.d. G-representation. Decomposition $V = E_1 \oplus \ldots \oplus E_n$ as a direct sum of irreducible f.d. G-representations, the number of $1 \le i \le n$ for which E_i is isomorphic to E is equal to $\dim_{\mathbb{C}} \mathrm{Hom}_G(E, V)$, and in particular it does not depend on the decomposition. It is called the* ***multiplicity of** E **in** V and denoted $[V : E]$.*

Proof. We have

$$\mathrm{Hom}_G(E, V) = \mathrm{Hom}_G(E, E_1 \oplus \ldots \oplus E_n) \cong \mathrm{Hom}_G(E, E_1) \oplus \ldots \oplus \mathrm{Hom}_G(E, E_n)$$

and by Schur's lemma the i-th summand is 1-dimensional if E_i is isomorphic to E and 0 otherwise. From this the claim is clear. □

Corollary 3.15. *Let G be a compact group. Let V and W be f.d. G-representations. Suppose that for every irreducible f.d. G-representation E we have $[V : E] = [W : E]$. Then V is isomorphic to W.*

Proof. We write V and W as direct sums of irreducible representations, and construct an isomorphism by adding isomorphisms between the various summands. □

3.3 Character

Let us fix a compact group G, and let us fix the Haar measure $\int \in \mathcal{M}(G)$ normalized to have total mass 1.

Definition 3.16. Let V be a f.d. G-representation. The **character** of V is the function $\mathrm{ch}_V \in C(G)$ given by

$$\mathrm{ch}_V(g) := \mathrm{Tr}\left(V \to V : \ v \mapsto gv\right).$$

Example 3.17. *The character of $\mathbb{C}_\chi$ is χ.*

Claim 3.18. *Let V be a f.d. G-representation.*

1. *The character ch_V is a **class function**, which means $\mathrm{ch}_V(hgh^{-1}) = \mathrm{ch}_V(g)$ for all $g, h \in G$.*
2. *We have $\mathrm{ch}_V(g^{-1}) = \overline{\mathrm{ch}_V(g)}$ for all $g \in G$.*

Proof. The first item follows immediately from the property $\mathrm{Tr}(STS^{-1}) = \mathrm{T}$ for $\mathbb{C}$-linear endomorphisms $T, S : V \to V$ of a f.d. $\mathbb{C}$-vector space. As for the second item, let $\langle -, - \rangle$ be a G-invariant inner product on V and denote by $T \in \mathrm{End}_{\mathbb{C}}(V)$ the operator $T(v) := gv$. We know that T is a unitary operator w.r.t. the inner product $\langle -.- \rangle$ and we want to see that $\mathrm{Tr}(T^{-1}) = \overline{\mathrm{Tr}(T)}$. This is an exercise in linear algebra (recall that a unitary operator is diagnolizable with all eigenvalues being complex numbers of absolute value 1). □

Given a f.d. G-representation V, let us denote by $V^G \subset V$ the $\mathbb{C}$-linear subspace of G-invariants:

$$V^G := \{v \in V \mid gv = v, \ \forall g \in G\}.$$

Lemma 3.19. *Let V be a f.d. G-representation. We have*

$$\int \mathrm{ch}_V = \dim_{\mathbb{C}} V^G.$$

Proof. Denote by $\pi(g) \in \mathrm{End}_{\mathbb{C}}(V)$ the operator given by $\pi(g)v := gv$. Let us define $P \in \mathrm{End}_{\mathbb{C}}(V)$ by (see Exercise 3.1 for integration of vector-valued functions)

$$P = \int (g \mapsto \pi(g)).$$

We claim that P is a projection operator onto V^G. Notice that for every $v \in V$ we have $P(v) = \int (g \mapsto gv)$. First, let us check that for any $v \in V$ we have $P(v) \in V^G$. Indeed, let $g \in G$. Then $gP(v) = g \int (h \mapsto hv) = \int (h \mapsto ghv) = \int (h \mapsto hv) = P(v)$, where in the third equality we used $\int$ being a Haar measure. Next, let us check that for $v \in V^G$ we have $P(v) = v$. Indeed, $P(v) = \int (g \mapsto gv) = \int (g \mapsto v) = (\int (g \mapsto 1))v = v$. Thus indeed P is a projection operator onto V^G. Therefore, $\mathrm{Tr}(P) = \dim_{\mathbb{C}} V^G$. But, on the other hand, we also have

$$\mathrm{Tr}(P) = \mathrm{Tr}\left(\int (g \mapsto \pi(g)) \right) = \int (g \mapsto \mathrm{Tr}(\pi(g))) = \int \mathrm{ch}_V.$$

□

Given two f.d. G-representations V and W, we construct a linear G-action on $\mathrm{Hom}_{\mathbb{C}}(V, W)$ as follows:

$$(gT)(v) := gT(g^{-1}v).$$

In this way we make $\mathrm{Hom}_{\mathbb{C}}(V, W)$ a G-representation. Notice that the subspace of G-invariants, $\mathrm{Hom}_{\mathbb{C}}(V, W)^G$ is equal to $\mathrm{Hom}_G(V, W)$, the space of morphisms of G-representations from V to W.

Lemma 3.20. *Let V and W be f.d. G-representations. We have*

$$\mathrm{ch}_{\mathrm{Hom}_{\mathbb{C}}(V,W)}(g) = \mathrm{ch}_W(g) \cdot \mathrm{ch}_V(g^{-1}) = \mathrm{ch}_W(g) \cdot \overline{\mathrm{ch}_V(g)} \quad \forall g \in G.$$

Proof. The second equality is just item (2) of Claim 3.18. The first equality follows from the following exercise in linear algebra: Let $T \in \mathrm{End}_{\mathbb{C}}(V)$ and $S \in \mathrm{End}_{\mathbb{C}}(W)$. Define $R \in \mathrm{End}_{\mathbb{C}}(\mathrm{Hom}_{\mathbb{C}}(V, W))$ by

$$R(A) := S \circ A \circ T.$$

Then

$$\mathrm{Tr}(R) = \mathrm{Tr}(S) \cdot \mathrm{Tr}(T).$$

□

Definition 3.21. We define an inner product $\langle -, - \rangle_G$ on $C(G)$ by:

$$\langle f_1, f_2 \rangle_G := \int (g \mapsto f_1(g) \cdot \overline{f_2(g)}).$$

Corollary 3.22. *Let V and W be f.d. G-representations. Then*

$$\langle ch_W, ch_V \rangle = \dim_{\mathbb{C}} \mathrm{Hom}_G(V, W).$$

Proof. We have

$$\dim_{\mathbb{C}} \mathrm{Hom}_G(V, W) = \dim_{\mathbb{C}}(\mathrm{Hom}_{\mathbb{C}}(V, W)^G) = \int (g \mapsto \mathrm{ch}_{\mathrm{Hom}_{\mathbb{C}}(V,W)}(g)) =$$

$$= \int (g \mapsto \mathrm{ch}_W(g) \cdot \overline{\mathrm{ch}_V(g)}) = \langle \mathrm{ch}_W, \mathrm{ch}_V \rangle_G.$$

□

Corollary 3.23 (Orthogonality relations)**.** *The functions $\mathrm{ch}_E \in C(G)$ as E runs over non-isomorphic irreducible f.d. G-representations form an orthonormal set, and thus in particular a linearly independnt set.*

Proof. By the previous corollary and by Schur's lemma (Claim 3.13) we have for an irreducible f.d. G-representation E

$$\langle \mathrm{ch}_E, \mathrm{ch}_E \rangle = \dim_{\mathbb{C}} \mathrm{Hom}_G(E, E) = 1$$

and for non-isomorphic irreducible f.d. G-representations E and F we have

$$\langle \mathrm{ch}_E, \mathrm{ch}_F \rangle = \dim_{\mathbb{C}} \mathrm{Hom}_G(E, F) = 0.$$

□

Remark 3.24. It is also true that $\{\mathrm{ch}_E\}_{[E] \in \mathrm{Irr}(G)}$ form a "complete" orthonormal system in the space of class functions on G, that is, given $f \in C(G)$ which is a class function, if $\langle f, \mathrm{ch}_E \rangle = 0$ for all $[E] \in \mathrm{Irr}(G)$ then $f = 0$. In a different terminology, $\{\mathrm{ch}_E\}_{[E] \in \mathrm{Irr}(G)}$ forms a Hilbert basis for the Hilbert space $L^2(G)^G$ of square-integrable class functions on G.

Claim 3.25. *Let V and W be f.d. G-representations. If $\mathrm{ch}_V = \mathrm{ch}_W$ then V is (non-canonically) isomorphic to W.*

Proof. As explained above, to see that V is isomorphic to W it is enough to see that $[V : E] = [W : E]$ for all irreducible f.d. G-representations E. We saw that $[V : E] = \dim_{\mathbb{C}} \mathrm{Hom}_G(E, V) = \langle \mathrm{ch}_E, \mathrm{ch}_V \rangle$ and from this the claim follows. □

Thus, at least in some basic sense, the problem of finding all irreducible f.d. G-representations up to isomorphism can be thought of as solved once we are able to write down all their characters.

Remark 3.26. We calculate easily that given a f.d. G-representation V, we have that V is irreducible if and only if $\langle \mathrm{ch}_V, \mathrm{ch}_V \rangle = 1$, i.e. the character of V has length 1.

3.4 Character in the case of $\mathrm{SU}(n)$

In this course we will focus on the representation theory of the groups $\mathrm{SU}(n)$. The difference with $\mathrm{U}(n)$ is not big, but $\mathrm{SU}(n)$ has a finite center, making things a bit more tidy. Also, one could study $\mathrm{SO}(n)$ instead, but, again, things will be a bit more tidy for $\mathrm{SU}(n)$.

We denote by $T \subset \mathrm{SU}(n)$ the subgroup consisting of diagonal matrices. We have an isomorphism of topological groups

$$(\mathbb{C}^{\times}_{|-|=1})^{n-1} \xrightarrow{\sim} T$$

given by

$$(t_1, \ldots, t_{n-1}) \mapsto \mathrm{diag}\left(t_1, \ldots, t_{n-1}, \frac{1}{t_1 \cdot \ldots \cdot t_{n-1}}\right)$$

(where $\mathrm{diag}(a_1, \ldots, a_n)$ stands for the diagonal matrix with values $a_1, \ldots, a_n$ on the diagonal).

Claim 3.27. *Every element in $\mathrm{SU}(n)$ is conjugate to an element in T.*

Proof. Let $g \in \mathrm{SU}(n)$. By linear algebra (every unitary transformation is unitarily diagnolizable) there exists $h \in \mathrm{U}(n)$ such that $g' := hgh^{-1}$ is diagonal. Since $\det(g') = 1$, i.e. $g' \in \mathrm{SU}(n)$, we have $g' \in T$. Denote $c := \det(h)$. Then $c \in \mathbb{C}^{\times}_{|-|=1}$. Denote $h' := c^{-1/n} \cdot h$. Then $\det(h') = 1$, i.e. $h' \in \mathrm{SU}(n)$, and still $h'g(h')^{-1} = g'$. □

Corollary 3.28. *Let V and W be f.d. $\mathrm{SU}(n)$-representations. If $(\mathrm{ch}_V)|_T = (\mathrm{ch}_W)|_T$ then $\mathrm{ch}_V = \mathrm{ch}_W$ and thus V is isomorphic to W.*

Therefore, a uniquely determining attribute of a f.d. $\mathrm{SU}(n)$-representation V is $(\mathrm{ch}_V)|_T$. Let us recall some linear algebra:

Exercise-Definition 3.29. *Let V be a f.d. $\mathbb{C}$-vector space. Let $S \subset \mathrm{End}_{\mathbb{C}}(V)$ be a subset consisting of pairwise commuting diagnolizable operators. Then, given a function $\chi : S \to \mathbb{C}$ denoting*

$$V_{S,\chi} := \{v \in V \mid Tv = \chi(T)v \ \forall T \in S\},$$

we have

$$V = \bigoplus_{\chi : S \to \mathbb{C}} V_{S,\chi}.$$

In fact, if for some χ we have $V_{S,\chi} \neq 0$ then χ is continuous and also whenever $T_1, T_2, T_1T_2 \in S$ we have $\chi(T_1T_2) = \chi(T_1)\chi(T_2)$. In particular, if S is a subgroup in $\mathrm{GL}_{\mathbb{C}}(V)$ then if for some χ we have $V_{S,\chi} \neq 0$ then χ must be a topological group morphism $S \to \mathbb{C}^{\times}$ and therefore we have

$$V = \bigoplus_{\chi \in \mathrm{qCh}(S)} V_{S,\chi}.$$

*Let us denote by $\mathrm{wt}_S(V) \subset \mathrm{qCh}(S)$ the subset of χ's for which $V_{S,\chi} \neq 0$ (the subset of **weights**).*

Exercise-Definition 3.30. *Let S be an abelian toploogical group. Show that every irreducible f.d. S-representation is 1-dimensional. Thus given a f.d. S-representation V we can write $V = L_1 \oplus \ldots \oplus L_n$ where L_i are 1-dimensional S-subrepresentations. So each L_i is isomorphic to $\mathbb{C}_{\chi_i}$ for some $\chi_i \in \mathrm{Ch}(S)$. See that for every $\chi \in \mathrm{Ch}(S)$ we have*

$$V_{S,\chi} = \bigoplus_{\substack{1 \le i \le n \\ [L_i] = [\mathbb{C}_\chi]}} L_i.$$

Now let us go back to a f.d. $\mathrm{SU}(n)$-representation V. Since the operators by which elements of G act on V are unitary w.r.t. some inner product, they are all diagnolizable. In particular, the operators by which elements of T act on V form a subgroup of $\mathrm{GL}_{\mathbb{C}}(V)$ consisting of pairwise commuting diagnolizable operators. Hence we can write

$$V = \bigoplus_{\chi \in \mathrm{Ch}(T)} V_{T,\chi}$$

and we have, for $t \in T$,

$$\mathrm{ch}_V(t) = \sum_{\chi \in \mathrm{Ch}(T)} (\dim_{\mathbb{C}} V_{T,\chi}) \cdot \chi(t)$$

Hence a uniquely determining attribute of a f.d. $\mathrm{SU}(n)$-representation V is the vector of dimensions

$$\left(\dim_{\mathbb{C}}(V_{T,\chi})\right)_{\chi \in \mathrm{Ch}(T)}.$$

The vector of dimensions recovers $\mathrm{ch}_V|_T$ and, conversely, $\mathrm{ch}_V|_T$ recovers the vector of dimensions:

Exercise 3.5. *Using Corollary 3.23 notice that $\mathrm{Ch}(T)$ is a linearly independent subset of $C(T)$.*

Another important piece of symmetry we have is as follows. Let W denote the group S_n of permutations on $\{1, \ldots, n\}$. In this context it is called the **Weyl group**. We have an action of W on T by:

$$w \cdot \mathrm{diag}(t_1, \ldots, t_n) = \mathrm{diag}(t_{w^{-1}(1)}, \ldots, t_{w^{-1}(n)}).$$

It is an action by topological group automorphisms. In the following exercise we give another way to look at W.

Exercise 3.6. *Consider $N_{\mathrm{SU}(n)}(T)$, the **normalizer** of T in $\mathrm{SU}(n)$. Show that it is equal to subgroup of "permutation matrices", i.e. matrices whose every row contains exactly one non-zero entry. Consider $Z_{\mathrm{SU}(n)}(T)$, the **centralizer** of T in $\mathrm{SU}(n)$. Show that $Z_{\mathrm{SU}(n)}(T) = T$. In general, recall that given a subgroup $H \subset G$ then $Z_G(H)$ is normal in $N_G(H)$, and we have a natural action of $N_G(H)/Z_G(H)$ on H by group automorphisms, via conjugation. This action is*

***faithful**, in the sense that if an element acts trivially then it is trivial. Back to our case, identify $N_{\mathrm{SU}(n)}(T)/Z_{\mathrm{SU}(n)}(T)$ with W via their actions on T (i.e. both identify with the same subgroup of the group of automoprhisms of the group T).*

Given $w \in W$, we will denote $\dot{w} \in N_{\mathrm{SU}(n)}(T)$ an element such that $\dot{w}t\dot{w}^{-1} = wt$ for all $t \in T$ (all what we will say will not depend on this choice). The action of W on T induces an action of W on $\mathrm{Ch}(T)$, by $(w\chi)(t) := \chi(w^{-1}t)$. Now, given a f.d. G-representation V, notice that we have

$$\sum_{\chi \in \mathrm{Ch}(T)} \dim_{\mathbb{C}} V_{T,\chi} \cdot \chi(t) = \mathrm{ch}_V(t) = \mathrm{ch}_V(\dot{w}t\dot{w}^{-1}) = \sum_{\chi \in \mathrm{Ch}(T)} \dim_{\mathbb{C}} V_{T,\chi} \cdot \chi(\dot{w}t\dot{w}^{-1}) =$$

$$= \sum_{\chi \in \mathrm{Ch}(T)} \dim_{\mathbb{C}} V_{T,\chi} \cdot (w^{-1}\chi)(t) = \sum_{\chi \in \mathrm{Ch}(T)} \dim_{\mathbb{C}} V_{T,w\chi} \cdot \chi(t)$$

and therefore $\dim_{\mathbb{C}} V_{T,w\chi} = \dim_{\mathbb{C}} V_{T,\chi}$ for all $w \in W$ and $\chi \in \mathrm{Ch}(T)$. Another way to explain this equality is to notice that $\dot{w}V_{T,\chi} = V_{T,w\chi}$.

3.5 Example: Irreps of SU(2)

Let us consider SU(2). Then matrices in T look like

$$\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix}$$

for $t \in \mathbb{C}^{\times}_{|-|=1}$, and every character of T looks like

$$\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} \mapsto t^m$$

for some uniquely defined $m \in \mathbb{Z}$. Let us also denote by χ_1 the character corresponding to $m := 1$. The non-trivial element in W sends $\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix}$ to $\begin{pmatrix} t^{-1} & 0 \\ 0 & t \end{pmatrix}$, and therefore sends χ_1^m to χ_1^{-m}.

Let $m \in \mathbb{Z}_{\geq 0}$. Consider the space $\mathcal{P}_m$ of homogeneous complex polynomials in two variables x, y, of degree n. So:

$$\mathcal{P}_0 = \mathrm{span}_{\mathbb{C}}\{1\},$$

$$\mathcal{P}_1 = \mathrm{span}_{\mathbb{C}}\{x, y\},$$

$$\mathcal{P}_2 = \mathrm{span}_{\mathbb{C}}\{x^2, xy, y^2\}$$

and so on. The natural action of $\mathrm{SU}(2) \subset \mathrm{GL}_2(\mathbb{C})$ on $\mathbb{C}^2$ (by multiplying a vector by a matrix) gives rise to a natural $\mathbb{C}$-linear action of SU(2) on $\mathcal{P}_m$ given by:

$$(gf)\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) := f\left(g^{-1}\begin{pmatrix} x \\ y \end{pmatrix}\right).$$

Let us calculate the character of $\mathcal{P}_m$. For $0 \le i \le m$, denote by $f_m^i \in \mathcal{P}_m$ the polynomial $f_m^i\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) := x^i y^{m-i}$ (those form a $\mathbb{C}$-basis for $\mathcal{P}_m$). Notice that

$$\left(\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} f_m^i\right)\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = f_m^i\left(\begin{pmatrix} t^{-1} & 0 \\ 0 & t \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\right) = f_m^i\left(\begin{pmatrix} t^{-1}x \\ ty \end{pmatrix}\right) =$$

$$= (t^{-1}x)^i (ty)^{m-i} = t^{m-2i} \cdot x^i y^{m-i} = t^{m-2i} \cdot f_m^i\left(\begin{pmatrix} x \\ y \end{pmatrix}\right)$$

i.e. we got

$$\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} f_m^i = t^{m-2i} \cdot f_m^i.$$

This shows that

$$\mathrm{ch}_{\mathcal{P}_m}\left(\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix}\right) = t^{-m} + t^{-m+2} + \ldots + t^{m-2} + t^m$$

i.e.

$$(\mathrm{ch}_{\mathcal{P}_m})|_T = \chi_1^{-m} + \chi_1^{-m+2} + \ldots + \chi_1^{m-2} + \chi_1^m.$$

Claim 3.31. *Each $\mathcal{P}_m$ is an irreducible f.d. $\mathrm{SU}(2)$-representation. Every irreducible f.d. $\mathrm{SU}(2)$-representation is isomorphic to some $\mathcal{P}_m$.*

Proof. We will explain it later. □

3.6 A glimpse at Weyl's integration formula

In this subsection, we abbreviate $G := \mathrm{SU}(n)$. When we speak of an action of G on G, we always mean the conjugation action. Also, we denote by $\int_G$ the Haar measure of mass 1 on G and by $\int_T$ the Haar measure of mass 1 on T.

Claim 3.32. *Restriction of functions from G to T yields an isomorphism of $\mathbb{C}$-vector spaces*

$$\mathrm{Res} : C(G)^G \xrightarrow{\sim} C(T)^W.$$

Proof. Clearly the restriction of a G-invariant function on G is a W-invariant function on T. Let us see that Res is bijective. Let us consider the orbit spaces $G\backslash G$ (under conjugation action) and $W\backslash T$. Then we can interpret $C(G)^G$ as $C(G\backslash G)$ and $C(T)^W$ as $C(W\backslash T)$, and Res is given by precomposing with the natural map $W\backslash T \to G\backslash G$. Therefore, it is enough to see that this last map is an isomorphism of topological spaces. This map is continuous. It is surjective by Claim 3.27 and injective because if two elements in T are conjugate in G then they have the same multisets of eigenvalues and therefore the same diagonal values up to permutation. Recall that a bijective continuous map between compact spaces is a homeomorphism. Hence, it is enough to check that $W\backslash T$ and $G\backslash G$ are compact. If we can show those are Hausdorff, then those are compact as Hausdorff quotients of the compact spaces G and T. To show that

$W\backslash T$ is Hausdorff, we need to take $t_1, t_2 \in T$ such that $Wt_1 \neq Wt_2$ and find disjoint W-invariant open subsets $U_1, U_2 \subset T$ such that $t_1 \in U_1$ and $t_2 \in U_2$. Take $U_1', U_2' \subset T$ be disjoint open subsets such that $Wt_1 \in U_1'$ and $Wt_2 \in U_2'$. Set $U_i := \cap_{w \in W} wU_i'$ (those are open subsets(!) as the intersections of finitely many open subsets). Those are as required. To show that $G\backslash G$ is Hausdorff, notice that it is enough to produce a continuous map $\phi : G \to X$ to some Hausdorff topological space X, which is G-invariant (i.e. $\phi(gg'g^{-1}) = \phi(g')$ for all $g, g' \in G$) and with the property that given $g_1, g_2 \in G$ such that $Gg_1 \neq Gg_2$ we have $\phi(g_1) \neq \phi(g_2)$. Consider the characteristic polynomial map $\phi : G \to \mathrm{Pol}_n(\mathbb{C})$ where $\mathrm{Pol}_n(\mathbb{C})$ is the $(n+1)$-dimensional $\mathbb{C}$-vector space of polynomials of degree $\leq n$. It has the desired properties (we could also use it for $W\backslash T$, but wanted to demonstrate another principle there, when a finite group acts). $\square$

Now, let us define a $\mathbb{C}$-linear map

$$\mathrm{Av}_W : C(T) \to C(T)^W$$

by

$$\mathrm{Av}_W(f)(t) := \frac{1}{|W|} \sum_{w \in W} f(wt)$$

(i.e. it is the averaging map).

Claim 3.33. *Let I be a G-invariant signed Radon measure on G. There exists a unique W-invariant signed Radon measure $I^{(T)}$ on T such that*

$$I(f) = I^{(T)}(\mathrm{Res}(f)) \quad \forall f \in C(G)^G.$$

Also, $I^{(T)}$ is a Radon measure if I is.

Proof. Let us show uniqueness first. If we have two such J_1, J_2, then $J_1(h) = J_2(h)$ for all $h \in C(T)^W$. But then for any $h \in C(T)$ we obtain $J_1(h) = J_1(\mathrm{Av}_W(h)) = J_2(\mathrm{Av}_W(h)) = J_2(h)$ and so $J_1 = J_2$.

Let us show existence now. Define $I^{(T)}$ by

$$I^{(T)}(h) := I(\mathrm{Res}^{-1}(\mathrm{Av}_W(h))).$$

It is clearly a W-invariant functional on $C(T)$, and it satisfies the desired property: For $f \in C(G)^G$ we have

$$I^{(T)}(\mathrm{Res}(f)) = I(\mathrm{Res}^{-1}(\mathrm{Av}_W(\mathrm{Res}(f)))) = I(\mathrm{Res}^{-1}(\mathrm{Res}(f))) = I(f).$$

So it only is left to see that $I^{(T)}$ satisfies the continuity property required from a Radon measure. For this it is enough to check that if $\{h_n\}$ is a sequence in $C(T)$ converging uniformly on T to $h \in C(T)$ then $\mathrm{Res}^{-1}(\mathrm{Av}_W(h_n))$ converges uniformly on G to $\mathrm{Res}^{-1}(\mathrm{Av}_W(h))$. This is immediate to see.

That $I^{(T)}$ is a Radon measure if I is a Radon measure is immediate to see. $\square$

Applying the claim to $I := \int_G$, we obtain a W-invariant Radon measure $\int_G^{(T)}$ on T satisfying

$$\int_G f = \int_G^{(T)} f|_T \quad \forall f \in C(G)^G.$$

Weyl's integration formula gives a formula for that $\int_G^{(T)}$. Let us state it for SU(2). In the statement we use a general notation - if $\mu \in \mathcal{M}(X)$ and $h \in C(X)$ then $h\mu \in \mathcal{M}(X)$ denotes the signed Radon measure given by $(h\mu)(f) := \mu(hf)$.

Theorem 3.34 (Weyl's integration formula for SU(2))**.** *Assume $G = \mathrm{SU}(2)$. We have*

$$\int_G^{(T)} = \tfrac{1}{2}|\chi_1 - \chi_1^{-1}|^2 \int_T.$$

Proof. Omitted. □

Let us now use this theorem to see that the SU(2)-representations $\mathcal{P}_m$ we constructed are irreducible. For that, it is enough to see that $\langle \mathrm{ch}_{\mathcal{P}_m}, \mathrm{ch}_{\mathcal{P}_m} \rangle = 1$. And indeed, we calculate:

$$\langle \mathrm{ch}_{\mathcal{P}_m}, \mathrm{ch}_{\mathcal{P}_m} \rangle = \int_G |\mathrm{ch}_{\mathcal{P}_m}|^2 = \int_T \frac{1}{2}|\chi_1 - \chi_1^{-1}|^2 |\chi_1^m + \chi_1^{m-2} + \ldots + \chi_1^{-m+2} + \chi_1^{-m}|^2 =$$

$$= \frac{1}{2}\int_T |\chi_1^{m+1} - \chi_1^{-(m+1)}|^2 = \frac{1}{2}\int_T (\chi_1^{m+1} - \chi_1^{-(m+1)})(\chi_1^{-(m+1)} - \chi_1^{m+1}) =$$

$$= \frac{1}{2}\int_T \left(2 + \chi_1^{2(m+1)} + \chi_1^{-2(m+1)}\right) = 1.$$

Here we used the orthogonality relations

$$\int_T \chi_1^m = \int_T \chi_1^m \cdot \overline{1} = \langle \chi_1^m, 1 \rangle = \begin{cases} 1 & m = 0 \\ 0 & m \neq 0 \end{cases}.$$

In fact, we can also use the formula to see that there are no irreducible f.d. SU(2)-representations except from the $\mathcal{P}_m$'s. To that end, let us consider an irreducible f.d. SU(2)-representation E and try to see what it means for it to be different from all the $\mathcal{P}_m$'s. This means that $\langle \mathrm{ch}_E, \mathrm{ch}_{\mathcal{P}_m} \rangle = 0$ for all $m \in \mathbb{Z}_{\geq 0}$. Let us write $(\mathrm{ch}_E)|_T = \sum_m d_m \cdot \chi_1^m$ for $d_m \in \mathbb{Z}_{\geq 0}$. Here, as we remember, $d_m = \dim_{\mathbb{C}} E_{T,\chi_1^m}$. The sum is finite (i.e. $d_m = 0$ except for finitely many m's), and we have the symmetry property $d_{-m} = d_m$ for all m. We have:

$$\langle \mathrm{ch}_E, \mathrm{ch}_{\mathcal{P}_m} \rangle = \int_G \mathrm{ch}_E \cdot \overline{\mathrm{ch}_{\mathcal{P}_m}} = \int_T \frac{1}{2}|\chi_1 - \chi_1^{-1}|^2 (\mathrm{ch}_E)|_T (\chi_1^m + \chi_1^{m-2} + \ldots + \chi_1^{-m+2} + \chi_1^{-m}) =$$

$$= \frac{1}{2}\int_T (\mathrm{ch}_E)|_T (\chi_1^m + \chi_1^{-m} - \chi_1^{m+2} - \chi_1^{-(m+2)}) = \frac{1}{2}\left(d_{-m} + d_m - d_{-(m+2)} - d_{m+2}\right).$$

Thus, the condition $\langle \mathrm{ch}_E, \mathrm{ch}_{\mathcal{P}_m} \rangle = 0$ for all $m \in \mathbb{Z}_{\geq 0}$ translates into:

$$d_{m+2} + d_{-(m+2)} = d_m + d_{-m}$$

for all $m \in \mathbb{Z}_{\geq 0}$ or, by the W-symmetry,

$$d_{m+2} = d_m$$

for all $m \in \mathbb{Z}_{\geq 0}$. This implies that $d_m = 0$ for all $m \in \mathbb{Z}$, because otherwise we would have infinitely many m's for which $d_m \neq 0$.

3.7 Classification of irreps of $\mathrm{SU}(n)$ and Weyl's character formula (without proofs)

Remark 3.35. Let us say that an abelian group A is a **lattice** if it is isomorphic to $\mathbb{Z}^n$ for some $n \in \mathbb{Z}_{\geq 0}$. Since $\mathbb{Z}^n$ is not isomorphic to $\mathbb{Z}^m$ whenever $n \neq m$, the number n does not depend on the choice of an isomorphism; It is called the **rank** of the lattice A. Recall that an abelian group is a lattice if and only if it is finitely generated and torsion-free. A list $e_1, \ldots, e_n$ of elements in A is said to be a **$\mathbb{Z}$-basis for** A if the morphism of abelian groups $\mathbb{Z}^n \to A$ given by $(d_1, \ldots, d_n) \mapsto d_1 e_1 + \ldots + d_n e_n$ is an isomorphism (in other words, every element in A can be expressed as a $\mathbb{Z}$-linear combination of the list $e_1, \ldots, e_n$, uniquely so).

We can always think of $\mathbb{Z}^n$ as sitting inside $\mathbb{R}^n$, giving us a geometric picture (for example, we can picture whether an element in $\mathbb{Z}^n$ sits in the convex hull of some set of elements in $\mathbb{Z}^n$). In fact, for any lattice A, of rank n, we can construct an n-dimensional $\mathbb{R}$-vector space $A_{\mathbb{R}}$ together with a morphism of abelian groups $\iota : A \to A_{\mathbb{R}}$ with the property that given a $\mathbb{Z}$-basis $\{e_i\}_{1 \leq i \leq n}$ of A, $\{\iota(e_i)\}_{1 \leq i \leq n}$ is an $\mathbb{R}$-basis of $A_{\mathbb{R}}$. In other words, instead of only linear combinations of $e_1, \ldots, e_n$ with integer coefficients, we now allow linear combinations of $e_1, \ldots, e_n$ with real coefficients. To construct $A_{\mathbb{R}}$, just choose an isomorphism $e : A \xrightarrow{\sim} \mathbb{Z}^n$, take $A_{\mathbb{R}} := \mathbb{R}^n$ and set ι to be the composition of the isomorphism e with the natural embedding of $\mathbb{Z}^n$ in $\mathbb{R}^n$. If the reader knows tensor products, a description of $A_{\mathbb{R}}$ which does not depend on choices is as $\mathbb{R} \underset{\mathbb{Z}}{\otimes} A$ (another description is $\mathrm{Hom}_{\mathbb{Z}}(\mathrm{Hom}_{\mathbb{Z}}(A, \mathbb{Z}), \mathbb{R})$).

Another way to describe the pair $(A_{\mathbb{R}}, \iota)$ is via a **universal property**. Namely, given any pair (V, μ) consisting of a $\mathbb{R}$-vector space V and a morphism of abelian groups $\mu : A \to V$, there exists a unqiue morphism of $\mathbb{R}$-vector spaces $\iota_\mu : A_{\mathbb{R}} \to V$ such that $\iota_\mu \circ \iota = \mu$.

In particular, one sees that given an automorphism (of abelian groups) $T : A \to A$, there exists a unique automorphism (of $\mathbb{R}$-vector spaces) $T_{\mathbb{R}} : A_{\mathbb{R}} \to A_{\mathbb{R}}$ extending it (i.e. $T_{\mathbb{R}}(\iota(a)) = \iota(T(a))$ for all $a \in A$). This yields in fact a group homomorphism $\mathrm{Aut}(A) \to \mathrm{Aut}(A_{\mathbb{R}})$, where $\mathrm{Aut}(A)$ denotes the group of automorphisms of A as an abelian group, while $\mathrm{Aut}(A_{\mathbb{R}})$ denotes the group of automorphisms of $A_{\mathbb{R}}$ as an $\mathbb{R}$-vector space (i.e. another name for $\mathrm{Aut}(A_{\mathbb{R}})$ is $\mathrm{GL}_{\mathbb{R}}(A_{\mathbb{R}})$).

The action of W on $\mathrm{Ch}(T)$ induces an action of W on $\mathrm{Ch}(T)_{\mathbb{R}}$. Let us fix a W-invariant inner product $\langle -, - \rangle$ on $\mathrm{Ch}(T)_{\mathbb{R}}$.

Definition 3.36. Let V be a f.d. $\mathrm{SU}(n)$-representation. We say that $\chi \in \mathrm{wt}_T(V)$ is **extremal** if, working in $\mathrm{Ch}(T)_{\mathbb{R}}$, every $\theta \in \mathrm{wt}_T(V)$ belongs to the convex hull of $\{w\chi\}_{w\in W}$.

Lemma 3.37. *Let V be a f.d. $\mathrm{SU}(n)$-representation. If $\chi \in \mathrm{wt}_T(V)$ is extremal, then $\theta \in \mathrm{wt}_T(V)$ is extremal if and only if $\theta = w\chi$ for some $w \in W$.*

Proof. It is clear that $w\chi$ is also extremal for each $w \in W$. Given $\theta \in \mathrm{wt}_T(V)$, consider the length $||\theta||$ (with respect to our W-invariant inner product $\langle -, - \rangle$) on $\mathrm{Ch}(T)_{\mathbb{R}}$). Notice that, since $||w\chi|| = ||\chi||$ for all $w \in W$ and θ lies in the convex hull of $\{w\chi\}_{w\in W}$, we have $||\theta|| \leq ||\chi||$. If $||\theta|| < ||\chi||$, then the convex hull of $\{w\theta\}_{w\in W}$ clearly does not contain χ, and so θ is not extremal. If $||\theta|| = ||\chi||$ then, since θ lies in the convex hull of the points $w\chi$ which all lie on the sphere of radius $||\chi||$ around the origin, an exercise shows that we must have $\theta = w\chi$ for some $w \in W$ (the exercise is that if a point on the sphere is a convex combination of a finite collection of points on the sphere, then it must be equal to one of them). $\square$

We have:

Proposition 3.38. *Every irreducible f.d. $\mathrm{SU}(n)$-representation admits extremal weights.*

Proof. Omitted. $\square$

Denote by $W\backslash\mathrm{Ch}(T)$ the orbit space, i.e. the quotient of $\mathrm{Ch}(T)$ by the equivalence relation $\chi \sim \theta$ if there exists $w \in W$ such that $\chi = w\theta$. The above shows that we have a map

$$\mathcal{E} : \mathrm{Irr}(\mathrm{SU}(n)) \to W\backslash\mathrm{Ch}(T)$$

given by sending the isomorphism class of an irreducible f.d. $\mathrm{SU}(n)$-representation V to the orbit of an extremal weight of V.

Theorem 3.39. *The map $\mathcal{E}$ is a bijection.*

In other words, the theorem classifies irreducible f.d. representations of $\mathrm{SU}(n)$. Given $\chi \in \mathrm{Ch}(T)$, by V_χ we will denote an irreducible f.d. $\mathrm{SU}(n)$-representation such that $\mathcal{E}([V_\chi]) = W\chi$. Next, after we have parametrized irreducible f.d. $\mathrm{SU}(n)$-representations, we would like to describe their characters.

It is convenient to choose representatives for W-orbits in $\mathrm{Ch}(T)$:

Definition 3.40. Let us say that $\chi \in \mathrm{Ch}(T)$ is **dominant** if, writing

$$\chi(\mathrm{diag}(t_1, \ldots, t_n)) = t_1^{m_1} \cdot \ldots \cdot t_n^{m_n},$$

we have $m_1 \geq m_2 \geq \ldots \geq m_n$.

Exercise 3.7. *Any W-orbit in $\mathrm{Ch}(T)$ contains a unique dominant element.*

We have a bijection

$$(\mathbb{Z}_{\geq 0})^{n-1} \xrightarrow{\sim} \{\chi \in \mathrm{Ch}(T) \mid \chi \text{ is dominant}\}$$

which we denote by $(d_1, \ldots, d_{n-1}) \mapsto \chi_{d_1,\ldots,d_{n-1}}$, given by

$$\chi_{d_1,\ldots,d_{n-1}}(\mathrm{diag}(t_1, \ldots, t_n)) := t_1^{d_1+d_2+\ldots+d_{n-1}} \cdot \ldots \cdot t_{n-2}^{d_{n-2}+d_{n-1}} \cdot t_{n-1}^{d_{n-1}}.$$

We will also have some special characters we will need:

Exercise-Definition 3.41. *Given $1 \leq i < j \leq n$, let us denote by $\alpha_{i,j} \in \mathrm{Ch}(T)$ the character given by*

$$\alpha_{i,j}(\mathrm{diag}(t_1, \ldots, t_n)) := \frac{t_i}{t_j}.$$

Let $\Delta \in \mathrm{Ch}(T)$ be given by

$$\Delta(t) := \prod_{1 \leq i < j \leq n} \alpha_{i,j}(t).$$

Show that there exists a unique character $\sqrt{\Delta} \in \mathrm{Ch}(T)$ satisfying $\sqrt{\Delta}^2 = \Delta$. Namely, $\sqrt{\Delta} = \chi_{1,\ldots,1}$.

Theorem 3.42 (Weyl's character formula)**.** *Let $\chi \in \mathrm{Ch}(T)$ be dominant. We have (for $t \in T$ for which the denominator does not vanish, which happens on the dense subset of T consisting of elements whose entries are pairwise non-equal)*

$$\mathrm{ch}_{V_\chi}(t) = \frac{\sum_{w \in W} \mathrm{sgn}(w) \cdot (w(\chi\sqrt{\Delta}))(t)}{\sqrt{\Delta}(t) \prod_{1 \leq i < j \leq n} (1 - \alpha_{i,j}(t)^{-1})}.$$

Example 3.43. *Let us see what the above means for $\mathrm{SU}(2)$. Dominant characters are χ_d for $d \in \mathbb{Z}_{\geq 0}$. Notice that $\sqrt{\Delta}(\mathrm{diag}(t, t^{-1})) = t$. The character of V_{χ_d} is given by:*

$$\mathrm{ch}_{V_{\chi_d}}(\mathrm{diag}(t, t^{-1})) = \frac{t^d \cdot t - t^{-d} \cdot t^{-1}}{t(1 - t^{-2})} = t^d \frac{1 - t^{-2(d+1)}}{1 - t^{-2}} =$$

$$= t^d(1 + t^{-2} + t^{-4} + \ldots + t^{-2d}) = t^d + t^{d-2} + \ldots + t^{-d}.$$

Deduce from this that V_{χ_d} is isomorphic to $\mathcal{P}_d$ that we had above.

Another way to interpret the deduction of the previous example is

$$\mathrm{ch}_{V_{\chi_d}}(\mathrm{diag}(t, t^{-1})) = (t^d - t^{-d-2})(1 + t^{-2} + t^{-4} + \ldots) =$$

$$= (t^d + t^{d-2} + \ldots) - (t^{-d-2} + t^{-d-4} + \ldots) = t^d + t^{d-2} + \ldots + t^{-d}.$$

In other words, we have two infinite series, whose difference happens to be finite (there are a lot of cancellations). This hints at the approach to the proof of

Weyl's character formula we want to take in this course eventually - that V_{χ_d} itself is, in some sense, the "difference" of two infinite-dimensional representations with "characters" our two infinite series. Let us again emphasize the issue of cancellations of infinities, for the general case of $\mathrm{SU}(n)$. We can rewrite Weyl's character formula in the following way, operating formally at least:

$$\mathrm{ch}_V = \sum_{w \in W} \mathrm{sgn}(w) \cdot w\chi \cdot \frac{w\sqrt{\Delta}}{\sqrt{\Delta}} \cdot \prod_{1 \leq i < j \leq n} \left(1 + \alpha_{i,j}^{-1} + \alpha_{i,j}^{-2} + \ldots\right). \tag{3.1}$$

This is an alternating sum of shifts of some "cone-like" expression. Somehow, everything cancels except things lying inside the convex hull of finitely many points $\{w\chi\}_{w \in W}$. We will illustrate this in the next subsection.

3.8 Some illustrations for Weyl's character formula for SU(3)

Let us consider SU(3) now. Let us define $\omega_1, \omega_2 \in \mathrm{Ch}(T)$ by

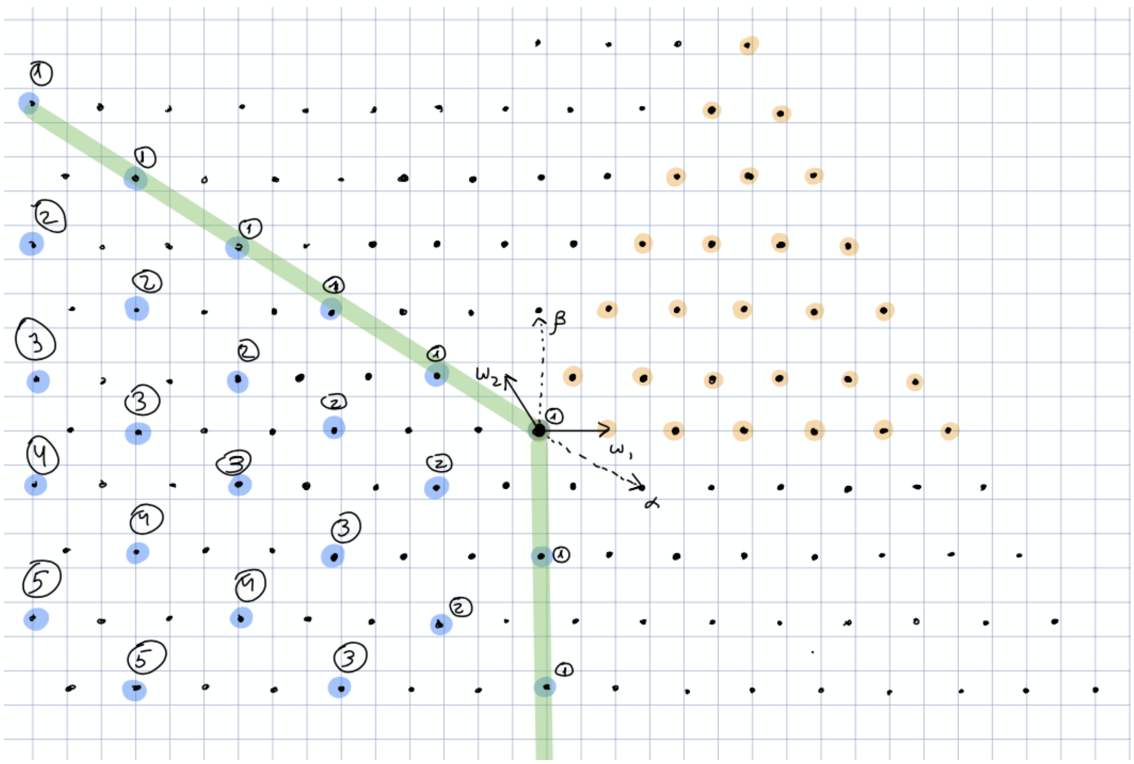
$$\omega_1(\mathrm{diag}(t_1, t_2, t_3)) := t_1, \quad \omega_2(\mathrm{diag}(t_1, t_2, t_3)) := t_2.$$

Then ω_1, ω_2 is a $\mathbb{Z}$-basis for $\mathrm{Ch}(T)$. In order to understand how to draw, we need to understand a W-invariant inner product on $\mathrm{Ch}(T)_{\mathbb{R}}$. Since the lengths of ω_1, ω_2 and $-(\omega_1 + \omega_2)$ are equal, we see that the angle between ω_1 and ω_2 should be $2\pi/3$. An element $\omega_1^{m_1}\omega_2^{m_2}$ is dominant if and only if $m_1 \geq m_2 \geq 0$. Let us also denote $\alpha := \alpha_{1,2}$ and $\beta := \alpha_{2,3}$. The product appearing in (3.1) is in our case seen to be:

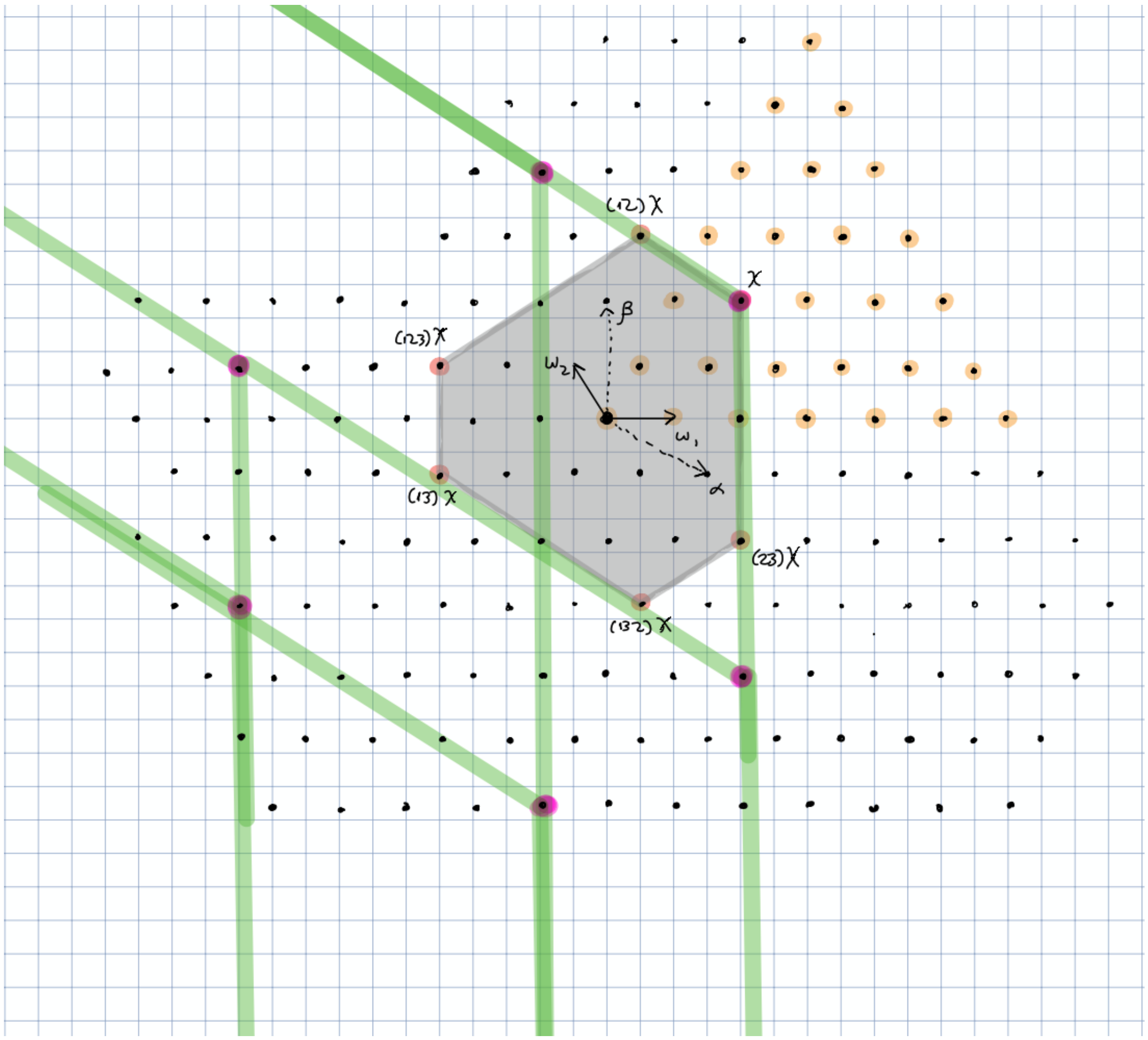
$$\left(1 + \alpha^{-1} + \alpha^{-2} + \ldots\right) \cdot \left(1 + \beta^{-1} + \beta^{-2} + \ldots\right) \cdot \left(1 + \alpha^{-1}\beta^{-1} + \alpha^{-2}\alpha^{-2} + \ldots\right) =$$

$$= \sum_{m_1, m_2 \geq 0} (1 + \min\{m_1, m_2\}) \cdot \alpha^{-m_1}\beta^{-m_2}. \tag{3.2}$$

We depict all this information as follows:



The black points represent elements of $\mathrm{Ch}(T)$. The yellowed points represent dominant elements. The blued elements represent elements appearing in (3.8), and I also put in a circle their multiplicity. The green lines form the boundary of the "cone-like" area defined by the blue points. Let us now choose some dominant character, for example $\chi := \omega_1^3\omega_2^2$. Consider the following illustration:



We drew the six points $w\chi$ for $w \in W$, and in grey we drew the convex hull of these points. The six points in violet are the points $w\chi \cdot \frac{w\sqrt{\Delta}}{\sqrt{\Delta}}$ for $w \in W$. We then in green emphasized the cones which appear in the formula. Thus, the formula has the sum of elements in three cones minus the difference of elements in three cones, and somehow everything cancels except from, potentially, things in the grey area.

3.9 Some further notes regarding Weyl's character formula

Exercise 3.8. *By plugging in 1 in place of χ in Weyl's character formula, obtain **Weyl's denominator formula***

$$\sum_{w\in W}\operatorname{sgn}(w)\cdot\sqrt{\Delta}(w^{-1}t)=\sqrt{\Delta}(t)\prod_{1\le i<j\le n}(1-\alpha_{i,j}(t)^{-1})$$

or, plugging in t^2 in place of t,

$$\sum_{w\in W}\operatorname{sgn}(w)\cdot\Delta(w^{-1}t)=\prod_{1\le i<j\le n}(\alpha_{i,j}(t)-\alpha_{i,j}(t)^{-1}).$$

Writing this concretely gives

$$\sum_{w\in W}\operatorname{sgn}(w)\cdot t_{w(1)}^{n-1}t_{w(2)}^{n-3}\cdot\ldots\cdot t_{w(n)}^{-(n-1)}=\prod_{1\le i<j\le n}\left(\frac{t_i}{t_j}-\frac{t_j}{t_i}\right).$$

Recall that we work under the condition $t_1\cdot\ldots\cdot t_n=1$, but it is immediate to see that our current equation bears multiplying all t_i's by some fixed t, hence we can drop the condition $t_1\cdot\ldots\cdot t_n=1$. Clearing denominators our equation is rewritten as

$$\sum_{w\in W}\operatorname{sgn}(w)\cdot(t_{w(1)}^2)^{n-1}(t_{w(2)}^2)^{n-2}\cdot\ldots\cdot(t_{w(n)}^2)^0=\prod_{1\le i<j\le n}\left(t_i^2-t_j^2\right).$$

*Setting $s_i:=t_i^2$, recognize this as the **Van-der-Monde determinant** equality*

$$\begin{vmatrix} s_1^{n-1} & s_1^{n-2} & \ldots & s_1 & 1\\ s_2^{n-1} & s_2^{n-2} & \ldots & s_2 & 1\\ \ldots & \ldots & \ldots & \ldots & \ldots\\ s_n^{n-1} & s_n^{n-2} & \ldots & s_n & 1\end{vmatrix}=\prod_{1\le i<j\le n}(s_i-s_j)$$

Exercise 3.9. *Given $\chi\in\operatorname{Ch}(T)$, we would like to find a formula for $\dim_{\mathbb{C}}V_\chi$. Notice that $\dim_{\mathbb{C}}V_\chi=\operatorname{ch}_{V_\chi}(1)$. We can assume that χ is dominant. Let us denote*

$$A_\chi(t):=\sum_{w\in W}\operatorname{sgn}(w)\cdot(w\chi)(t).$$

Weyl's character formula can be written, in view of Exercise 3.8, as

$$\operatorname{ch}_{V_\chi}(t)=\frac{A_{\chi\sqrt{\Delta}}(t)}{A_{\sqrt{\Delta}}(t)}.$$

Let us now denote by $m_1\ge m_2\ge\ldots\ge m_n$ a sequence such that

$$\chi(\operatorname{diag}(t_1,\ldots,t_n))=t_1^{m_1}\cdot\ldots\cdot t_n^{m_n}$$

and by $k_1 \geq k_2 \geq \ldots \geq k_n$ *a sequence such that*

$$\sqrt{\Delta}(\mathrm{diag}(t_1, \ldots, t_n)) = t_1^{k_1} \cdot \ldots \cdot t_n^{k_n}.$$

For $\tau \in \mathbb{C}^{\times}_{|-|=1}$*, we have*

$$A_\chi(\mathrm{diag}(\tau^{k_1}, \ldots, \tau^{k_n})) = \sum_{w \in W} \mathrm{sgn}(w) \cdot \tau^{k_{w(1)} m_1} \cdot \ldots \cdot \tau^{k_{w(n)} m_n} =$$

$$= \sum_{w \in W} \mathrm{sgn}(w) \cdot \tau^{k_1 m_{w^{-1}(1)}} \cdot \ldots \cdot \tau^{k_n m_{w^{-1}(n)}} = A_{\sqrt{\Delta}}(\mathrm{diag}(\tau^{m_1}, \ldots, \tau^{m_n})).$$

Therefore, noticing that for τ *close to* 1*, but not equal to it, the components of* $(\tau^{k_1}, \ldots, \tau^{k_n})$ *are pairwise non-equal, we have*

$$\mathrm{ch}_{V_\chi}(\mathrm{diag}(\tau^{k_1}, \ldots, \tau^{k_n})) = \frac{A_{\chi\sqrt{\Delta}}(\mathrm{diag}(\tau^{k_1}, \ldots, \tau^{k_n}))}{A_{\sqrt{\Delta}}((\mathrm{diag}(\tau^{k_1}, \ldots, \tau^{k_n}))} = \frac{A_{\sqrt{\Delta}}(\mathrm{diag}(\tau^{m_1+k_1}, \ldots, \tau^{m_n+k_n}))}{A_{\sqrt{\Delta}}(\mathrm{diag}(\tau^{k_1}, \ldots, \tau^{k_n}))} =$$

$$= \frac{\sqrt{\Delta}(\mathrm{diag}(\tau^{m_1+k_1}, \ldots, \tau^{m_n+k_n}))}{\sqrt{\Delta}(\mathrm{diag}(\tau^{k_1}, \ldots, \tau^{k_n}))} \cdot \prod_{1 \leq i < j \leq n} \frac{1 - \tau^{(m_j - m_i) + (k_j - k_i)}}{1 - \tau^{k_j - k_i}}.$$

Hence

$$\dim_{\mathbb{C}}(V_\chi) = \mathrm{ch}_{V_\chi}(1) = \lim_{\tau \to 1} \mathrm{ch}_{V_\chi}(\mathrm{diag}(\tau^{k_1}, \ldots, \tau^{k_n})) = \lim_{\tau \to 1} \prod_{1 \leq i < j \leq n} \frac{1 - \tau^{(m_j - m_i) + (k_j - k_i)}}{1 - \tau^{k_j - k_i}} = \ldots$$

Since $1 - \tau^N \sim N(1 - \tau)$ *as* $\tau \to 1$*, we can continue*

$$\ldots = \prod_{1 \leq i < j \leq n} \frac{(m_j - m_i) + (k_j - k_i)}{k_j - k_i} = \prod_{1 \leq i < j \leq n} \left(1 + \frac{m_i - m_j}{j - i}\right).$$

to conclude:

$$\dim_{\mathbb{C}} V_{d_1, \ldots, d_{n-1}} = \prod_{1 \leq i < j \leq n} \left(1 + \frac{d_i + \ldots + d_{j-1}}{j - i}\right).$$

For example, let us consider SU(2)*. We obtain*

$$\dim_{\mathbb{C}} V_d = 1 + d.$$

Considering SU(3)*, we obtain*

$$\dim_{\mathbb{C}} V_{d_1, d_2} = (1 + d_1)(1 + d_2)(1 + \frac{d_1 + d_2}{2}).$$

4 Manifolds, Lie groups and Lie algebras

4.1 Manifolds

To minimize background, we will only deal with embedded manifolds in this course. We assume that the reader knows, given f.d. $\mathbb{R}$-vector spaces E and F and open subsets $U \subset E$ and $V \subset F$, what a smooth map from U to V is. We also assume that given a smooth map $f : U \to V$ and a point $p \in U$, the reader knows what is the **differential** $D_p f : E \to F$ of f at the point p (it is an $\mathbb{R}$-linear map[14]). We recall that a smooth map $\phi : U \to V$ is a **diffeomorphism** if there exists a smooth map $\psi : V \to U$ such that $\phi \circ \psi = \mathrm{id}_V$ and $\psi \circ \phi = \mathrm{id}_U$.

Definition 4.1.

- An **embedded manifold** is a pair (E, M) consisting of a f.d. $\mathbb{R}$-vector space E and a subset $M \subset E$, such that for every $p \in M$ there exist $0 \leq m \leq \dim_{\mathbb{R}} E$, an open subset $p \in U \subset E$, an open subset $0 \in V \subset \mathbb{R}^n$ and a diffeomorphism $\phi : V \to U$, such that

  $$\phi^{-1}(M) = \{(x_1, \ldots, x_n) \in V \mid x_{m+1} = \ldots = x_n = 0\}.$$

  To abbreviate, we will often-times call an embedded manifold simply a **manifold**, and will write M instead of (E, M) (i.e., E is implicit). We also say that M is a **manifold embedded into** E.

- Let (E_1, M_1) and (E_2, M_2) be embedded manifolds. A **morphism of manifolds** (or simply a **smooth map**) from (E_1, M_1) to (E_2, M_2) is a map $\phi : M_1 \to M_2$ satisfying the following condition: For every f.d. $\mathbb{R}$-vector space F and every open subset $U \subset F$ and every map $\psi : U \to M_1$ such that the composition $U \xrightarrow{\psi} M_1 \subset E_1$ is smooth, we have that the composition $U \xrightarrow{\psi} M_1 \xrightarrow{\phi} M_2 \subset E_2$ is smooth. An isomorphism of manifolds is also called a **diffeomorphism**.

Definition 4.2. Let M be a manifold. A smooth map $M \to \mathbb{C}$ is called a **smooth function** on M and we denote by $C^\infty(M)$ the $\mathbb{C}$-vector space of smooth functions on M.

Remark 4.3. Given an embedded manifold (E, M), we always treat M as a topological space, with the subspace topology induced by the inclusion $M \subset E$.

Example 4.4.

- *(E, U) is an embedded manifold whenever U is an open subset in E.*

- *Given an embedded manifold (E, M) and an open subset $U \subset M$, (E, U) is also an embedded manifold.*

[14] $D_p f$ is the unique $\mathbb{R}$-linear map from E to F satisfying $f(p+x) - f(p) - (D_p f)(x) = o(||x||)$ as $x \to 0$.

- *(E, D) is an embedded manifold whenever D is a discrete subset in E.*
- *Given embedded manifolds (E_1, M_1) and (E_2, M_2), we have an embedded manifold $(E_1 \times E_2, M_1 \times M_2)$.*

Claim 4.5 (Implicit function theorem)**.** *Let E and F be f.d. $\mathbb{R}$-vector spaces. Let $U \subset E$ be an open subset. Let $q \in F$. Let $\phi : U \to F$ be a smooth map. Assume that $D_p\phi : E \to F$ is surjective for every $p \in \phi^{-1}(q)$. Then $(E, \phi^{-1}(q))$ is an embedded manifold.*

Proof. Omitted. □

Example 4.6. *$(\mathbb{R}^n, S^{n-1})$ is an embedded manifold. Indeed, let us consider the smooth map $f : \mathbb{R}^n \to \mathbb{R}$ given by $(x_1, \ldots, x_n) \mapsto x_1^2 + \ldots + x_n^2$. Then $S^{n-1} = f^{-1}(1)$. So, by Claim 4.5 we will know that $(\mathbb{R}^n, S^{n-1})$ is an embedded manifold if we will show that $D_p f : \mathbb{R}^n \to \mathbb{R}$ is non-zero (and hence surjective) for every $p \in S^{n-1}$. The matrix representing this $D_p f$ in the standard basis is $(2x_1, \ldots, 2x_n)$, where $p = (x_1, \ldots, x_n)$. Clearly $(2x_1, \ldots, 2x_n) \neq 0$ if and only if $p \neq 0$, in particular when $p \in S^{n-1}$.*

Example 4.7. *The requirement of surjectivity in Claim 4.5 is necessary. For example, we can consider $\phi : \mathbb{R}^2 \to \mathbb{R}$ given by $(x, y) \mapsto xy$. Then $(\mathbb{R}^2, \phi^{-1}(0))$ is not an embedded manifold.*

Remark 4.8. One should be careful that, for example, a bijective smooth map is not necessarily a diffeomorphism. A standard example is the map $\mathbb{R} \to \mathbb{R}$ given by $x \mapsto x^3$.

If (E, M) is an embedded manifold and $N \subset M$ is an open subset, then as we mentioned above (E, N) is an embedded manifold - we say that N is an **open submanifold** of M. If $N \subset M$ is a closed subset, then in general (E, N) is not an embedded manifold. If it is, we say that N is a **closed submanifold** of M. One can show that if $N \subset M$ is a subset such that (E, N) is an embedded manifold then N is **locally closed** in M, meaning that there exists an open subset $U \subset M$ such that $N \subset U$ and $N \cap U$ is closed in U. If $N \subset M$ is such a locally closed submanifold, then one sees that the inclusion map $N \to M$ is smooth and given a manifold L and a smooth map $L \to M$ whose image lies in N, its corestriction $L \to N$ is a smooth map.

Remark 4.9. Let M be a manifold and let $p \in M$. Then there exist an open $p \in U \subset M$, an open $0 \in V \subset \mathbb{R}^n$ (for some $n \in \mathbb{Z}_{\geq 0}$) and a diffeomorphism of U and V. In other words, manifolds "look locally" like open subsets in Euclidean spaces.

Exercise 4.1. *Let (E, M) be an embedded manifold. Let $f \in C^\infty(M)$ and let $p \in M$. Then there exists an open subset $p \in U \subset E$ and $\widetilde{f} \in C^\infty(U)$ such that $\widetilde{f}|_{U \cap M} = f|_{U \cap M}$.*

4.2 Tangent spaces and tangent maps

Recall directional derivatives:

Definition 4.10. Let E be a f.d. $\mathbb{R}$-vector space. Let $p \in E$ and let f be a smooth function on an open neighbourhood U of p in E. Given $v \in E$ the **directional derivative of f at p in the direction of v** is:

$$\partial_v f := \lim_{\epsilon \to 0} \frac{1}{\epsilon}(f(p + \epsilon \cdot v) - f(p)) \in \mathbb{C}.$$

Exercise 4.2. *Recall that the assignment $v \mapsto \partial_v f$ is $\mathbb{R}$-linear.*

Definition 4.11. Let (E, M) be an embedded manifold and let $p \in M$.

- A vector $v \in E$ is said to be a **tangent vector to M at p** if for every smooth function f on an open neighbourhood U of p in E satisfying $f|_{M \cap U} = 0$ we have $\partial_v f = 0$.
- The **tangent space to M at p** is the set of tangent vectors to M at p, which is an $\mathbb{R}$-linear subspace of E. It is denoted $T_p M$.

Example 4.12. *Let (E, U) be an embedded manifold with U open in E. Then for every $p \in U$ we have $T_p U = E$.*

Remark 4.13. Let (E, M) be an embedded manifold and let $p \in M$. A "geometric" description of $T_p M$ is as follows. A vector $0 \neq v \in E$ lies in $T_p M$ if and only if we can find a sequence p_n of points in M such that $p_n \to p$ and $\frac{1}{||p_n - p||}(p_n - p) \to \frac{1}{||v||} v$. Or, a vector $v \in E$ lies in $T_p M$ if and only if we can find a smooth map $g : (-\epsilon, \epsilon) \to M$ such that $\lim_{t \to 0} \frac{1}{t}(g(t) - g(0)) = v$.

Claim 4.14. *Let E and F be f.d. $\mathbb{R}$-vector spaces. Let $U \subset E$ be an open subset. Let $q \in F$. Let $\phi : U \to F$ be a smooth map. Assume that $D_p \phi : E \to F$ is surjective for every $p \in \phi^{-1}(q)$. Denote $M := \phi^{-1}(q)$ (we saw that M is a manifold embedded into E). Then, for every $p \in M$, $T_p M = \mathrm{Ker}(D_p f)$.*

Proof. Omitted. □

Example 4.15. *Consider $(\mathbb{R}^n, S^{n-1})$. The tangent space to S^{n-1} at $(1, 0, \ldots, 0)$ is $\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_1 = 0\}$.*

Example 4.16. *The requirement of surjectivity in Claim 4.14 is necessary. For example, consider $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) := x^2$. Then $f^{-1}(0) = \{0\}$ is a submanifold in $\mathbb{R}$. We have $T_0(f^{-1}(0)) = 0$ but $D_0 f = 0$ and so $\mathrm{Ker}(D_0 f) = \mathbb{R}$.*

Using Exercise 4.1 we see that given a manifold M, a point $p \in M$, $f \in C^\infty(M)$ and $v \in T_p M$ we can define unambiguously $\partial_v f \in \mathbb{C}$, by locally extending f to a function $\widetilde{f}$ on an open neighbourhood of p in E and defining $\partial_v f$ to be $\partial_v \widetilde{f}$. The assignment

$$T_p M \times C^\infty(M) \to \mathbb{R}, \quad (v, f) \mapsto \partial_v(f)$$

is $\mathbb{R}$-bilinear and satisfies the **Leibnitz rule** $\partial_v(f_1 f_2) = \partial_v(f_1) f_2(p) + f_1(p) \partial_v(f_2)$.

Proposition-Definition 4.17. *Let M_1 and M_2 be manifolds. Let $\phi : M_1 \to M_2$ be a smooth map and let $p \in M_1$. There exists a unique $\mathbb{R}$-linear map $D_p\phi : T_pM_1 \to T_{\phi(p)}M_2$ satisfying:*

- *For every smooth function f on an open neighbourhood of $\phi(p)$ in M_2 and every $v \in T_pM_1$, we have $\partial_v(f \circ \phi) = \partial_{(D_p\phi)(v)}(f)$.*

We call it the ***tangent map of*** *ϕ* ***at*** *p (or* ***differential of*** *ϕ* ***at*** *p).*

Exercise 4.3. *Let M_1,M_2 and M_3 be manifolds and let $\phi : M_1 \to M_2$ and $\psi : M_2 \to M_3$ be smooth maps. Let $p \in M_1$. Then $D_{\phi(p)}\psi \circ D_p\phi = D_p(\psi \circ \phi)$.*

Remark 4.18. Let E be a f.d. $\mathbb{R}$-vector space and let $U \subset E$ be an open subset. Let (F, M) be an embedded manifold. Let $\phi : E \to M$ be a smooth map. Let $p \in U$. Then $D_p\phi : E \to T_{\phi(p)}M \subset F$ is the usual differential. Namely, we for example can characterize

$$(D_p\phi)(v) := \lim_{t\to 0} \frac{1}{t}(\phi(p + tv) - \phi(p)).$$

Remark-Definition 4.19. Let $I \subset \mathbb{R}$ be an open subset (viewed as an embedded manifold $(\mathbb{R}, I)$). Given $t \in I$, we have $T_tI = \mathbb{R}$. Given a manifold M and a smooth map $\phi : I \to M$, we therefore have the tangent map $D_t\phi : \mathbb{R} \to T_{\phi(t)}M$. Let us denote in such a situation $d_t\phi := (D_t\phi)(1)$ and call it the **derivative of** ϕ **at** t. The information of the $\mathbb{R}$-linear map $D_t\phi : \mathbb{R} \to T_{\phi(t)}M$ is the same as the information of the vector $d_t\phi \in T_{\phi(t)}M$. If M is an embedded manifold (E, M), we have:

$$d_t\phi = \lim_{s\to 0} \frac{1}{s}(\phi(t + s) - \phi(t)).$$

Exercise 4.4. *Let $I \subset \mathbb{R}$ be an open subset and M and N manifolds. Let $\phi : I \to M$ and $\psi : M \to N$ be smooth maps. Show that, for any $t \in I$, we have $(D_{\phi(t)}\psi)(d_t\phi) = d_t(\psi \circ \phi)$.*

Remark 4.20. Let (E_1, M_1) and (E_2, M_2) be embedded manifolds and let $\phi : M_1 \to M_2$ be a smooth map. Let $p \in M_1$. A "geometric" description of $D_p\phi : T_pM_1 \to T_pM_2$ is as follows. Given $v \in T_pM_1$ we find a smooth map $g : (-\epsilon, \epsilon) \to M_1$ such that $v = d_0g$. Then $(D_p\phi)(v) = d_0(\phi \circ g)$.

The following theorem gives the basic understanding of smooth maps with surjective differential:

Theorem 4.21. *Let M and N be manifolds, let $p \in M$, let $\phi : M \to N$ be a smooth map, and assume that $D_p\phi : T_pM \to T_{\phi(p)}N$ is surjecitve. Then there exist $n, m \in \mathbb{Z}_{\geq 0}$, open subsets $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ open subsets $p \in U' \subset M$, $\phi(p) \in U'' \subset N$, diffeomorphisms $\alpha : U'' \to V$ and $\beta : U' \to U \times V$ such that $\phi(U') \subset U''$ and we have $\alpha(\phi(m)) = \pi(\beta(m))$ for all $m \in U'$, where $\pi : U \times V \to V$ denotes the projection onto the second variable.*

Remark 4.22. In other words, at a neighbourhood of a point where the differential is surjective, a smooth map look like a projection.

Let us state the inverse function theorem:

Theorem 4.23 (Inverse function theorem)**.** *Let M and N be manifolds. Let $\phi : M \to N$ be a smooth map, and let $p \in M$. Suppose that $D_p\phi : T_pM \to T_{\phi(p)}N$ is invertible. Then there exist open subsets $p \in U \subset M$ and $\phi(p) \in V \subset N$ such that $\phi(U) \subset V$ and the smooth map $\phi|_U : U \to V$ is a diffeomorphism.*

Proof. This is easy to prove given the previous theorem. □

4.3 Vector fields and flows

Definition 4.24. Let (E, M) be an embedded manifold. A **vector field on** M is a smooth map $\xi : M \to E$ such that for every $p \in M$ we have $\xi(p) \in T_pM$.

Remark 4.25. So, informally a vector field on M is a collection $(v_p)_{p\in M}$ with $v_p \in T_pM$ which "varies smoothly with p" (here we denoted v_p for the above $\xi(p)$).

Definition 4.26. Let M be a manifold and let ξ be a vector field on M. Let $I \subset \mathbb{R}$ be an open interval and let $\phi : I \to M$ be a smooth map. We say that ϕ **obeys** ξ if for every $t \in I$ we have $d_t\phi = \xi(\phi(t))$.

The basic theorem of the subject of ordinary differential equations is:

Theorem 4.27 (Uniqueness and existence of solutions of ODE's)**.** *Let M be a manifold and let ξ be a vector field on M. Let $p \in M$.*

- *(uniqueness) Let $I \subset \mathbb{R}$ be an open interval and let $t_0 \in I$. Let $\phi, \psi : I \to M$ be two smooth maps obeying ξ and satisfying $\phi(t_0) = p$ and $\psi(t_0) = p$. Then $\phi = \psi$.*
- *(existence) Let $t_0 \in \mathbb{R}$. There exists $r > 0$ and a smooth map $\phi : (t_0 - r, t_0 + r) \to M$ obeying ξ such that $\phi(t_0) = p$.*

Proof. Omitted. □

We also want to know that the solution in the previous theorem "varies smoothly with a smooth parameter". For that, we first define:

Definition 4.28. Let N be a manifold and (E, M) an embedded manifold. An **N-parametrized vector field on** M is a smooth map $\xi : N \times M \to E$ such that for every $(q, p) \in N \times M$ we have $\xi(q, p) \in T_pM$.

And now we can state:

Theorem 4.29 (smooth dependence of solutions of ODE's on parameters)**.** *Let N and M be manifolds. Let ξ be an N-parametrized vector field on M and let $p \in M$. Let $I \subset \mathbb{R}$ be an open interval and let $t_0 \in I$. Let us be given, for every $n \in N$, a smooth map $\phi_n : I \to M$ obeying ξ_n (where $\xi_n(m) := \xi(n, m)$) and satisfying $\phi_n(t_0) = p$. Define a map $\phi : N \times I \to M$ by $\phi(n, t) := \phi_n(t)$. Then ϕ is smooth.*

Proof. Omitted. □

4.4 Lie groups

Definition 4.30.

- A **Lie group** is a manifold G equipped with a group structure, such that the multiplication map $G \times G \to G$ and the inverse map $G \to G$ are smooth maps.
- Let G_1 and G_2 be Lie groups. A **morphism of Lie groups** from G_1 to G_2 is a map $\phi : G_1 \to G_2$ which is both a smooth map and a group morphism.

Example 4.31.

- $\mathbb{R}$ *(a manifold embedded into* $\mathbb{R}$*) with addition is a Lie group. Similarly,* $\mathbb{C}$ *is a Lie group.*
- $\mathbb{R}^\times$ *(a manifold embedded into* $\mathbb{R}$*) with multiplication is a Lie group. Similarly,* $\mathbb{C}^\times$ *is a Lie group.*
- $\mathrm{GL}_n(\mathbb{R})$ *(a manifold embedded into* $M_n(\mathbb{R})$*) with matrix multiplication is a Lie group. Similarly,* $\mathrm{GL}_n(\mathbb{C})$ *(a manifold embedded into* $M_n(\mathbb{C})$*) is a Lie group.*

We have the following theorem; we will prove it later.

Theorem 4.32.

1. *Let G be a Lie group. Let $H \subset G$ be a closed subgroup. Then H is a closed submanifold of G, and therefore H is a Lie group itself (we say that H is a* ***closed Lie subgroup*** *of G).*
2. *("automatic smoothness") Let G and H be Lie groups. Let $\phi : G \to H$ be a morphism of topological groups. Then ϕ is a morphism of Lie groups (i.e. it is smooth).*

Proof. Given in §4.8. □

Corollary 4.33. *All the closed subgroups of* $\mathrm{GL}_n(\mathbb{R})$ *and* $\mathrm{GL}_n(\mathbb{C})$ *we considered (such as* $\mathrm{SU}(n), \mathrm{SL}_n(\mathbb{R})$ *etc.) are Lie groups*[15].

Remark 4.34. A consequence of part 2 of Theorem 4.32 is that, given a Lie group G and a representation of G on a f.d. $\mathbb{C}$-vector space V, the corresponding morphism of topological group $\pi : G \to \mathrm{GL}_{\mathbb{C}}(V)$ is automatically a morphism of Lie groups. In other words, defining in the obvious way smooth f.d. G-representations, we see that in fact there is no difference between those and our previously defined ("continuous") f.d. G-representations.

[15] Of course one can also check these directly.

4.5 The exponential map

Definition 4.35. Let (E, G) be an embedded Lie group. For $g \in G$, let us denote by $m_g : G \to G$ the diffeomorphism given by $x \mapsto gx$. One can check that the map $\lambda : T_1G \times G \to E$ given by $(X, g) \mapsto (D_1 m_g)(X)$ is smooth and so defines an T_1G-parametrized vector field on G. Given $X \in T_1G$ we denote by λ_X the vector field on G given by $\lambda(X, -)$.

Claim 4.36. *Let G be a Lie group and let $X \in T_1G$. There exists a unique morphism of Lie groups $e_X : \mathbb{R} \to G$ satisfying $d_0 e_X = X$. It can also be characterized as the unique smooth map $e_X : \mathbb{R} \to G$ obeying λ_X and satisfying $e_X(0) = 1$.*

Proof. Given $r > 0$ denote by $e^r_X : (-r, r) \to G$ the smooth map obeying λ_X and satisfying $e^r_X(0) = 1$, if it exists (it is unique if exists, by the uniqueness part of Theorem 4.27). In particular, we can talk about $e_X := e^\infty_X$ (if it exists).

Let us first see that if for a given $r > 0$ the map e^r_X exists, given $-r < t, s < r$ such that $-r < t + s < r$ we have $e^r_X(t)e^r_X(s) = e^r_X(t + s)$. Put differently, we fix $-r < t < r$, and we want to show that given $s \in (-r', r'')$, where $r' := \min\{r, r + t\}$ and $r'' := \min\{r, r - t\}$, we have $e^r_X(t + s) = e^r_X(t)e^r_X(s)$. Let us denote by $f : (-r', r'') \to G$ the map $f(s) := e^r_X(t)^{-1}e^r_X(t + s)$. We want to see that $f(s) = e^r_X(s)$ for all $s \in (-r', r'')$. By the uniqueness part of Theorem 4.27 it is enough to show that f obeys λ_X and that $f(0) = 1$. The second is clear. As for the first, we notice that $f : (-r', r'') \to G$ can be written as the composition

$$(-r', r'') \xrightarrow{s \mapsto s+t} (-r, r) \xrightarrow{e^r_X} G \xrightarrow{m_{e^r_X(t)^{-1}}} G.$$

Computing using this we see that $d_s f = \lambda_X(f(s))$ for all $s \in (-r', r'')$.

Now, let again $r > 0$ be such that the map e^r_X exists. We want to show that $e^{r+r/2}_X$ exists as well. Define $f^+ : (r/2, r + r/2) \to G$ by $f^+(t) := e^r_X(r/2)e^r_X(t - r/2)$. Similarly, define $f^- : (-r - r/2, -r/2) \to G$ by $f^-(t) := e^r_X(-r/2)e^r_X(t + r/2)$. Notice that for $t \in (r/2, r)$ we have $f^+(t) = e^r_X(r/2)e^r_X(t - r/2) = e^r_X(t)$ in view of the established additivity property above. Similarly $f^-(t) = e^r_X(t)$ for $t \in (-r, -r/2)$. Thus, we can patch f^-, e^r_X, f^+ into one smooth map $h : (-r - r/2, r + r/2) \to G$. It is left as a small exercise to see that h suits the conditions to be $e^{r+r/2}_X$.

Notice that e^r_X exists for some $r > 0$, by the existence part of Theorem 4.27. Using the above, we see that $e^{r'}_X$ then exists for r' as large as wanted, and therefore $e_X := e^\infty_X$ exists (by patching $e^{r'}_X$'s for r' bigger and bigger). By the additivity property above we deduce that e_X is a morphism of Lie groups, and clearly $d_0 e_X = X$.

It is left to see that given a morphism of Lie groups $f : \mathbb{R} \to G$ which satisfies $d_0 f = X$, we have $f = e_X$. Of course $f(0) = 1$. Let $t \in \mathbb{R}$. Let us notice that $f : \mathbb{R} \to G$ is equal to the composition

$$\mathbb{R} \xrightarrow{s \mapsto s-t} \mathbb{R} \xrightarrow{f} G \xrightarrow{m_{f(t)}} G.$$

Computing using this, we find

$$d_t f = \lambda_X(f(t)).$$

Therefore we conclude that $f = e_X$.

□

Definition 4.37. Let G be a Lie group. The **exponential map** $\exp : T_1G \to G$ (or $\exp_G$ if we want to emphasize G) is defined as sending X to $e_X(1)$, where e_X is as in Claim 4.36.

Exercise 4.5. *Let G be a Lie group and let $X \in T_1G$. Show that $e_X(t) = \exp(tX)$ for all $t \in \mathbb{R}$.*

Exercise 4.6.

- *Check that for the Lie group $\mathbb{R}$, the exponential map is the identity map $\mathbb{R} \to \mathbb{R}$.*
- *Check that for the Lie group $\mathbb{R}^\times$, the exponential map is the usual exponential map $\mathbb{R} \to \mathbb{R}^\times$.*
- *Check that for the Lie group $\mathbb{C}^\times_{|-|=1}$, the exponential map is the map $i\mathbb{R} \to \mathbb{C}^\times_{|-|=1}$ given by $u \mapsto e^u$.*

Example 4.38. *Let $G := \mathrm{GL}_n(\mathbb{R})$. We have $T_1G = M_n(\mathbb{R})$. Recall that we have the exponentiation of matrices $M_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$ given by*

$$X \mapsto e^X := I + X + \frac{1}{2!}X^2 + \dots.$$

Let $X \in M_n(\mathbb{R})$. Then the map $\phi : \mathbb{R} \to \mathrm{GL}_n(\mathbb{R})$ given by $t \mapsto e^{tX}$ is a morphism of Lie groups, satisfying $d_0\phi = X$. Hence we have $\exp_{\mathrm{GL}_n(\mathbb{R})}(X) = e^X$.

Lemma 4.39. *Let G be a Lie group. The exponential map $\exp : T_1G \to G$ is smooth, and we have $D_0 \exp = \mathrm{Id}_{T_1G}$.*

Proof. Let us consider the map $ee : T_1G \times \mathbb{R} \to G$ given by $(X, t) \mapsto e_X(t)$. It obeys the T_1G-parametrized vector field λ on G and therefore by Theorem 4.29 we see that ee is smooth. Therefore the composition $T_1G \xrightarrow{X \mapsto (X,1)} T_1G \times \mathbb{R} \to G$ is smooth. Notice that this composition is exp.

As regarding $D_0 \exp$, fix $X \in T_1G$. Consider the smooth map $\phi_X : \mathbb{R} \to T_1G$ given by $\phi_X(t) := tX$. Then

$$X = d_0 e_X = d_0(\exp \circ \phi_X) = (D_0 \exp)(X),$$

as desired.

□

Exercise 4.7. *Let* $\phi : H \to G$ *be a morphism of Lie groups. Then we have*

$$\phi \circ e_X = e_{(D_1\phi)(X)} \quad \forall X \in T_1 H$$

and

$$\phi \circ \exp_H = \exp_G \circ (D_1\phi).$$

Claim 4.40. *Let* G *be a Lie group and let* $H \subset G$ *be a closed Lie subgroup. Then given* $X \in T_1 G$ *we have* $X \in T_1 H$ *if and only if* $\exp_G(tX) \in H$ *for all* $t \in \mathbb{R}$*.*

Proof. Suppose that $X \in T_1 H$. Denoting the inclusion $i : H \to G$ (it is a morphism of Lie groups), notice that $D_1 i$ is the inclusion of $T_1 H$ in $T_1 G$, and we have

$$\exp_G(tX) = \exp_G((D_1 i)(tX)) = i(\exp_H(tX)) \in H.$$

Conversely, let $X \in T_1 G$ and suppose that $\exp_G(tX) \in H$ for all $t \in \mathbb{R}$. Define a smooth map $\phi : \mathbb{R} \to H$ by $\phi(t) := \exp_G(tX)$. Then

$$(D_1 i)(d_0\phi) = d_0(i \circ \phi) = X$$

so X is in the image of the inclusion $D_1 i : T_1 H \to T_1 G$, i.e. $X \in T_1 H$. $\square$

Example 4.41. *Let us consider* $\mathrm{SL}_n(\mathbb{R}) \subset \mathrm{GL}_n(\mathbb{R})$*. Given* $X \in M_n(\mathbb{R})$*, by Claim 4.40 we have* $X \in T_1\mathrm{SL}_n(\mathbb{R})$ *if and only if* $\exp(tX) \in \mathrm{SL}_n(\mathbb{R})$ *for all* $t \in \mathbb{R}$*. Recall that we have* $\det(\exp(Y)) = \exp(\mathrm{tr}(Y))$ *for all* $Y \in M_n(\mathbb{R})$*. Therefore*

$$\exp(tX) \in \mathrm{SL}_n(\mathbb{R})\ \forall t \iff \det(\exp(tX)) = 1\ \ \forall t \iff \exp(t{\cdot}\mathrm{tr}(X)) = 1\ \forall t \iff \mathrm{tr}(X) = 0.$$

Thus $T_1\mathrm{SL}_n(\mathbb{R}) = \{X \in M_n(\mathbb{R}) \mid \mathrm{tr}(X) = 0\}$*.*

4.6 The Lie algebra of a Lie group

Let G be a Lie group. Since, by Lemma 4.39, $\exp : T_1 G \to G$ has invertible differential at $0 \in T_1 G$, by Theorem 4.23 we obtain that there exist open subsets $0 \in U \subset T_1 G$ and $1 \in V \subset G$ such that $\exp(U) \subset V$ and $\exp|_U : U \to V$ is a diffeomorphism. By slight abuse of notation, let us denote by $\exp^{-1} : V \to U$ the smooth map which is inverse to $\exp|_U : U \to V$. Thus, we think of G, at the vicinity of 1, and $T_1 G$, at the vicinity of 0, as identified. What is the relation between the additive group structure on $T_1 G$ (which is "simple") and the group structure on G (which is, generally speaking, "complicated")? To compare, let us "transport" the group structure on G to $T_1 G$, locally near 0. We have an open subset $0 \in U_1 \subset U$ such that $\exp(U_1) \cdot \exp(U_1) \subset \exp(U) = V$. We define $m : U_1 \times U_1 \to T_1 G$ by

$$m(X, Y) := \exp^{-1}(\exp(X) \cdot \exp(Y)).$$

Up to first approximation, there is no difference between the two group structures:

Lemma 4.42. *We have* $(D_{(0,0)}m)(X,Y) = X+Y$*. In other words,*

$$m(X,Y) = X + Y + o(||(X,Y)||)$$

as $(X,Y) \to (0,0)$*.*

Proof. Notice that the composition

$$U_1 \xrightarrow{X\mapsto(X,0)} U_1 \times U_1 \xrightarrow{m} T_1G$$

is equal to $X \mapsto X$. Therefore, taking the tangent map at 0, we obtain that the composition

$$T_1G \xrightarrow{X\mapsto(X,0)} T_1G \oplus T_1G \xrightarrow{D_{(0,0)}m} T_1G$$

is equal to $X \mapsto X$. In other words, $(D_{(0,0)}m)(X,0) = X$ for all $X \in T_1G$. Completely symmetrically we have $(D_{(0,0)}m)(0,Y) = Y$ for all $Y \in T_1G$. Thus

$$(D_{(0,0)}m)(X,Y) = (D_{(0,0)}m)(X,0) + (D_{(0,0)}m)(0,Y) = X + Y.$$

□

The Lie algebra concept appears when we consider the second approximation. From multivariable calculus, there exists a unique $\mathbb{R}$-bilinear symmetric map

$$B : (T_1G \oplus T_1G) \times (T_1G \oplus T_1G) \to T_1G,$$

such that

$$m(X,Y) = X + Y + \frac{1}{2}B((X,Y),(X,Y)) + o(||(X,Y)||^2)$$

as $(X,Y) \to (0,0)$.

Lemma 4.43. *We have* $B((X_1,0),(X_2,0)) = 0$ *for all* $X_1, X_2 \in T_1G$ *and* $B((0,Y_1),(0,Y_2)) = 0$ *for all* $Y_1, Y_2 \in T_1G$*.*

Proof. The second claim is analogous to the first, so let us show just the first. By the polarization identity, it is enough to see that $B((X,0),(X,0)) = 0$ for all $X \in T_1G$. We have

$$m(X,0) = X + \frac{1}{2}B((X,0),(X,0)) + o(||X||^2)$$

and on the other hand

$$m(X,0) = \exp^{-1}(\exp(X)\cdot\exp(0)) = \exp^{-1}(\exp(X)\cdot 1) = \exp^{-1}(\exp(X)) = X.$$

Comparing, we obtain

$$B((X,0),(X,0)) = o(||X||^2),$$

forcing the desired. □

Thus, we have

$$m(X,Y) = X + Y + B((X,0),(0,Y)) + o(||(X,Y)||^2)$$

as $(X,Y) \to (0,0)$. Let us define an $\mathbb{R}$-bilinear map $C : T_1G \times T_1G \to T_1G$ by $C(X,Y) := 2B((X,0),(0,Y))$. So, we have

$$m(X,Y) = X + Y + \frac{1}{2}C(X,Y) + o(||(X,Y)||^2) \tag{4.1}$$

as $(X,Y) \to (0,0)$.

Claim 4.44. *We have:*

- *(alternativity)* $C(X,X) = 0$ *for all* $X \in T_1G$.
- *(Jacobi identity)* $C(X, C(Y,Z)) = C((X,Y),Z) + C(Y, C(X,Z))$ *for all* $X,Y,Z \in T_1G$.

Proof. Let us show alternativity. It is enough to check it for X close to 0. We have then

$$m(X,-X) = \exp^{-1}(\exp(X) \cdot \exp(-X)) = \exp^{-1}(1) = 0.$$

Plugging this into (4.1) we get

$$0 = C(X,-X) + o(||X||^2)$$

as $X \to 0$, i.e.

$$C(X,X) = o(||X||)^2$$

as $X \to 0$. This implies that $X \mapsto C(X,X)$ is equal to zero, as desired.

The Jacobi identity is a little bit more complicated to establish. We will do it in the next subsection. □

One defines:

Definition 4.45. Let k be a field.

- A **Lie algebra over** k (or a k**-Lie algebra**) is a k-vector space $\mathfrak{g}$ equipped with a k-bilinear map

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

  satisfying:

  1. (alternativity) $[X,X] = 0$ for all $X \in \mathfrak{g}$.
  2. (Jacobi identity) $[X,[Y,Z]] = [[X,Y],Z] + [Y,[X,Z]]$ for all $X,Y,Z \in \mathfrak{g}$.

  The map $[-,-]$ is called the **Lie bracket**. One usually abuses notation and denotes $[-,-]$ in the same way for different Lie algebras.

- Let $\mathfrak{g}_1, \mathfrak{g}_2$ be Lie algebras over k. A **morphism of Lie algebras** (or more precisely a **morphism of k-Lie algebras**) from $\mathfrak{g}_1$ to $\mathfrak{g}_2$ is a k-linear map $\alpha : \mathfrak{g}_1 \to \mathfrak{g}_2$ satisfying

$$\alpha([X,Y]) = [\alpha(X), \alpha(Y)], \quad \forall X, Y \in \mathfrak{g}_1.$$

Thus, given our Lie group G, we have the structure of an $\mathbb{R}$-Lie algebra on T_1G, with C being the Lie bracket. But, as we remarked, we will always denote it by $[-,-]$, i.e. $[X,Y] := C(X,Y)$.

Definition 4.46. Let G be a Lie group. We denote the $\mathbb{R}$-Lie algebra which is T_1G equipped with the Lie bracket described above by $\mathrm{Lie}(G)$.

So, to repeat, we have the smooth map $\exp : \mathrm{Lie}(G) \to G$ which is a diffeomorphism onto the open image at a neighbourhood of $0 \in \mathrm{Lie}(G)$ and the Lie bracket $[-,-] : \mathrm{Lie}(G) \times \mathrm{Lie}(G) \to \mathrm{Lie}(G)$ is characterized as the unique $\mathbb{R}$-bilinear map satisfying

$$\exp(X)\exp(Y) = \exp(X + Y + \frac{1}{2}[X,Y] + o(||(X,Y)||^2))$$

when $(X,Y) \to (0,0)$.

Lemma-Definition 4.47. Let $\phi : H \to G$ be a morphism of Lie groups. Then $D_1\phi : T_1H \to T_1G$ is a morphism of $\mathbb{R}$-Lie algebras. We denote it by $\mathrm{Lie}(\phi) : \mathrm{Lie}(H) \to \mathrm{Lie}(G)$.

Proof. Let us abbreviate $T := D_1\phi$. We want to check that

$$[T(X), T(Y)] = T([X,Y])$$

for all $X, Y \in T_1H$. We have

$$\exp(X)\exp(Y) = \exp(X + Y + \frac{1}{2}[X,Y] + o(||(X,Y)||^2)).$$

Applying ϕ we obtain:

$$\exp(T(X))\exp(T(Y)) = \exp(T(X) + T(Y) + \frac{1}{2}T([X,Y]) + o(||(X,Y)||^2)).$$

On the other hand, we have

$$\exp(T(X))\exp(T(Y)) = \exp(T(X) + T(Y) + \frac{1}{2}[T(X),T(Y)] + o(||(X,Y)||^2)).$$

For small enough X, Y we can therefore compare and obtain $T([X,Y]) - [T(X),T(Y)] = o(||(X,Y)||^2)$, implying that the $\mathbb{R}$-bilinear map $(X,Y) \mapsto T([X,Y]) - [T(X),T(Y)]$ must be zero, as desired. ∎

Example 4.48. *Let us consider* $G := \mathrm{GL}_n(\mathbb{R})$*. Recall that* $T_1G = M_n(\mathbb{R})$ *and that* $\exp : M_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$ *is given by the usual exponentiation of matrices. We start with*

$$\exp(X)\exp(Y) = \exp(X + Y + \frac{1}{2}[X,Y] + o(||(X,Y)||^2))$$

and expand both sides into power series. We obtain:

$$(I + X + \frac{1}{2}X^2 + o(||X||^2))(I + Y + \frac{1}{2}Y^2 + o(||Y||^2)) =$$

$$= I + X + Y + \frac{1}{2}[X,Y] + \frac{1}{2}(X + Y + \frac{1}{2}[X,Y])^2 + o(||(X,Y)||^2)$$

and simplifying

$$I + X + Y + \frac{1}{2}X^2 + \frac{1}{2}Y^2 + XY + o(||(X,Y)||^2) =$$

$$= I + X + Y + \frac{1}{2}[X,Y] + \frac{1}{2}(X^2 + XY + YX + Y^2) + o(||(X,Y)||^2)$$

and so we obtain

$$[X,Y] = XY - YX + o(||(X,Y)||^2),$$

yielding

$$[X,Y] = XY - YX.$$

Thus, the Lie bracket in the case of $\mathrm{Lie}(\mathrm{GL}_n(\mathbb{R})) = M_n(\mathbb{R})$ *is the commutator.*

Example 4.49. *Similarly,* $\mathrm{Lie}(\mathrm{GL}_n(\mathbb{C})) = M_n(\mathbb{C})$ *and the Lie bracket is again the commutator. Notice here, interestingly, that we consider this as an* $\mathbb{R}$*-Lie algebra, but it is in fact naturally a* $\mathbb{C}$*-Lie algebra.*

Definition 4.50. Let V be a vector space over a field k. We denote the k-Lie algebra $\mathrm{End}_k(V)$, equipped with the Lie bracket $[T,S] := T \circ S - S \circ T$, by $\mathfrak{gl}_k(V)$, or $\mathfrak{gl}(V)$ if k is understood. One also denotes $\mathfrak{gl}_n(k) := \mathfrak{gl}(k^n)$.

Remark 4.51. If H is a closed Lie subgroup of $GL_n(\mathbb{R})$ by all that we have seen we obtain that $\mathrm{Lie}(H)$ is a $\mathbb{R}$-Lie subalgebra of $M_n(\mathbb{R})$, so the Lie bracket is given by the usual commutator of matrices.

4.7 Proof of the Jacobi identity

To prove the Jacobi identity, we will use the **adjoint representation**. Namely, given $g \in G$ consider the smooth map $c_g : G \to G$ given by $c_g(h) := ghg^{-1}$. It sends 1 to 1 and hence we obtain an $\mathbb{R}$-linear map

$$\mathrm{Ad}(g) := D_1c_g : \mathrm{Lie}(G) \to \mathrm{Lie}(G).$$

We obtain thus a map $\mathrm{Ad} : G \to \mathrm{End}_{\mathbb{R}}(\mathrm{Lie}(G))$.

Lemma 4.52. *The map* Ad *is a morphism of Lie groups*

$$\mathrm{Ad} : G \to \mathrm{Aut}_{\mathbb{R}}(\mathrm{Lie}(G)).$$

Proof. It is easy to see that Ad is multiplicative, and hence in particular its image lies in $\mathrm{Aut}_{\mathbb{R}}(\mathrm{Lie}(G))$. We leave the verification that it is smooth for now. □

We can therefore think of Ad as a representation of G on $\mathrm{Lie}(G)$ (it is a representation over $\mathbb{R}$) - called the **adjoint representation**. We can now consider

$$\mathrm{ad} := D_1\mathrm{Ad} : \mathrm{Lie}(G) \to \mathrm{End}_{\mathbb{R}}(\mathrm{Lie}(G)).$$

Recall that, since Ad is a Lie gruop morphism, ad is a Lie algebra morphism, where the Lie algebra structure on the target is that of the commutator.

Claim 4.53. *We have* $\mathrm{ad}(X)(Y) = [X, Y]$ *for all* $X, Y \in \mathrm{Lie}(G)$.

Proof. We have

$$\mathrm{ad}(X) = \lim_{t\to 0} \frac{1}{t}(\mathrm{Ad}(\exp(tX)) - \mathrm{Id}_{\mathrm{Lie}(G)})$$

and so

$$\mathrm{ad}(X)(Y) = \lim_{t\to 0} \frac{1}{t}(\mathrm{Ad}(\exp(tX))(Y) - Y).$$

Let us use $\exp^{-1}$ on some small enough neighbourhood of 1 in G as we did before. We have

$$\mathrm{Ad}(g)(Y) = (D_1c_g)(Y) = D_1(\exp^{-1} \circ c_g)(Y) = \lim_{t\to 0} \frac{1}{t} \exp^{-1}(g\exp(tY)g^{-1}).$$

Now fix $X \in \mathrm{Lie}(G)$ and let us consider $g := \exp(sX)$ in the formula above. We have

$$\exp(sX)\exp(tY)\exp(-sX) = \exp(tY + st[X, Y] + o(||(s, t)||^2))$$

and so

$$\exp^{-1}(\exp(sX)\exp(tY)\exp(-sX)) = tY + st[X, Y] + h(s, t)$$

where $h(s, t) = o(||(s, t)||^2)$. Notice that $h(0, t) = 0$ and $h(s, 0) = 0$. Hence $h(s, t) = st \cdot k(s, t)$ where k is also a smooth function from a neighbourhood of $(0, 0)$ in the (s, t)-plane to $\mathrm{Lie}(G)$. Since

$$0 = \lim_{(s,t)\to(0,0)} \frac{1}{s^2 + t^2} h(s, t) = \lim_{(s,t)\to(0,0)} \frac{st}{s^2 + t^2} k(s, t)$$

we must have $k(0, 0) = 0$ (we see this by, for example, plugging in $s = t$). We now calculate

$$\mathrm{Ad}(\exp(sX))(Y) = \lim_{t\to 0} \frac{1}{t}(tY + st[X, Y] + h(s, t)) = Y + s[X, Y] + \lim_{t\to 0} \frac{1}{t} h(s, t).$$

We obtain:

$$\operatorname{ad}(X)(Y)=\lim_{s\to 0}\frac{1}{s}(\operatorname{Ad}(\exp(sX))(Y)-Y)=[X,Y]+\lim_{s\to 0}\lim_{t\to 0}\frac{1}{st}h(s,t))=[X,Y].$$

□

Claim 4.53, coupled with ad being a Lie algebra morphism immediately results in Jacobi's identity:

$$[X,[Y,Z]]=-[[Y,Z],X]=-\operatorname{ad}([Y,Z])(X)=-[\operatorname{ad}(Y),\operatorname{ad}(Z)](X)=$$
$$=-(\operatorname{ad}(Y)(\operatorname{ad}(Z)(X))-\operatorname{ad}(Z)(\operatorname{ad}(Y)(X)))=-([Y,[Z,X]]-[Z,[Y,X]])=$$
$$=[[X,Y],Z]+[Y,[X,Z]].$$

4.8 Proof of Theorem 4.32

Claim 4.54. *Let G be a Lie group. Let $X,Y\in\operatorname{Lie}(G)$. Then*

$$\exp(X+Y)=\lim_{n\to\infty}\left(\exp(\tfrac{1}{n}X)\exp(\tfrac{1}{n}Y)\right)^n.$$

Proof. We have

$$\exp(\tfrac{1}{n}X)\exp(\tfrac{1}{n}Y)=\exp(\tfrac{1}{n}X+\tfrac{1}{n}Y+\tfrac{1}{2n^2}[X,Y]+o(\tfrac{1}{n^2}))$$

as $n\to\infty$. Therefore

$$\left(\exp(\tfrac{1}{n}X)\exp(\tfrac{1}{n}Y)\right)^n=\exp(X+Y+\tfrac{1}{2n}[X,Y]+o(\tfrac{1}{n})).$$

Taking the limit as $n\to\infty$ we get what we want. □

Proof (of Theorem 4.32, part 1). First, let us notice that it is enough to see that, for some open $1\in U\subset G$, we have that $H\cap U$ is a closed submanifold in U. Then by applying diffeomorphisms of G of translating by elements in H, we obtain that H is a closed submanifold also around all of its other points.

Let us denote by $\mathfrak{h}\subset\operatorname{Lie}(G)$ the subset consisting of X for which $\exp(tX)\in H$ for all $t\in\mathbb{R}$. Clearly this subset contains 0 and is closed under multiplication by scalars in $\mathbb{R}$. From Claim 4.54 we also see that this subset is closed under addition, so it is an $\mathbb{R}$-linear subspace of $\operatorname{Lie}(G)$.

Let $V\subset\operatorname{Lie}(G)$ be an $\mathbb{R}$-linear complement to $\mathfrak{h}$ in $\operatorname{Lie}(G)$. Define a smooth map $\phi:\operatorname{Lie}(G)\to G$ by

$$\operatorname{Lie}(G)\xleftarrow{X+Y\leftarrow\!\shortmid(X,Y)}\mathfrak{h}\times V\xrightarrow{(X,Y)\mapsto\exp(X)\cdot\exp(Y)}G.$$

We claim now that for some open $0\in U\subset V$ we have $\phi(U)\cap H=\{1\}$. Let us see first that this will finish the proof. Notice that $D_0\phi=\operatorname{Id}_{\operatorname{Lie}(G)}$. Therefore there exists an open neighbourhood of 0 in $\operatorname{Lie}(G)$, which we can assume to be of the form $U' \times U$ where U' is an open neighbourhood of 0 in $\mathfrak{h}$ and U is an open neighbourhood of 0 in V (taken small enough so that $\phi(U) \cap H = \{1\}$), such that $\phi|_{U' \times U}$ is a diffeomorphism onto its open image in G. Then to show that $H \cap \phi(U' \times U)$ is a closed submanifold of $\phi(U' \times U)$ is the same as to show that $(\phi|_{U' \times U})^{-1}(H)$ is a closed submanifold of $U' \times U$. Notice that $(\phi|_{U' \times U})^{-1}(H) = U' \times \{0\}$. Clearly $U' \times \{0\}$ is a closed submanifold of $U' \times U$, as desired.

Thus, it is left to see that there exists an open $0 \in U \subset V$ such that $\phi(U) \cap H = \{1\}$. Let us denote $C := \{v \in V \mid \phi(v) \in H\}$. Then C is a closed subset in V which is also closed under multiplication by scalars in $\mathbb{Z}$. We also know that C does not contain any non-trivial $\mathbb{R}$-linear subspace of V. Then an exercise shows that 0 is a discrete point in C, as desired. □

We have the following important result, which we don't prove:

Theorem 4.55 (Sard's theorem, weak version)**.** *Let M and N be non-empty manifolds and $\phi : M \to N$ a smooth map. There exists $q \in N$ such that for every $p \in \phi^{-1}(q)$ the differential $D_p\phi : T_pM \to T_qN$ is surjective.*

Remark 4.56. Sard's theorem in fact says that the set of points $q \in N$ for which the property we stated holds is in fact "almost all" of N (its complement has measure zero).

Using Sard's theorem, we can prove:

Claim 4.57. *Let G and H be Lie groups and let $\phi : G \to H$ be a morphism of Lie groups. If ϕ is surjective then $D_p\phi$ is surjective for all $p \in G$. If ϕ is bijective then ϕ is an isomorphism of Lie groups.*

Proof. By Sard's theorem, there exists $p \in G$ such that $D_p\phi$ is surjecitve. Then for any other $p' \in G$, we can write $gp = p'$ for $g \in G$ and then, writing $\phi = m_{\phi(g)} \circ \phi \circ m_g$ we see that $D_{p'}\phi$ is surjective. If now ϕ is in fact bijective, we want to see that its inverse is smooth. This is seen using Theorem 4.21. From it, we see that the differential of ϕ must be in fact an isomorphism at each point, and then using the inverse function theorem (which itself is a consequence of Theorem 4.21) we deduce the desired. □

Proof (of Theorem 4.32, part 2). Let us consider the map $\widetilde{\phi} : G \to G \times H$ given by $g \mapsto (g, \phi(g))$. The image Γ of $\widetilde{\phi}$ is a closed subset in $G \times H$ (called the **graph** of ϕ). In fact, clearly in our case Γ is also a subgroup of $G \times H$. Therefore, by Theorem 4.32, Γ is a closed Lie subgroup of $G \times H$. Let us consider the projections $p_1 : \Gamma \to G$ and $p_2 : \Gamma \to H$, which are clearly morphisms of Lie groups . The projection p_1 is bijective. By Claim 4.57 we obtain that p_1 is an isomorphism of Lie groups. Therefore $p_2 \circ (p_1)^{-1} : G \to H$ is a morphism of Lie groups. But clearly $p_2 \circ (p_1)^{-1} = \phi$ and we are done. □

4.9 Some of Lie's theorems

Theorem 4.58 (Lie's theorems)**.** *Let G and H be Lie groups.*

1. *Suppose that G is connected. Let $\phi_1, \phi_2 : G \to H$ be two morphisms of Lie groups. Suppose that* $\mathrm{Lie}(\phi_1) = \mathrm{Lie}(\phi_2)$*. Then $\phi_1 = \phi_2$.*
2. *Suppose that G is simply connected*[16]*. Let $\alpha : \mathrm{Lie}(G) \to \mathrm{Lie}(H)$ be a morphism of $\mathbb{R}$-Lie algebras. Then there exists a morphism of Lie groups $\phi : G \to H$ such that* $\mathrm{Lie}(\phi) = \alpha$*.*

Proof.

1. Since, for $X \in \mathrm{Lie}(G)$, we have $\phi_i(\exp(X)) = \exp((\mathrm{Lie}(\phi_i))(X))$, we deduce $\phi_1(\exp(X)) = \phi_2(\exp(X))$ for all $X \in \mathrm{Lie}(G)$. Since exp is a diffeomorphism onto the open image at a neighbourhood of $0 \in \mathrm{Lie}(G)$, we obtain $\phi_1(g) = \phi_2(g)$ for $g \in U$, where $0 \in U \subset G$ is some open subset. Hence, the subset $S \subset G$ consisting of g for which $\phi_1(g) = \phi_2(g)$ is a closed and open subgroup in G. Since G is connected, we must have $S = G$, as desired.
2. Omitted.

□

Remark 4.59. It is easy to see why the conditions in the theorem are necessary. To give a contraexample to item (1) when G is not connected, consider any two different group morphisms $\phi_1, \phi_2 : G \to H$ where G and H are finite groups. We can view G and H as Lie groups, and then of course $\mathrm{Lie}(H) = 0$ and $\mathrm{Lie}(H) = 0$, so $\mathrm{Lie}(\phi_1) = \mathrm{Lie}(\phi_2)$. To give a contra-example to item (2) when G is not simply connected, we consider $G = \mathbb{C}^\times_{|-|=1}$ and $H = G$. The Lie algebra $\mathrm{Lie}(G)$ is a one-dimensional $\mathbb{R}$-vector space, and the bracket therefore (by the alternativity axiom) must be trivial: $[X, Y] = 0$ for all $X, Y \in \mathrm{Lie}(G)$. Therefore any $\mathbb{R}$-linear map $\mathrm{Lie}(G) \to \mathrm{Lie}(G)$ is a morphism of Lie algebras. But the Lie group morphisms $G \to G$ are all of the form $\phi_n : z \mapsto z^n$ for some $n \in \mathbb{Z}$, and $\mathrm{Lie}(\phi_n)$ is the $\mathbb{R}$-linear map of multiplication by n. Therefore, if we consider any $\mathbb{R}$-linear map $\alpha : \mathrm{Lie}(G) \to \mathrm{Lie}(G)$ given by multiplication by some $c \in \mathbb{R} \smallsetminus \mathbb{Z}$, it provides a contraexample.

4.10 Representations of Lie groups versus representations of Lie algebras

Definition 4.60. Let k be a field and let $\mathfrak{g}$ be a Lie algebra over k.

- Let V be a k-vector space. A **$\mathfrak{g}$-action** on V is a map $a : \mathfrak{g} \times V \to V$ which satisfies:
    1. a is k-bilinear.

[16]For us, simply connected means connected and simply connected.

2. $a([X,Y],v) = a(X,a(Y,v)) - a(Y,a(X,v))$ for all $X,Y \in \mathfrak{g}$ and $v \in V$.

As usual, we usually write Xv instead of $a(X,v)$, so the second condition is then written $[X,Y]v = XYv - YXv$.

- A **$\mathfrak{g}$-representation** (over k), or a **$\mathfrak{g}$-module**, is a k-vector space equipped with a $\mathfrak{g}$-action.
- Let V_1 and V_2 be two $\mathfrak{g}$-representations. A **morphism of $\mathfrak{g}$-representations** from V_1 to V_2 is a k-linear map $T : V_1 \to V_2$ such that $T(Xv) = XT(v)$ for all $X \in \mathfrak{g}$ and $v \in V_1$.

It is important to understand the following exercise:

Exercise 4.8. *Let k be a field and let $\mathfrak{g}$ be a Lie algebra over k. Given a k-vector space V, check that there is a bijection between the set of Lie algebra morphisms $\mathfrak{g} \to \mathfrak{gl}(V)$ and the set of $\mathfrak{g}$-actions on V, given by sending a morphism $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ to the action $a : \mathfrak{g} \times V \to V$ given by $a(X,v) := \rho(X)(v)$.*

Example 4.61. *Let $\mathfrak{g}$ be a Lie algebra over a field k. An important example of a $\mathfrak{g}$-module is the **adjoint representation**, where $\mathfrak{g}$ acts on $\mathfrak{g}$ by setting the result of X acting on Y to be $[X,Y]$.*

Since our Lie algebras are usually over $\mathbb{R}$, but our representation spaces are usually over $\mathbb{C}$, we need to discuss complexification.

Definition 4.62. Let V be an $\mathbb{R}$-vector space. A **complexification** of V is a pair (W,ι) consisting of a $\mathbb{C}$-vector space W and an $\mathbb{R}$-linear map $\iota : V \to W$ satisfying the following universal property:

- Let U be a $\mathbb{C}$-vector space and let $\iota' : V \to U$ be an $\mathbb{R}$-linear map. Then there exists a unique $\mathbb{C}$-linear map $T : W \to U$ such that $T \circ \iota = \iota'$.

Exercise-Definition 4.63. *Let V be an $\mathbb{R}$-vector space. Given two complexifiactions of V, (W,ι) and (W',ι'), there exists a unique $\mathbb{C}$-linear map $T : W \to W'$ satisfying $T \circ \iota = \iota'$ and there exists a unique $\mathbb{C}$-linear map $S : W' \to W$ satisfying $S \circ \iota' = \iota$. Show that $S \circ T = \mathrm{id}_W$ and $T \circ S = \mathrm{id}_{W'}$. Hence T and S are isomorphisms of $\mathbb{C}$-vector spaces, and W and W' are canonically isomorphic. Thus we can speak about <u>the</u> complexification of V. We denote it by $V_{\mathbb{C}}$ (and ι is usually implicit, since it is injective as we will see immediately and hence one simply identifies V with its image in $V_{\mathbb{C}}$).*

Exercise 4.9. *Let V be an $\mathbb{R}$-vector space. We can construct $V_{\mathbb{C}}$ as follows. As an abelian group, $V_{\mathbb{C}} := V \times V$. We think of $(v_1,v_2) \in V_{\mathbb{C}}$ as $v_1 + iv_2$. Then it is clear how to define multiplication by scalar from $\mathbb{C}$: Given $a,b \in \mathbb{R}$ and $(v_1,v_2) \in V_{\mathbb{C}}$, we define*

$$(a+ib)(v_1,v_2) := (av_1 - bv_2, av_2 + bv_1).$$

Show that indeed $(V_{\mathbb{C}},\iota)$, where $\iota(v) := (v,0)$, is a complexification of V.

Exercise 4.10. *Let V be an $\mathbb{R}$-vector space. Suppose that $e_1, \dots, e_n$ is an $\mathbb{R}$-basis for V. Show that $e_1, \dots, e_n$ is a $\mathbb{C}$-basis for $V_{\mathbb{C}}$ (as we said, we mean that $\iota(e_1), \dots, \iota(e_n)$ is a $\mathbb{C}$-basis for $V_{\mathbb{C}}$, but we identify V with $\iota(V)$ with ι and keep ι implicit).*

Exercise 4.11. *Let V_1 and V_2 be $\mathbb{R}$-vector spaces and let W be a $\mathbb{C}$-vector space. Show that there is a bijection between the sets of $\mathbb{C}$-bilinear maps $(V_1)_{\mathbb{C}} \times (V_2)_{\mathbb{C}} \to W$ and of $\mathbb{R}$-bilinear maps $V_1 \times V_2 \to W$, given by restriction along the natural $V_1 \times V_2 \to (V_1)_{\mathbb{C}} \times (V_2)_{\mathbb{C}}$.*

Now, we can also complexify Lie algebras. Given an $\mathbb{R}$-Lie algebra $\mathfrak{g}$, we define its complexification $\mathfrak{g}_{\mathbb{C}}$ as a $\mathbb{C}$-Lie algebra equipped with a morphism of $\mathbb{R}$-Lie algebras $\iota : \mathfrak{g} \to \mathfrak{g}_{\mathbb{C}}$ satisfying a universal property (as an exercise, fill it in). It is constructed by taking the complexification of $\mathfrak{g}$ as an $\mathbb{R}$-vector space, and the Lie bracket $\mathfrak{g}_{\mathbb{C}} \times \mathfrak{g}_{\mathbb{C}} \to \mathfrak{g}_{\mathbb{C}}$ is obtained from Exercise 4.11 applied to the $\mathbb{R}$-bilinear map $\mathfrak{g} \times \mathfrak{g} \xrightarrow{[-,-]} \mathfrak{g} \hookrightarrow \mathfrak{g}_{\mathbb{C}}$. The (quite trivial) details are left to the reader.

The following proposition is basic for us, explaining that we can study Lie algebra representations instead of Lie group representations.

Proposition 4.64. *Let G be a simply connected Lie group.*

1. *Let V be a f.d. $\mathbb{C}$-vector space. There is a natural bijection between the set of $\mathbb{C}$-linear G-actions on V and the set of $\mathrm{Lie}(G)_{\mathbb{C}}$-actions on V. More precisely, the bijection is given as follows. Given a morphism of topological groups $\rho : G \to \mathrm{GL}_{\mathbb{C}}(V)$, we recall that it is a morphism of Lie groups, we consider the associated morphism of $\mathbb{R}$-Lie algebras, $\mathrm{Lie}(\rho) : \mathrm{Lie}(G) \to \mathfrak{gl}_{\mathbb{C}}(V)$, and then consider the unique morphism of $\mathbb{C}$-Lie algebras $\mathrm{Lie}(G)_{\mathbb{C}} \to \mathfrak{gl}_{\mathbb{C}}(V)$ whose restriction along $\mathrm{Lie}(G) \to \mathrm{Lie}(G)_{\mathbb{C}}$ is $\mathrm{Lie}(\rho)$. This last morphism of $\mathbb{C}$-Lie algebras is the $\mathrm{Lie}(G)_{\mathbb{C}}$-action on V we associate to ρ.*

2. *Let V and W be f.d. G-representations, so also considered as $\mathrm{Lie}(G)_{\mathbb{C}}$-representations by part 1 of this proposition. A $\mathbb{C}$-linear map $T : V \to W$ is a morphism of G-representations if and only if it is a morphism of $\mathrm{Lie}(G)_{\mathbb{C}}$-representations.*

Proof.

1. The described procedure is bijective, by Lie's theorem and by the universal property of complexification.

2. Let us denote the morphisms corresponding to the G-actions on V and W by $\pi_V : G \to \mathrm{Aut}_{\mathbb{C}}(V)$ and $\pi_W : G \to \mathrm{Aut}_{\mathbb{C}}(W)$. It is immediate to see that T is a moprhism of $\mathrm{Lie}(G)_{\mathbb{C}}$-representations (over $\mathbb{C}$) if and only if T is a moprhism of $\mathrm{Lie}(G)$-representations (over $\mathbb{R}$). Suppose first that T is

a morphism of G-representations. Then we have a commutative diagram

$$\begin{array}{ccc} G & \xrightarrow{\pi_V} & \mathrm{Aut}_{\mathbb{C}}(V) \\ \Big\downarrow{\scriptstyle \pi_W} & & \Big\downarrow{\scriptstyle T\circ -} \\ \mathrm{Aut}_{\mathbb{C}}(W) & \xrightarrow{-\circ T} & \mathrm{Hom}_{\mathbb{C}}(V,W) \end{array} .$$

Taking the differential at $1 \in G$ we obtain a commutative diagram

$$\begin{array}{ccc} \mathrm{Lie}(G) & \xrightarrow{D_1(\pi_V)} & \mathrm{End}_{\mathbb{C}}(V) \\ \Big\downarrow{\scriptstyle D_1(\pi_W)} & & \Big\downarrow{\scriptstyle T\circ -} \\ \mathrm{End}_{\mathbb{C}}(W) & \xrightarrow{-\circ T} & \mathrm{Hom}_{\mathbb{C}}(V,W) \end{array}$$

which precisely means that $T \circ \mathrm{Lie}(\pi_V) = \mathrm{Lie}(\pi_W) \circ T$, i.e. T is a morphism of $\mathrm{Lie}(G)$-representations. Conversely, suppose that T is a morphism of $\mathrm{Lie}(G)$-representations. Let us denote by $S \subset G$ the subset consisting of g for which $T \circ \pi_V(g) = \pi_W(g) \circ T$. We want to see that $S = G$. Notice that S is a closed subgroup of G. Also, notice that, for $X \in \mathrm{Lie}(G)$,

$$T\circ\pi_V(\exp(X)) = T\circ\exp(\mathrm{Lie}(\pi_V)(X)) = \exp(\mathrm{Lie}(\pi_W)(X))\circ T = \pi_W(\exp(X))\circ T$$

(let us leave as an exercise the the middle equality follows from $T \circ (\mathrm{Lie}(\pi_V))(X) = (\mathrm{Lie}(\pi_W))(X) \circ T$). Hence, the image of exp is contained in S. Since exp is a diffeomorphism onto the open image in some neighbourhood of $1 \in G$, we deduce that S is open in G. Since S is open and closed in G, and non-empty, and G is connected, we deduce $S = G$, as desired.

□

4.11 The case of $\mathrm{SU}(n)$

As we have seen, $\mathrm{Lie}(\mathrm{SU}(n))$ is the $\mathbb{R}$-Lie subalgebra of $\mathfrak{gl}_n(\mathbb{C})$ consisting of matrices X for which $e^{tX} \in \mathrm{SU}(n)$ for all $t \in \mathbb{R}$. In other words, the conditions are:

- $e^{tX}\overline{(e^{tX})^{\mathrm{tr}}} = 1$.
- $\det(e^{tX}) = 1$.

As an exercise, check that $\det(e^Y) = e^{\mathrm{tr}(Y)}$. Hence the second condition is equivalent to $t \cdot \mathrm{tr}(X) \in 2\pi i\mathbb{Z}$ for all $t \in \mathbb{R}$, which is equivalent to $\mathrm{tr}(X) = 0$. To check what the first condition means, we will differentiate it with respect to t. We obtain

$$X \cdot e^{tX} \cdot \overline{(e^{tX})^{\mathrm{tr}}} + e^{tX} \cdot \overline{(X \cdot e^{tX})^{\mathrm{tr}}} = 0, \tag{4.2}$$

and substituting $t := 0$ we obtain $X + \overline{X^{\mathrm{tr}}} = 0$. Conversely, it is easy to check (as an exercise) that if $X + \overline{X^{\mathrm{tr}}} = 0$ then (4.2) holds for all $t \in \mathbb{R}$ and therefore $e^{tX}\overline{(e^{tX})^{\mathrm{tr}}} = 1$ (since both sides agree for $t := 0$ and have the same derivative for all $t \in \mathbb{R}$).

To conclude, we see that $\mathrm{Lie}(\mathrm{SU}(n))$ is the $\mathbb{R}$-Lie subalgebra of $\mathfrak{gl}_n(\mathbb{C})$ consisting of matrices X satisfying $\mathrm{tr}(X) = 0$ and $X + \overline{X^{\mathrm{tr}}} = 0$.

Let us now describe the complexification $\mathrm{Lie}(\mathrm{SU}(n))_{\mathbb{C}}$. By the universal property of complexification, the morphism of $\mathbb{R}$-Lie algebras

$$\mathrm{Lie}(\mathrm{SU}(n)) \to \mathfrak{gl}_n(\mathbb{C})$$

(which is simply the embedding) induces a morphism of $\mathbb{C}$-Lie algebras

$$\mathrm{Lie}(\mathrm{SU}(n))_{\mathbb{C}} \to \mathfrak{gl}_n(\mathbb{C}). \tag{4.3}$$

We first claim that (4.3) is injective. Indeed, for that we need to check that if $X, Y \in \mathrm{Lie}(\mathrm{SU}(n))$ and $X + iY = 0$, then $X = 0$ and $Y = 0$. But if $X + iY = 0$ then we have $\overline{(X + iY)^{\mathrm{tr}}} = 0$, and since $X, Y \in \mathrm{Lie}(\mathrm{SU}(n))$ the left-hand-side is equal to $-X - i(-Y) = -(X - iY)$, so we obtain $X - iY = 0$. The equalities $X + iY = 0$ and $X - iY = 0$ of course imply $X = 0$ and $Y = 0$. Next, we notice that the image of (4.3) lies in $\mathfrak{sl}_n(\mathbb{C})$, which is the $\mathbb{C}$-Lie subalgebra of $\mathfrak{gl}_n(\mathbb{C})$ consisting of matrices with trace 0. Finally, we leave as an exercise to check that the $\mathbb{C}$-dimension of $\mathrm{Lie}(\mathrm{SU}(n))_{\mathbb{C}}$, which is the same as the $\mathbb{R}$-dimension of $\mathrm{Lie}(\mathrm{SU}(n))$, is the same as the $\mathbb{C}$-dimension of $\mathfrak{sl}_n(\mathbb{C})$. Therefore (4.3) is an isomorphism, of $\mathbb{C}$-Lie algebras.

To conclude, we have a natural isomorphism of $\mathrm{Lie}(\mathrm{SU}(n))_{\mathbb{C}}$ with $\mathfrak{sl}_n(\mathbb{C})$. We also have:

Claim 4.65. *The topological group* $\mathrm{SU}(n)$ *is simply connected.*

Proof. Omitted for now. □

Therefore, by Proposition 4.64, we obtain:

Corollary 4.66. *On finite-dimensional complex vector spaces,* $\mathrm{SU}(n)$*-representations are "the same" as* $\mathfrak{sl}_n(\mathbb{C})$*-representations, in the sense of Proposition 4.64, once we recall that we have a canonical isomorphism of the complexification of* $\mathrm{Lie}(\mathrm{SU}(n)) \subset \mathfrak{gl}_n(\mathbb{C})$ *with* $\mathfrak{sl}_n(\mathbb{C})$*, induced by the inclusion* $\mathrm{Lie}(\mathrm{SU}(n)) \subset \mathfrak{sl}_n(\mathbb{C})$*.*

5 Representation theory of $\mathfrak{sl}_2$

Throughout, we work over $\mathbb{C}$. We set $\mathfrak{g} := \mathfrak{sl}_2 := \mathfrak{sl}_2(\mathbb{C})$. We consider the following $\mathbb{C}$-basis for $\mathfrak{g}$:

$$H := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \ E := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ F := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

We have the following relations:

$$[H, E] = 2E, \ [H, F] = -2F, \ [E, F] = H. \tag{5.1}$$

5.1 Finite-dimensional irreducible modules

Lemma 5.1. *Let V be a $\mathfrak{g}$-module.*

1. *Let $v \in V$ and suppose that $Hv = cv$ for $c \in \mathbb{C}$. Then $HEv = (c+2) \cdot Ev$ and $HFv = (c-2) \cdot Fv$. In other words, E and F shift us between eigenspaces of H.*

2. *Suppose that V is finite-dimensional and non-zero. There exists $v \in V$ which is both a non-zero eigenvector of H and satisfying $Ev = 0$.*

Proof.

1. Immediate, using the relations of (5.1).

2. There exists some non-zero eigenvector of H, say $w \in V$ satisfying $Hw = dw$ for some $d \in \mathbb{C}$. Considering $(E^n w)_{n \in \mathbb{Z}_{\geq 0}}$, the non-zero vectors in this list are linearly independent, since those are eigenvectors of H with different eigenvalues $c + 2n$. Thus, there are only finitely many non-zero vectors in this list, meaning that $E^n w = 0$ for some $n \in \mathbb{Z}_{\geq 1}$. Let n minimal with that property. Denote $v := E^{n-1}w$. Then $v \neq 0$, v is an eigenvector of H, and $Ev = 0$.

□

Lemma 5.2. *Let V be a $\mathfrak{g}$-module. Let $0 \neq v \in V$ and $c \in \mathbb{C}$ be such that $Hv = cv$ and $Ev = 0$.*

1. *We have*
$$EF^n v = n(c - (n-1))F^{n-1}v$$
for all $n \in \mathbb{Z}_{\geq 1}$.

2. *Suppose that V is finite-dimensional. Then $c \in \mathbb{Z}_{\geq 0}$, $F^n v \neq 0$ for $0 \leq n \leq c$ and $F^{c+1}v = 0$.*

3. *Suppose again that V is finite-dimensional. Let $W \subset V$ be the subspace spanned by $\{F^m v\}_{0 \leq m \leq c}$. Then W is a $\mathfrak{g}$-submodule of V which is irreducible.*

Proof.

1. Let us denote $v_n := \frac{1}{n!}F^n v$. We calculate
$$Ev_1 = EFv = FEv + Hv = cv_0,$$
$$Ev_2 = \frac{1}{2}EFv_1 = \frac{1}{2}(FEv_1 + Hv_1) = \frac{1}{2}(c + (c-2))v_1 = (c-1)v_1.$$
Guessing
$$Ev_n = (c - (n-1))v_{n-1}, \tag{5.2}$$

we inductively then verify:

$$Ev_{n+1} = \frac{1}{n+1}EFv_n = \frac{1}{n+1}(FEv_n+Hv_n) = \frac{1}{n+1}((c-(n-1))Fv_{n-1}+(c-2n)v_n) =$$

$$= \frac{1}{n+1}(n(c-(n-1)) + (c-2n))v_n = (c-n)v_n.$$

This is equivalent to $EF^n v = n(c-(n-1))F^{n-1}v$ for all $n \in \mathbb{Z}_{\geq 1}$.

2. From (5.2) we find

$$\frac{1}{n!}E^n F^n v = \left(\prod_{0 \leq i \leq n-1} (c-i) \right) \cdot v. \tag{5.3}$$

Let now $n_0 \in \mathbb{Z}_{\geq 1}$ be such that $F^{n_0}v = 0$ (such n exists because V is finite-dimensional and the non-zero elements in $\{F^n v\}_{n \in \mathbb{Z}_{\geq 0}}$ are linearly independent, as eigenvectors of H with different eigenvalues). Then from (5.3) we obtain $c = i$ for some $0 \leq i < n_0 - 1$, so $c \in \mathbb{Z}_{\geq 0}$ indeed. Taking $0 \leq n \leq c$, (5.3) gives $F^n v \neq 0$. It remains to understand why $F^{c+1}v = 0$. Supposing the opposite, we would obtain inductively from (5.2) that $F^{c+n}v \neq 0$ for all $n \in \mathbb{Z}_{\geq 1}$, which would contradict V being finite-dimensional.

3. Clearly W is a $\mathfrak{g}$-submodule of V, by formulas we just seen. To see that W is an irreducible $\mathfrak{g}$-module, we consider a non-zero $\mathfrak{g}$-submodule $U \subset W$. Since the action of H on W is diagnolizable, so is the action of H on U. Therefore U contains $F^m v$ for some $0 \leq m \leq c$. Then U also contains $F^\ell F^m v$ for any $\ell \in \mathbb{Z}_{\geq 0}$, and also U contains $E^\ell F^m v$ for any $\ell \in \mathbb{Z}_{\geq 0}$, and from formulas we have seen this clearly shows that U contains $F^{m'}v$ for all $0 \leq m' \leq c$, so $U = W$.

□

Corollary 5.3. *Let V be an irreducible f.d. $\mathfrak{g}$-module. There exist $v \in V$ and $m \in \mathbb{Z}_{\geq 0}$ such that $v, Fv, F^2v, \ldots, F^m v$ is a basis for V, we have*

$$HF^n v = (m-2n)F^n v \quad \forall 0 \leq n \leq m$$

and

$$EF^n v = n(m-(n-1))F^{n-1}v \quad \forall 0 \leq n \leq m.$$

Proof. This follows from the above lemmas. □

Corollary 5.4. *For every $m \in \mathbb{Z}_{\geq 0}$ there is precisely one, up to isomorphism, irreducible $\mathfrak{g}$-module of dimension $m+1$, and we wrote above explicitly its "multiplication table".*

Proof. In view of the above, the uniqueness is clear, but still need to see existence. One possible approach is simply to define a module dictated by Corollary 5.3 and check that it is well-defined. Another approach is to consider the representation, say of SU(2), or of $\mathrm{SL}_2(\mathbb{R})$, or of $\mathrm{SL}_2(\mathbb{C})$ on the space of homogeneous polynomials of degree m in two variables, as we did before, and differentiate it to obtain a module as desired. We basically already did it before, but the "problem" with that is that it does not work if we replace $\mathbb{C}$ with an arbitrary algebraically closed field of characteristic 0. But, in fact, if we do the version with $\mathrm{SL}_2(\mathbb{C})$, we can make sense of it over any such field (but we don't do it here). $\square$

Recall that we saw that f.d. $\mathfrak{sl}_2$-modules are "the same" as f.d. SU(2)-representations. Since every f.d. SU(2)-representation, as a f.d. representation of a compact group, is completely reducible, i.e. can be written as the direct sum of irreducible subrepresentations, we deduce that every f.d. $\mathfrak{sl}_2$-module is completely reducible, i.e. can be written as the direct sum of irreducible submodules. This is a "transcendental" approach, involving analysis (another way to say it is that this approach is not algebraic, in the sense that it does not carry over to working over an arbitrary algebraically closed field of characteristic 0 instead of $\mathbb{C}$). It is sometimes called "Weyl's unitary[17] trick". We next want to give a purely algebraic approach to this complete reducibility.

5.2 Detour 1 - tensor products

We fix a field k.

Definition 5.5. Let V and W be k-vector spaces. The **tensor product** of V and W (over k) is a pair (U, B) consisting of a k-vector space U and a k-bilinear map $B : V \times W \to U$, satisfying the following **universal property**: Given a pair (U', B') consisting of a k-vector space U' and a k-bilinear map $B' : V \times W \to U'$, there exists a unique k-linear map $T : U \to U'$ such that $T \circ B = B'$.

Exercise 5.1. *Let V and W be k-vector spaces. Let (U_1, B_1) and (U_2, B_2) be two tensor products of V and W. By the definition, there exists a unique k-linear map $T_{12} : U_1 \to U_2$ such that $T_{12} \circ B_1 = B_2$ and there exists a unique k-linear map $T_{21} : U_2 \to U_1$ such that $T_{21} \circ B_2 = B_1$. Show that $T_{21} \circ T_{12} = \mathrm{Id}_{U_1}$ and $T_{12} \circ T_{21} = \mathrm{Id}_{U_2}$, so that we have a canonical isomorphism of k-vector spaces between U_1 and U_2.*

In view of the exercise, we can talk about **the** tensor product of V and W - if a tensor product exists. The notation for the vector space is $V \otimes W$ (or $V \underset{k}{\otimes} W$ in a more complete notation) and for the bilinear form it is $(v, w) \mapsto v \otimes w$ (the bilinear form itself is not given a name usually).

[17] Serre writes that Weyl used the "more theological" word "unitarian".

Exercise 5.2. *Let V and W be k-vector spaces. Show that the tensor product of V and W exists as follows. For each element $(v, w) \in V \times W$ create a formal symbol $\delta_{(v,w)}$ and create formally the k-vector space $\widetilde{U}$ with basis the elements $\{\delta_{(v,w)}\}_{(v,w)\in V\times W}$. Let U denote the quotient of $\widetilde{U}$ by the k-linear subspace generated by elements of the form*

$$\delta_{(v_1+v_2,w)}-\delta_{(v_1,w)}-\delta_{(v_2,w)},\ \delta_{(cv,w)}-c\delta_{(v,w)},\ \delta_{(v,w_1+w_2)}-\delta_{(v,w_1)}-\delta_{(v,w_2)},\ \delta_{(v,cw)}-c\delta_{(v,w)}.$$

Consider the map $B : V \times W \to U$ given by sending (v, w) to the image of $\delta_{(v,w)} \in \widetilde{U}$ under the quotient map $\widetilde{U} \to U$. Show that (U, B) is a tensor product of V and W.

Exercise 5.3. *Let V and W be k-vector spaces. Show that the tensor product of V and W exists as follows. Choose a k-basis $\{e_i\}_{i\in I}$ of V and a k-basis $\{f_j\}_{j\in J}$ of W. For each element $(i, j) \in I \times J$ create a formal symbol $\delta_{(i,j)}$ and create formally the k-vector space U with basis the elements $\{\delta_{(i,j)}\}_{(i,j)\in I\times J}$. Consider the map $B : V \times W \to U$ characterized by sending (e_i, f_j) to $\delta_{(i,j)}$. Show that (U, B) is a tensor product of V and W.*

Exercise 5.4. *Deduce from the previous exercise that given k-vector spaces V and W and k-bases $\{e_i\}_{i\in I}$ of V and $\{f_j\}_{j\in J}$ of W, we have a k-basis of $V \otimes W$ given by $\{e_i \otimes f_j\}_{(i,j)\in I\times J}$.*

5.3 Detour 2 - the Casimir element

Given a group G and an abstract G-representation V, the correct structure of an abstract G-representation on V^* is given by $(g\zeta)(v) := \zeta(g^{-1}v)$. Given two abstract G-representations V and W, the correct structure of G-representation on $\mathrm{Hom}_{\mathbb{C}}(V, W)$ is given by $(gT)(v) := gT(g^{-1}v)$. The correct structure of G-representation on $V \otimes W$ is characterized by $g(v \otimes w) = gv \otimes gw$.

Given a Lie algebra $\mathfrak{g}$ and a $\mathfrak{g}$-module V, the correct structure of a $\mathfrak{g}$-module on V^* is given by $(X\zeta)(v) := -\zeta(Xv)$. Given two $\mathfrak{g}$-modules V and W, the correct structure of $\mathfrak{g}$-module on $\mathrm{Hom}_{\mathbb{C}}(V, W)$ is given by $(XT)(v) := XT(v) - T(Xv)$. The correct structure of $\mathfrak{g}$-module on $V \otimes W$ is characterized by $X(v \otimes w) = Xv \otimes w + v \otimes Xw$.

Suppose now that $\mathfrak{g}$ is a finite-dimensional Lie algebra and that we are given a $\mathfrak{g}$-invariant non-degenerate symmetric bilinear form $B : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$. The condition of $\mathfrak{g}$-invariancy conforms with the above constructions, concretely it means that

$$B([X, Y], Z) + B(Y, [X, Z]) = 0$$

for all $X, Y, Z \in \mathfrak{g}$. The form B induces an isomorphism of $\mathfrak{g}$-modules $\iota_B : \mathfrak{g} \to \mathfrak{g}^*$ given by $\iota_B(X)(Y) := B(X, Y)$. We can construct the following isomorphisms

$$\mathrm{End}_{\mathbb{C}}(\mathfrak{g}) \xleftarrow{\sim} \mathfrak{g} \otimes \mathfrak{g}^* \xleftarrow{\sim} \mathfrak{g} \otimes \mathfrak{g}.$$

Here the left isomorphism is characterized by sending $X \otimes \zeta$ to the endomorphism sending Y to $\zeta(Y)X$. The right isomorphism is characterized by sending $X \otimes Y$ to $X \otimes \iota_B(Y)$. Let us take the image on the right of the element $\mathrm{Id}_{\mathfrak{g}}$ on the left - call it $\mathcal{C}$. Since $\mathrm{Id}_{\mathfrak{g}}$ is $\mathfrak{g}$-invariant, so is $\mathcal{C}$. If we want a concrete description, let $X_1, \ldots, X_n$ be a basis for $\mathfrak{g}$ and let $X_1^*, \ldots, X_n^*$ the basis for $\mathfrak{g}$ which is dual to this basis with respect to B, i.e. we have $B(X_i, X_j^*) = \delta_{i,j}$ for all $1 \leq i, j \leq n$. Then we see that $\mathcal{C} = \sum_{1 \leq i \leq n} X_i \otimes X_i^*$.

Now, suppose that we are given a $\mathfrak{g}$-module V. Denoting the corresponding morphism of Lie algebras $\pi : \mathfrak{g} \to \mathrm{End}_{\mathbb{C}}(V)$, we have a morphism of $\mathfrak{g}$-modules $\mathfrak{g} \otimes \mathfrak{g} \to \mathrm{End}_{\mathbb{C}}(V)$ characterized by $X \otimes Y \mapsto \pi(X) \circ \pi(Y)$. Calling the image of $\mathcal{C}$ under this map C, we obtain an endomorphism of $\mathfrak{g}$-modules $C \in \mathrm{End}_{\mathfrak{g}}(V)$. It is called the **Casimir operator** (corresponding to B).

A standard way to produce a $\mathfrak{g}$-invariant symmetric bilinear form is given by the **Killing form**. It is the form $B : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ given by $B(X, Y) := \mathrm{Tr}(\mathrm{ad}(X) \circ \mathrm{ad}(Y))$. It is indeed $\mathfrak{g}$-invariant:

$$B([Z, X], Y) + B(X, [Z, Y]) = \mathrm{Tr}(\mathrm{ad}([Z, X]) \circ \mathrm{ad}(Y)) + \mathrm{Tr}(\mathrm{ad}(X) \circ \mathrm{ad}([Z, Y])) =$$

$$= \mathrm{Tr}([\mathrm{ad}(Z), \mathrm{ad}(X)]\mathrm{ad}(Y) + \mathrm{ad}(X)[\mathrm{ad}(Z), \mathrm{ad}(Y)]) =$$

$$= \mathrm{Tr}(\mathrm{ad}(Z)\mathrm{ad}(X)\mathrm{ad}(Y) - \mathrm{ad}(X)\mathrm{ad}(Y)\mathrm{ad}(Z)) + \mathrm{Tr}(-\mathrm{ad}(X)\mathrm{ad}(Z)\mathrm{ad}(Y) + \mathrm{ad}(X)\mathrm{ad}(Z)\mathrm{ad}(Y)) = 0.$$

One of the characterizations of $\mathfrak{g}$ being **semisimple** is that the Killing form B is non-degenerate.

Let us now realize this for $\mathfrak{g} := \mathfrak{sl}_2$. First, we see that the Killing form B is non-degenerate. We work with the basis H, E, F of $\mathfrak{g}$, and compute the matrices representing in this basis:

$$\mathrm{ad}(H) : \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \ \mathrm{ad}(E) : \begin{pmatrix} 0 & 0 & 1 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ \mathrm{ad}(F) : \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix}$$

and then we compute

$$B : \begin{pmatrix} 8 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \end{pmatrix}.$$

Thus indeed B is non-degenerate. Next, we compute the basis dual to (H, E, F) with respect to B:

$$(H^*, E^*, F^*) = (\frac{1}{8}H, \frac{1}{4}F, \frac{1}{4}E).$$

Hence

$$\mathcal{C} = H \otimes H^* + E \otimes E^* + F \otimes F^* = \frac{1}{8}H \otimes H + \frac{1}{4}E \otimes F + \frac{1}{4}F \otimes E.$$

Therefore given a $\mathfrak{g}$-module V, the Casimir operator $C \in \mathrm{End}_{\mathfrak{g}}(V)$ is given by

$$C = \frac{1}{8}(H^2 + 2EF + 2FE) = \frac{1}{8}(H^2 + 2H + 4FE).$$

It is interesting, I think, to notice how C can't be seen "inside" SU(2), or even inside $\mathfrak{sl}_2$, one has to extend the scope of possible operators in order to discover it.

5.4 Complete reducibility of finite-dimensional modules

Definition 5.6. Given a vector space V, an operator $T : V \to V$ and $c \in \mathbb{C}$, we will denote by $V_{T,c} \subset V$ the eigenspace of T with eigenvalue c and by $V_{(T,c)} \subset V$ the generalized eigenspace of T with eigenvalue c.

Lemma 5.7. *Let V be a f.d. $\mathfrak{g}$-module.*

1. *If V is irreducible, of dimension $m+1$, then C acts on V by the scalar $\frac{1}{8}(m^2+m)$.*
2. *The generalized eigenvalues of the action of C on V lie in $\{\frac{1}{8}(m^2+m)\}_{m\in\mathbb{Z}_{\geq 0}}$, and if some $\frac{1}{8}(m^2+m)$ is indeed a generalized eigenvalue then $V_{(H,m)} \neq 0$.*
3. *If the only generalized eigenvalue of C acting on V is $\frac{1}{8}(m^2+m)$, for some $m \in \mathbb{Z}_{\geq 0}$, then the generalized eigenvalues of H acting on V lie in $\{m, m-2, \dots, -m+2, -m\}$.*

Proof. Let us first show that (2) and (3) follow from (1). Using dimension reasoning, we can always find a chain of $\mathfrak{g}$-submodules

$$0 = K_0 \subset K_1 \subset \ldots \subset K_r = V$$

such that K_{i+1}/K_i is an irreducible $\mathfrak{g}$-module. This clearly shows that (2) follows from (1). In the case (3), all these irreducible $\mathfrak{g}$-modules must be $(m+1)$-dimensional, by (2), and since the generalized eigenvalues of H acting on an $(m+1)$-dimensional irreducible $\mathfrak{g}$-module lie in $\{m, m-2, \dots, -m+2, -m\}$, we also get (3).

Let thus V be irreducible, of dimension $m+1$, and let $0 \neq v \in V$ be such that $Ev = 0$ and $Hv = mv$. We calculate:

$$Cv = \frac{1}{8}(H^2v + Hv + 4FEv) = \frac{1}{8}(m^2+m)v.$$

Then $CF^nv = F^nCv = \frac{1}{8}(m^2+m)F^nv$ for all $n \in \mathbb{Z}_{\geq 0}$ and so, since $\{F^nv\}_{n\in\mathbb{Z}_{\geq 0}}$ spans V, we indeed obtain that C acts on V by the scalar $\frac{1}{8}(m^2+m)$. □

Lemma 5.8. *Let V be a f.d. $\mathfrak{g}$-module. Let $n \in \mathbb{Z}_{\geq 0}$. The operator*

$$F^n : V_{(H,n)} \to V_{(H,-n)}$$

is injective.

Proof. Let us argue by induction on the dimension of V. If $V = 0$ then the claim is clear. so assume $V \neq 0$. By dimension reasoning there exists a maximal proper $\mathfrak{g}$-submodule $K \subset V$; then V/K is necessarily an irreducible $\mathfrak{g}$-module. Denote by $[-] : V \to V/K$ the quotient map. Let $0 \neq v \in V_{(H,n)}$. If $v \in K$ then by the induction hypothesis applied to K we have $F^n v \neq 0$ and we are done. Otherwise, we have $[v] \neq 0$. If $K \neq 0$ then we can apply the induction hypothesis to V/K and obtain $[F^n v] = F^n[v] \neq 0$ and so $F^n v \neq 0$. So it remains to consider the case when $K = 0$, i.e. V is irreducible. But then the claim follows from direct observation, using Corollary 5.3. □

Lemma 5.9. *Let V be a f.d. $\mathfrak{g}$-module. Suppose that the only generalized eigenvalue of C acting on V is $\frac{1}{8}(m^2+m)$, for some $m \in \mathbb{Z}_{\geq 0}$. Then $V_{(H,m)} = V_{H,m}$.*

Proof. As operators on V, we have

$$[H, F^n] = [H,F]F^{n-1} + F[H,F]F^{n-2} + \ldots + F^{n-1}[H,F] = -2n \cdot F^n.$$

Then

$$\begin{gathered}[E, F^n] = [E,F]F^{n-1} + F[E,F]F^{n-2} + \ldots + F^{n-1}[E,F] = \\ = HF^{n-1} + FHF^{n-2} + \ldots + F^{n-1}H = \\ = \big([H, F^{n-1}] + F^{n-1}H\big) + \big(F[H,F^{n-2}] + F^{n-1}H\big) + \ldots + \big(F^{n-1}H\big) = \\ = -(2(n-1) + 2(n-2) + \ldots + 2 + 0)F^{n-1} + nF^{n-1}H = \\ = nF^{n-1}(H - (n-1)).\end{gathered}$$

Since F^{m+1} and E act by zero on $V_{(H,m)}$ (by Lemma 5.7(3), because F^{m+1} and E send $V_{(H,m)}$ to $V_{(H,-m-2)}$ and to $V_{(H,m+2)}$ respectively), we obtain that $F^m(H-m)$ acts by zero on $V_{(H,m)}$. Since, by Lemma 5.8, F^m acts injectively on $V_{(H,m)}$, we obtain that $H - m$ acts by zero on $V_{(H,m)}$, implying $V_{(H,m)} = V_{H,m}$. □

Lemma 5.10. *Let V be a f.d. $\mathfrak{g}$-module. Suppose that the only generalized eigenvalue of C acting on V is $\frac{1}{8}(m^2+m)$, for some $m \in \mathbb{Z}_{\geq 0}$. Then V is a direct sum of irreducible $\mathfrak{g}$-submodules of dimension $m+1$.*

Proof. We first claim that $V_{H,m}$ generates V as a $\mathfrak{g}$-module. Let $W \subset V$ be the $\mathfrak{g}$-submodule generated by $V_{H,m}$. Since the projection map $V_{(H,m)} \to (V/W)_{(H,m)}$ is surjective and, by Lemma 5.9, we have $V_{(H,m)} = V_{H,m} \subset W$, we deduce $(V/W)_{(H,m)} = 0$. Since the only generalized eigenvalue of C acting on V/W is $\frac{1}{8}(m^2+m)$, by Lemma 5.7(2) we see that we must have $V/W = 0$, i.e. $W = V$ as desired.

Let now $e_1, \ldots, e_r$ be a basis for $V_{H,m}$. Denote by L_i the span of $e_i, Fe_i, \ldots, F^m e_i$. Notice that $Ee_i = 0$. Therefore, by Lemma 5.2, we know that L_i is an irreducible $\mathfrak{g}$-submodule of V of dimension $m+1$. Since $V_{H,m}$ generates V as a $\mathfrak{g}$-module, we have $V = \sum_i L_i$. It is left to see that $\{L_i\}_i$ is a linearly independent family. Indeed, $\{L_i\}_i$ is linearly independent if $\{L_i \cap V_{(H,c)}\}_i$ is linearly independent for every $c \in \mathbb{C}$. For $c \notin \{m, m-2, \ldots, -m+2, m\}$ this is clear, while otherwise, denoting $n := (m-c)/2$, this is equivalent to the linear independence of $\{F^n e_i\}_i$. The latter linear independence would follow from the linear independence of $\{E^n F^n e_i\}_i$. But we saw that $E^n F^n e_i = \left(n! \prod_{0 \le i \le n-1} (m-i)\right) e_i$, so that the linear independence of $\{E^n F^n e_i\}_i$ is equivalent to the linear independence of $\{e_i\}_i$, and we are done. □

Corollary 5.11 (Complete reducibility)**.** *Every f.d. $\mathfrak{g}$-module can be written as a direct sum of irreducible $\mathfrak{g}$-submodules.*

Proof. Let V be a f.d. $\mathfrak{g}$-module. By Lemma 5.7, the generalized eigenvalues of C acting on V lie in $\{\frac{1}{8}(m^2+m)\}_{m \in \mathbb{Z}_{\ge 0}}$. We can decompose $V = \bigoplus_{m \in \mathbb{Z}_{\ge 0}} V_{(C, \frac{1}{8}(m^2+m))}$, and this is a decomposition into $\mathfrak{g}$-submodules. By Lemma 5.10, each $V_{(C, \frac{1}{8}(m^2+m))}$ can be written as a direct sum of irreducible $\mathfrak{g}$-submodules of dimension $m+1$. □

From the discussion we have:

Corollary 5.12. *The action of H on a f.d. $\mathfrak{g}$-module is diagnolizable. The eigenvalues of H acting on a f.d. $\mathfrak{g}$-module are in $\mathbb{Z}$.*

Claim 5.13. *Let V be a f.d. $\mathfrak{g}$-module. Let $n \in \mathbb{Z}_{\ge 0}$. The linear maps*

$$F^n : V_{H,n} \to V_{H,-n}$$

and

$$E^n : V_{H,-n} \to V_{H,n}$$

are isomorphisms.

Proof. By Corollary 5.11 we reduce to the case when V is an irreducible $\mathfrak{g}$-module. Then the claim follows by direct observation, using Corollary 5.3. □

6 The universal enveloping algebra

In the previous section we have used $C = \frac{1}{8}(H^2 + 2H + 4FE)$, as well as calculations involving terms like $[E, F^{n+1}] = EF^{n+1} - F^{n+1}E$ and so on. We had a formal meaning for those only after having a $\mathfrak{g}$-module at hand, computing then with operators on that module. The universal enevloping algebra is a "home" for expressions like $\frac{1}{8}(H^2 + 2H + 4FE)$ which is "abstract", in the sense that, in order to have meaning for the expression, we do not require a "realization" on a module.

6.1 Algebras and modules

Definition 6.1.

- A (associative and unital) k**-algebra** is a k-vector space A equipped with a k-bilinear map $A \times A \to A$ (which we simply denote $(a, b) \mapsto ab$) satisfying:

- $(a_1a_2)a_3 = a_1(a_2a_3)$ for all $a_1, a_2, a_3 \in A$.
- There exists an element $1 \in A$ such that $1a = a$ and $a1 = a$ for all $a \in A$.

- Let A and B be k-algebras. A **morphism of k-algebras** from A to B is a k-linear map $\phi : A \to B$ satisfying $\phi(a_1a_2) = \phi(a_1)\phi(a_2)$ for all $a_1, a_2 \in A$ and $\phi(1) = 1$.

Remark 6.2. In other words, a k-algebra is a ring which also has the structure of a k-vector space and for which the multiplication map is k-bilinear.

Definition 6.3. Let A be a k-algebra.

- An **A-module** is a k-vector space M equipped with a k-bilinear map $A \times M \to M$ (which we simply denote $(a, m) \mapsto am$) satisfying:
  - $(ab)m = a(bm)$ for all $a, b \in A$ and $m \in M$.
  - $1m = m$ for all $m \in M$.
- Let M and N be A-modules. A **morphism of A-modules** from M to N is a k-linear map $\phi : M \to N$ satisfying $\phi(am) = a\phi(m)$ for all $a \in A$ and $m \in M$.

6.2 The universal eneveloping algebra

Exercise 6.1. *Let A be a k-algebra. Show that $[-,-] : A \times A \to A$ given by $[a, b] := ab - ba$ is a Lie bracket. Thus A equipped with $[a, b] := ab - ba$ is a k-Lie algebra, which we denote A^{Lie}.*

The idea of the universal enveloping algebra is that given a k-Lie algebra $\mathfrak{g}$, we want to find a morphism of k-Lie algebras $\iota : \mathfrak{g} \to A^{\mathrm{Lie}}$ "as efficient as possible". This means that we want to "artificially" manufacture a place larger than $\mathfrak{g}$, where we can form expressions such as $XY + 2XY^2 - YXY + X^3YZ$ for $X, Y, Z \in \mathfrak{g}$ (where here the product is associative), with the rule that $XY - YX$ is equal to $[X, Y]$. Here, "efficient" has a "surjective" and an "injective" meaning. As for the first, if we have some $\mathfrak{g} \to A^{\mathrm{Lie}}$ we can always embed $A \hookrightarrow B$ into some bigger k-algebra and consider the composition $\mathfrak{g} \to B^{\mathrm{Lie}}$, and this is wasteful, so we want A to be "as small as possible". On the other extreme, we can take $0 : \mathfrak{g} \to A^{\mathrm{Lie}}$ and this "loses information".

An important idea is that the universal enveloping algebra is given by a universal property:

Definition 6.4. Let $\mathfrak{g}$ be a k-Lie algebra. The **universal eneveloping algebra** of $\mathfrak{g}$ is a pair (A, ι) consisting of a k-algebra A and a morphism of k-Lie algebras $\iota : \mathfrak{g} \to A^{\mathrm{Lie}}$, satisfying the following property (called a **universal property**):

- Let (B, ϵ) be a pair consisting of a k-algebra B and a morphism of k-Lie algebras $\epsilon : \mathfrak{g} \to B^{\text{Lie}}$. Then there exists a unique morphism of k-algebras $\phi : A \to B$ such that $\epsilon = \phi \circ \iota$.

Remark 6.5. To reformulate the above definition, The pair (A, ι) is universal if for every pair (B, ϵ) the map

$$\text{Hom}_{k\text{-algebras}}(A, B) \to \text{Hom}_{k\text{-Lie algebras}}(\mathfrak{g}, B^{\text{Lie}})$$

is a bijection. One usually says in words "to give a Lie algebra morphism from $\mathfrak{g}$ is the same as to give an algebra morphism from A".

So (ι, A) of the definition above is a "universal solution" to the problem of finding a k-algebra with a morphism of k-Lie algebras from $\mathfrak{g}$, in the sense that all other solutions factor uniquely via it. The next important part of this pattern is to explain in which sense it is unique:

Lemma 6.6. *Let $\mathfrak{g}$ be a k-Lie algebra. Let (A_1, ι_1) and (A_2, ι_2) be two universal enveloping algebras of $\mathfrak{g}$. Then there exists a unique isomorphism of k-algebras $\epsilon_{12} : A_1 \to A_2$ satisfying $\epsilon_{12} \circ \iota_1 = \iota_2$.*

Proof. There exists a unique morphism of k-algebras $\epsilon_{12} : A_1 \to A_2$ satisfying $\epsilon_{12} \circ \iota_1 = \iota_2$ by the universal property of (A_1, ι_1). On the other hand, there exists a unique morphism of k-algebras $\epsilon_{21} : A_2 \to A_1$ satisfying $\epsilon_{21} \circ \iota_2 = \iota_1$ by the universal property of (A_2, ι_2). Notice that $\epsilon_{12} \circ \epsilon_{21} \circ \iota_2 = \iota_2$ but also $\text{id}_{A_2} \circ \iota_2 = \iota_2$ and therefore by the uniqueness part of the universal property of (A_2, ι_2) we must have $\epsilon_{12} \circ \epsilon_{21} = \text{id}_{A_2}$. Symmetrically, we find that $\epsilon_{21} \circ \epsilon_{12} = \text{id}_{A_1}$. □

Therefore, it is justifiable to speak about <u>the</u> universal enveloping algebra of $\mathfrak{g}$, if it exists. We denote it by $(\mathcal{U}(\mathfrak{g}), \iota)$ (but one usually keeps ι implicit, especially after seeing that it is injective). We have:

Proposition 6.7. *Let $\mathfrak{g}$ be a k-Lie algebra. Then the universal enveloping algebra of $\mathfrak{g}$ exists.*

Proof. Not difficult, but omitted (one takes a quotient of the tensor algebra of $\mathfrak{g}$, killing expressions $XY - YX - [X, Y]$). □

Lemma 6.8. *Let $\mathfrak{g}$ be a k-Lie algebra. The k-span of elements in $\mathcal{U}(\mathfrak{g})$ of the form*

$$\iota(X_1) \cdot \ldots \cdot \iota(X_m)$$

for various sequences $X_1, \ldots, X_m \in \mathfrak{g}$ is the whole $\mathcal{U}(\mathfrak{g})$.

Proof. Let us denote by $\mathcal{U}(\mathfrak{g})'$ the k-span as in the formulation of the lemma. It is clear that $\mathcal{U}(\mathfrak{g})'$ is a k-subalgebra of $\mathcal{U}(\mathfrak{g})$ (notice that it contains 1 as the empty product of $\iota(X_i)$'s). Also, the image of ι lies in $\mathcal{U}(\mathfrak{g})'$; Let us denote by $\iota' : \mathfrak{g} \to (\mathcal{U}(\mathfrak{g})')^{\text{Lie}}$ the corestriction (it is a k-Lie algebra morphism). By the universal property of $(\mathcal{U}(\mathfrak{g}), \iota)$, there exists a k-algebra morphism $e : \mathcal{U}(\mathfrak{g}) \to \mathcal{U}(\mathfrak{g})'$ satisfying $e \circ \iota = \iota'$. Let us denote by $f : \mathcal{U}(\mathfrak{g})' \to \mathcal{U}(\mathfrak{g})$ the inclusion. Then we have $f \circ \iota' = \iota$ and therefore $f \circ e : \mathcal{U}(\mathfrak{g}) \to \mathcal{U}(\mathfrak{g})$ is a k-algebra morphism satisfying $(f \circ e) \circ \iota = \iota$. Since also $\mathrm{id}_{\mathcal{U}(\mathfrak{g})} \circ \iota = \iota$, by the uniqueness part of the universal property of $(\mathcal{U}(\mathfrak{g}), \iota)$ we obtain $f \circ e = \mathrm{id}_{\mathcal{U}(\mathfrak{g})}$. This implies that f is surjective, meaning $\mathcal{U}(\mathfrak{g}) = \mathcal{U}(\mathfrak{g})'$. □

Next, assume for simplicity that $\mathfrak{g}$ is finite-dimensional and let $Y_1, \dots, Y_n$ be a k-basis for $\mathfrak{g}$. Then clearly Lemma 6.8 shows that elements of the form

$$\iota(Y_{i_1}) \cdot \ldots \cdot \iota(Y_{i_m}), \quad 1 \leq i_1, \dots, i_m \leq n$$

k-span $\mathcal{U}(\mathfrak{g})$. However, clearly those are generally not linearly independent: Write $[Y_2, Y_1] = \sum_{1 \leq i \leq n} c_i Y_i$. Then

$$\iota(Y_2)\iota(Y_1) = \iota(Y_2)\iota(Y_1) - \iota(Y_1)\iota(Y_2) + \iota(Y_1)\iota(Y_2) = [\iota(Y_2), \iota(Y_1)] + \iota(Y_1)\iota(Y_2) =$$

$$= \iota([Y_2, Y_1]) + \iota(Y_1)\iota(Y_2) = \sum_{1 \leq i \leq n} c_i \iota(Y_i) + \iota(Y_1)\iota(Y_2).$$

More generally, $\iota(Y_i)\iota(Y_j)$ is expressible as a k-linear combination of $\iota(Y_\ell)$'s and $\iota(Y_j)\iota(Y_i)$. Similarly, one can convince oneself that elements of the form

$$\iota(Y_{i_1}) \cdot \ldots \cdot \iota(Y_{i_m}), \quad 1 \leq i_1 \leq i_2 \leq \ldots \leq i_m \leq n$$

k-span $\mathcal{U}(\mathfrak{g})$ - if we are given a product with a "non-correct" order, by operations as above, relying on the swapping possibility $YX = XY + [Y, X]$, we can eventually rewrite it in terms of products in the "correct" order. Those already are linearly independent, i.e. form a k-basis for $\mathcal{U}(\mathfrak{g})$, as given by the **PBW theorem**:

Theorem 6.9 (PBW theorem)**.** *Let $\mathfrak{g}$ be a k-Lie algebra, let us say finite-dimensional for simplicity of formulation. Let $Y_1, \dots, Y_n$ be a k-basis for $\mathfrak{g}$. Then*

$$\{\iota(Y_1)^{m_1} \iota(Y_2)^{m_2} \cdot \ldots \cdot \iota(Y_n)^{m_n}\}_{(m_1, \dots, m_n) \in (\mathbb{Z}_{\geq 0})^n}$$

is a k-basis for $\mathcal{U}(\mathfrak{g})$.

Proof. Omitted. □

Exercise 6.2. *See that the PBW theorem in particular shows that ι is injective.*

Finally, let $\mathfrak{g}$ be a k-Lie algebra and let M be a k-vector space. A $\mathfrak{g}$-action on M is encoded by a k-Lie algebra morphism $\mathfrak{g} \to (\mathrm{End}_k(M))^{\mathrm{Lie}}$. An $\mathcal{U}(\mathfrak{g})$-action on M is encoded by a k-algebra morphism $\mathcal{U}(\mathfrak{g}) \to \mathrm{End}_k(M)$. But by the universal property, those two are in bijection. In words, "to give a $\mathfrak{g}$-module is the same as to give a $\mathcal{U}(\mathfrak{g})$-module". To repeat, this can be reformulated as saying that given a $\mathfrak{g}$-module M, there exists a unique $\mathcal{U}(\mathfrak{g})$-action on M, for which $\iota(X)m = Xm$ given $X \in \mathfrak{g}, m \in M$. The identification also applies to morphisms - if M and N are two $\mathfrak{g}$-modules (and thus $\mathcal{U}(\mathfrak{g})$-modules) then a k-linear map $T : M \to N$ is a $\mathfrak{g}$-morphism if and only if it is a $\mathcal{U}(\mathfrak{g})$-morphism - this follows easily from Lemma 6.8.

6.3 The universal enveloping algebra as a "deformation"

Definition 6.10. A **grading** on a k-algebra A is a linearly independent sequence of k-linear subspaces of A

$$A_0, A_1, \ldots$$

such that:

- $1 \in A_0$.
- $A_n \cdot A_m \subset A_{n+m}$ for all $n, m \in \mathbb{Z}_{\geq 0}$.
- $\bigoplus_{n \in \mathbb{Z}_{\geq 0}} A_n = A$.

We for convenience always denote $A_n := 0$ for $n \in \mathbb{Z}_{<0}$. A k-algebra equipped with a grading is called a **graded** k-algebra. As an exercise, figure out what is a morphism of graded k-algebras.

Example 6.11. *A polynomial algebra $A := k[x_1, \ldots, x_r]$ is naturally graded, by letting A_n be the subspace of homogeneous polynomials of degree n.*

The example has the following more abstract incarnation. We will work with infinite fields for simplicity (one can formulate things so that this will not be required).

Example 6.12. *Assume that k is infinite. Let V be a finite-dimensional k-vector space. We can consider the k-algebra $k[V]$ of polynomial functions on V, graded by taking $k[V]_n$ to be the subspace of homogeneous polynomials of degree n. When we choose coordinates, it becomes isomorphic to $k[x_1, \ldots, x_r]$, where $r := \dim_k V$.*

Exercise 6.3. *Assume that k is infinite. Let $x_1, \ldots, x_r$ be a k-basis for V^*. Then a k-basis for $k[V]_n$ is given by $x_1^{m_1} \cdot \ldots \cdot x_r^{m_r}$, for $(m_1, \ldots, m_r) \in \mathbb{Z}_{\geq 0}^r$ and $m_1 + \ldots + m_r = n$.*

We have the following:

Claim 6.13. *Assume that k is infinite. Let V be a finite-dimensional k-vector space. Note that $k[V]_1$ is equal to V^*, the dual space. We have the following universal properties:*

- *Let A be a commutative k-algebra. Then*

  $$\mathrm{Hom}_{k\text{-algebras}}(k[V], A) \to \mathrm{Hom}_{k\text{-vector spaces}}(V^*, A),$$

  given by restricting to $k[V]_1 = V^*$, is a bijection.

- *Let A be a commutative graded k-algebra. Then*

  $$\mathrm{Hom}_{\text{graded } k\text{-algebras}}(k[V], A) \to \mathrm{Hom}_{k\text{-vector spaces}}(V^*, A_1),$$

  given by restricting to $k[V]_1 = V^*$, is a bijection.

Definition 6.14. A **filtration** on a k-algebra A is an increasing sequence of k-linear subspaces of A

$$A_{\leq 0} \subset A_{\leq 1} \subset \dots$$

such that:

- $1 \in A_{\leq 0}$.
- $A_{\leq n} \cdot A_{\leq m} \subset A_{\leq n+m}$ for all $n, m \in \mathbb{Z}_{\geq 0}$.
- $\cup_{n \in \mathbb{Z}_{\geq 0}} A_{\leq n} = A$.

We for convenience always denote $A_{\leq n} := 0$ for $n \in \mathbb{Z}_{<0}$. A k-algebra equipped with a filtration is called a **filtered** k-algebra. As an exercise, figure out what is a morphism of filtered k-algebras.

Example 6.15. *The universal enevloping algebra $U(\mathfrak{g})$ is naturally filtered. Namely, we define $U(\mathfrak{g})_{\leq n}$ to be the k-span of expressions $\iota(X_1) \cdot \ldots \cdot \iota(X_m)$ for $X_1, \ldots, X_m \in \mathfrak{g}$ and $m \leq n$.*

Definition 6.16. Let A be a filtered k-algebra. The **associated graded** k-algebra, denoted $\mathrm{gr}(A)$, is the graded k-algebra constructed as follows:

$$\mathrm{gr}(A) := \bigoplus_{n \in \mathbb{Z}_{\geq 0}} (A_{\leq n} / A_{\leq n-1}) \cdot t^n$$

where t^n is just a dummy which will prevent possible ambiguity. The product is given as follows. Given $a \in A_{\leq n}$ and $b \in A_{\leq m}$, we let

$$((a + A_{\leq n-1}) \cdot t^n)((b + A_{\leq m-1}) \cdot t^m) := (ab + A_{\leq n+m-1}) \cdot t^{n+m}.$$

The grading is given by $(\mathrm{gr}(A))_n := (A_{\leq n} / A_{\leq n-1}) \cdot t^n$.

We now ask what is the associated graded of $U(\mathfrak{g})$. Notice first that $\mathrm{gr}(U(\mathfrak{g}))$ is commutative. Indeed, this follows from the following property

$$[\iota(X_1) \cdot \ldots \iota(X_n), \iota(Y_1) \cdot \ldots \cdot \iota(Y_m)] \subset U(\mathfrak{g})_{\leq n+m-1}.$$

This property follows by induction from the property $[\iota(X), \iota(Y)] \subset U(\mathfrak{g})_{\leq 1}$, which is clear since $[\iota(X), \iota(Y)] = \iota([X, Y])$. Next, let us notice that $U(\mathfrak{g})_{\leq 0} = k \cdot t^0$ while $U(\mathfrak{g})_{\leq 1} = (k \oplus \iota(\mathfrak{g})) \cdot t^1$, and so $U(\mathfrak{g})_{\leq 1} / U(\mathfrak{g})_{\leq 0} \cong \mathfrak{g}$ (via $(\iota(X) + U(\mathfrak{g})_{\leq 0}) \cdot t^1 \leftarrow\!\shortmid X$). Therefore, by Claim 6.13, we obtain a morphism of graded k-algebras

$$k[\mathfrak{g}^*] \to \mathrm{gr}(U(\mathfrak{g})), \tag{6.1}$$

the unique one whose pre-composition with the natural $\mathfrak{g} \to k[\mathfrak{g}^*]$ is equal to the map $\mathfrak{g} \to \mathrm{gr}(U(\mathfrak{g}))$ given by $X \mapsto (\iota(X) + U(\mathfrak{g})_{\leq 0}) \cdot t^1$.

Exercise 6.4. *Deduce (assuming that k is infinite) from the PBW theorem that the map (6.1) is an isomorphism.*

Therefore, we can think of $U(\mathfrak{g})$ as a "non-commutative deformation" of $k[\mathfrak{g}^*]$.

7 Representation theory of $\mathfrak{sl}_n$

Throughout, we work over $\mathbb{C}$. We set $\mathfrak{g} := \mathfrak{sl}_n := \mathfrak{sl}_n(\mathbb{C})$. We denote by $\mathfrak{h} \subset \mathfrak{g}$ the Lie subalgebra of diagonal matrices. We denote by $\mathfrak{n} \subset \mathfrak{g}$ the Lie subalgebra of nilpotent upper-triangular matrices and by $\mathfrak{n}^- \subset \mathfrak{g}$ the Lie subalgebra of nilpotent lower-triangular matrices.

7.1 Weights

Our main tool in visualizing $\mathfrak{g}$-representations is by considering $\mathfrak{h}$-eigenspaces. Here the common terminology is "weights" rather than "eigenvalues" etc.

Definition 7.1. Let V be a $\mathfrak{g}$-module.

1. Let $v \in V$ and let $\lambda \in \mathfrak{h}^*$. We say that v is a **weight vector with weight** λ if $Hv = \lambda(H)v$ for all $H \in \mathfrak{h}$.

2. Let $v \in V$. We say that v is a **weight vector** if for some $\lambda \in \mathfrak{h}^*$ it is a weight vector with weight λ.

3. Let $\lambda \in \mathfrak{h}^*$. We denote

$$V_{\mathfrak{h},\lambda} := \{\text{weight vectors with weight } \lambda \text{ in } V\} \subset V.$$

   This is a linear subspace of V, called the **λ-weight space**.

4. We say that $\lambda \in \mathfrak{h}^*$ is a **weight** of V if $V_{\mathfrak{h},\lambda} \neq 0$. We denote by $\mathrm{wt}(V) \subset \mathfrak{h}^*$ the subset consisting of weights of V.

Exercise 7.1. *Let V be a $\mathfrak{g}$-module. The family of subspaces $\{V_{\mathfrak{h},\lambda}\}_{\lambda \in \mathfrak{h}^*}$ is linearly independent.*

Definition 7.2. A $\mathfrak{g}$-module is said to be a **weight module** if it is spanned by weight vectors.

Exercise 7.2. *Let V be a $\mathfrak{g}$-module. Then V is a weight module if and only if $V = \oplus_{\lambda \in \mathfrak{h}^*} V_{\mathfrak{h},\lambda}$ (and of course we can then also write $V = \oplus_{\lambda \in \mathrm{wt}(V)} V_{\mathfrak{h},\lambda}$).*

7.2 Roots

Definition 7.3. Considering $\mathfrak{g}$ as a $\mathfrak{g}$-module via the adjoint representation, the set of non-zero weights $\mathrm{wt}(\mathfrak{g})$ is called the set of **roots** of $\mathfrak{g}$. Let us denote it by R.

Let us see more precisely what are the roots of $\mathfrak{g}$. For $1 \leq i, j \leq n$ with $i \neq j$, denote by $E_{i,j} \in \mathfrak{g}$ the matrix whose (i,j)-entry is 1 and all other entries are 0. Then one calculates:

Exercise 7.3. *We have*

$$[\mathrm{diag}(h_1, \ldots, h_n), E_{i,j}] = (h_i - h_j)E_{i,j}.$$

Thus, denoting by $\alpha_{i,j} \in \mathfrak{h}^*$ the functional given by sending $\mathrm{diag}(h_1, \ldots, h_n)$ to $h_i - h_j$, we see that $\alpha_{i,j} \in R$. Notice that

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{i \neq j} \mathbb{C} \cdot E_{i,j}.$$

Since clearly $\mathfrak{h} \subset \mathfrak{g}_{\mathfrak{h},0}$, we deduce by linear independence of weight spaces that $\mathfrak{h} = \mathfrak{g}_{\mathfrak{h},0}$, $R = \{\alpha_{i,j}\}_{i \neq j}$ and $\mathfrak{g}_{\mathfrak{h},\alpha_{i,j}} = \mathbb{C} \cdot E_{i,j}$. Given $\alpha \in R$, let us also denote $E_\alpha := E_{i,j}$ for the pair (i,j) such that $\alpha = \alpha_{i,j}$.

Using root vectors and weight spaces, we can imagine "geometrically" how $\mathfrak{g}$ acts on weight modules:

Exercise 7.4. *Given a $\mathfrak{g}$-modules V and $\alpha, \lambda \in \mathfrak{h}^*$ we have*

$$\mathfrak{g}_{\mathfrak{h},\alpha} \cdot V_{\mathfrak{h},\lambda} \subset V_{\mathfrak{h},\lambda+\alpha}.$$

Given $\alpha, \beta \in \mathfrak{h}^*$ we have

$$[\mathfrak{g}_{\mathfrak{h},\alpha}, \mathfrak{g}_{\mathfrak{h},\beta}] \subset \mathfrak{g}_{\mathfrak{h},\alpha+\beta}.$$

Using the previous exercise, we can do the following exercise:

Exercise 7.5. *Let V be a $\mathfrak{g}$-module. Let $S \subset V$ be a subset consisting of weight vectors, and suppose that S generates V as a $\mathfrak{g}$-module. Then V is a weight module.*

Definition 7.4. The roots $\alpha_{i,j}$ with $i < j$ are called the **positive roots**. We denote By $R^+ \subset R$ the subset of positive roots. Similarly, the roots $\alpha_{i,j}$ with $i > j$ are called the **negative roots** and we denote by $R^- \subset R$ the subset of negative roots. We have $R^- = -R^+$. The roots $\alpha_{i,j}$ with $j = i+1$ are called the **simple roots**. We denote by $R^s \subset R^+$ the subset of simple roots.

Exercise 7.6. *We have $R = R^+ \coprod R^-$. The subset $R^s \subset \mathfrak{h}^*$ is a basis. The coefficients in an expression of an element of R^+ as a linear combination of elements in R^s are non-negative integers.*

Notice that

$$\mathfrak{n} = \bigoplus_{\alpha \in R^+} \mathfrak{g}_{\mathfrak{h},\alpha}$$

and

$$\mathfrak{n}^- = \bigoplus_{\alpha \in R^-} \mathfrak{g}_{\mathfrak{h},\alpha}.$$

Let us also write, for $\alpha \in R^+$, $F_\alpha := E_{-\alpha}$.

Definition 7.5. Let $\alpha \in R^+$ be a positive root, write $\alpha = \alpha_{i,j}$ for $1 \leq i < j \leq n$. The corresponding **co-root** $H_\alpha \in \mathfrak{h}$ is defined as having the entry 1 at the i-place, the entry -1 at the j-place and entries 0 at all other places.

Exercise 7.7. *Let $\alpha \in R^+$. Check that we have*

$$[H_\alpha, E_\alpha] = 2E_\alpha, \ [H_\alpha, F_\alpha] = -2F_\alpha, \ [E_\alpha, F_\alpha] = H_\alpha.$$

Definition 7.6. Let $\alpha \in R^+$. Let us denote by $\mathfrak{g}_\alpha \subset \mathfrak{g}$ the Lie subalgebra spanned by $H_\alpha, E_\alpha, F_\alpha$. Notice that, by Exercise 7.7, $\mathfrak{g}_\alpha$ is indeed a Lie sublagebra, and it is isomorphic to $\mathfrak{sl}_2$ by sending $H_\alpha, E_\alpha, F_\alpha$ to H, E, F respectively.

We can now show:

Proposition 7.7. *Let V be a f.d. $\mathfrak{g}$-module. Then V is a $\mathfrak{h}$-weight module.*

Proof. One possible proof is again via Weyl's "unitary trick", considering on V the $\mathrm{SU}(n)$-action corresponding to the $\mathfrak{g}$-action, and noticing that, denoting by $T \subset \mathrm{SU}(n)$ the subgroup of diagonal matrices as before, the eigenspaces of T are eigenspaces of $\mathfrak{h}$, since $\mathfrak{h}$ is the complexification of the Lie algebra of T. Again, this proof is not algebraic.

An algebraic proof is given as follows. For $\alpha \in R^s$, consider V as a $\mathfrak{g}_\alpha$-module. Then we have already seen (Corollary 5.12) that the action of H_α on V is diagnolizable. Since the H_α's, for $\alpha \in R^s$, span $\mathfrak{h}$ (and $\mathfrak{h}$ is abelian) this implies that $\mathfrak{h}$ acts on V diagonlizably as well. □

We again consider the **Weyl group** $W := S_n$ acting (linearly) on $\mathfrak{h}$ by permuting the entries:

$$\sigma(\mathrm{diag}(x_1, \ldots, x_n)) := (x_{\sigma^{-1}(1)}, \ldots, x_{\sigma^{-1}(n)}).$$

We get an induced (linear) action of W on $\mathfrak{h}^*$.

Definition 7.8. Let $\alpha \in R$. Write $\alpha = \alpha_{i,j}$ for $1 \leq i, j \leq n$ with $i \neq j$. We define $s_\alpha \in W$ to be the permutation sending i to j and j to i and fixing all other elements.

Exercise 7.8. *Recall that the group W is generated by $\{s_\alpha\}_{\alpha \in R^s}$.*

We have the following formula for the action of s_α:

Lemma 7.9. *Let $\alpha \in R$. Then*

$$s_\alpha(\lambda) = \lambda - \lambda(H_\alpha)\alpha$$

for all $\lambda \in \mathfrak{h}^*$.

Proof. Notice that $\alpha(H_\alpha) = 2$ and both sides are seen to send α to $-\alpha$. If $\lambda \in \mathfrak{h}^*$ is such that $\lambda(H_\alpha) = 0$ then it is easy to see that both sides send λ to itself. □

7.3 Highest weights

Definition 7.10. Let V be a $\mathfrak{g}$-module.

1. A vector $v \in V$ is called a **extremal** if v is a weight vector and $\mathfrak{n}v = 0$.

2. $\lambda \in \mathfrak{h}^*$ is called an **extremal weight** of V if there exists a non-zero extremal vector in V with weight λ. We denote by $\mathrm{ext}(V) \subset \mathfrak{h}^*$ the set of extremal weights of V.

3. A vector $v \in V$ is called a **highest weight vector** if v is an extremal vector and v generates V as a $\mathfrak{g}$-module.

4. V is called a **highest weight module** if it is non-zero and it has a highest weight vector.

5. If V is a highest weight module then the weight of a highest weight vector in V is called a **highest weight** of V.

Lemma 7.11. *Let V be a $\mathfrak{g}$-module. A weight vector $v \in V$ is extremal if and only if $E_\alpha v = 0$ for all $\alpha \in R^s$.*

Proof. Since $\mathfrak{n}$ is spanned by vectors E_α for $\alpha \in R^+$, it is enough to see that if $E_\alpha v = 0$ for all $\alpha \in R^s$ then also $E_\alpha v = 0$ for all $\alpha \in R^+$. A calculation gives that, given $1 \leq i < j < k \leq n$, we have

$$[E_{\alpha_{i,j}}, E_{\alpha_{j,k}}] = E_{\alpha_{i,k}}.$$

Thus we obtain

$$E_{\alpha_{i,j}} = [E_{\alpha_{i,i+1}}, [E_{\alpha_{i+1,i+2}}, [\ldots [E_{\alpha_{j-2,j-1}}, E_{\alpha_{j-1,j}}] \ldots]].$$

Notice that if $Xv = 0$ and $Yv = 0$ for some $X, Y \in \mathfrak{g}$ then also $[X,Y]v = 0$, since $[X,Y]v = XYv - YXv$, and hence the claim is clear. ☐

Definition 7.12. We define a partial order on $\mathfrak{h}^*$ as follows. We set $\lambda \leq \mu$ if $\mu - \lambda \in \sum_{\alpha \in R^+} \mathbb{Z}_{\geq 0} \cdot \alpha$.

Lemma 7.13. *Let V be a highest weight $\mathfrak{g}$-module, with highest weight $\lambda \in \mathfrak{h}^*$.*

1. *V is a weight $\mathfrak{g}$-module, and all weight spaces of V are finite-dimensional.*

2. *We have $\mathrm{wt}(V) \subset \{\lambda' \in \mathfrak{h}^* \mid \lambda' \leq \lambda\}$.*

3. *λ is the unique highest weight of V (so that we can speak of <u>the</u> highest weight of a highest weight module).*

4. *We have $\dim_{\mathbb{C}} V_{\mathfrak{h},\lambda} = 1$.*

Proof. By the PBW theorem V is spanned by vectors of the form

$$F_{\beta_m} \cdot \ldots \cdot F_{\beta_1} \cdot H_l \cdot \ldots \cdot H_1 \cdot E_{\alpha_k} \cdot \ldots \cdot E_{\alpha_1} \cdot v.$$

Those are scalar multiples of

$$F_{\beta_m} \cdot \ldots \cdot F_{\beta_1} v.$$

Those vectors have weight $\lambda - \sum_{1 \leq i \leq m} \beta_i$, which sits in $\{\mu \in \mathfrak{h}^* \mid \mu \leq \lambda\}$. Only one of them has weight λ, namely v. ☐

Claim 7.14. *Every irreducible f.d. $\mathfrak{g}$-module is a highest weight module.*

Proof. Let E be an irreducible f.d. $\mathfrak{g}$-module. The set $\mathrm{wt}(E)$ is finite and non-empty and hence contains maximal elements with respect to the partial order $\leq$. Let $\lambda \in \mathrm{wt}(E)$ be such a maximal element. Let $0 \neq v \in E_{\mathfrak{h},\lambda}$. As E is irreducible, v generates E as a $\mathfrak{g}$-module. We claim that v is extremal. Indeed, we want to see that $E_\alpha v = 0$ for $\alpha \in R^+$. But $E_\alpha v \in E_{\mathfrak{h},\lambda+\alpha}$ and by the maximality of λ we have $\lambda + \alpha \notin \mathrm{wt}(E)$, i.e. $E_{\mathfrak{h},\lambda+\alpha} = 0$ and so $E_\alpha v = 0$. □

7.4 Irreducible highest weight modules

Lemma 7.15. *The highest weights of non-isomorphic irreducible highest weight $\mathfrak{g}$-modules are non-equal.*

Proof. Let E and F be irreducible highest weight $\mathfrak{g}$-modules, both with highest weight $\lambda \in \mathfrak{h}^*$. We want to see that E and F are isomorphic.

Let $v_1 \in E$ and $v_2 \in F$ be non-zero highest weight vectors (with weight λ). Let us consider the $\mathfrak{g}$-module $E \oplus F$ and the vector $(v_1, v_2) \in E \oplus F$. Let us consider the $\mathfrak{g}$-submodule $M \subset E \oplus F$ generated by (v_1, v_2). We claim that $(0, v_2) \notin M$ (and analogously $(v_1, 0) \notin M$). Indeed, M is generated by the highest weight vector (v_1, v_2) with highest weight λ and therefore by Lemma 7.13 we have $\dim M_{\mathfrak{h},\lambda} = 1$ (so $M_{\mathfrak{h},\lambda}$ is spanned by (v_1, v_2)). Since $(0, v_2) \in E \oplus F$ is also a highest weight vector with highest weight λ, would it lie in M it would be a scalar multiple of (v_1, v_2), which is of course not correct. Thus $(0, v_2) \notin M$.

Let us now consider the projection $\phi : M \hookrightarrow E \oplus F \to E$. We have $\mathrm{Im}(\phi) \neq 0$ since $v_1 = \phi(v_1, v_2)$ and therefore, since E is irreducible, we have $\mathrm{Im}(\phi) = E$. We also have $\mathrm{Ker}(\phi) = M \cap F$ (where we identify $F \cong 0 \oplus F \subset E \oplus F$). Since $(0, v_2) \notin M$, we have $\mathrm{Ker}(\phi) \neq F$ and therefore, since F is irreducible, we have $\mathrm{Ker}(\phi) = 0$. Thus, we have obtained that ϕ is an isomorphism from M to E. Completely analogously we obtain that ϕ is an isomorphism from M to F, deducing that E and F are isomorphic. □

We therefore see that irreducible highest weight $\mathfrak{g}$-modules are classified by their highest weights. We still don't know whether, given $\lambda \in \mathfrak{h}^*$, there exists an irreducible highest weight $\mathfrak{g}$-module with highest weight λ - we will see later that the answer is indeed positive. We also don't know yet for which $\lambda \in \mathfrak{h}^*$ an irreducible highest weight $\mathfrak{g}$-module with highest weight λ is finite-dimensional - we will see later the answer to that also.

7.5 Verma modules

To exhibit the existence of highest weight modules with a given highest weight, we exhibit a **univeral** one. Let $\lambda \in \mathfrak{h}^*$. The **Verma $\mathfrak{g}$-module** M_λ is characterized by a universal property:

Definition 7.16. Let $\lambda \in \mathfrak{h}^*$. A pair (M, v) consisting of a $\mathfrak{g}$-module M and an extreme vector $v \in M_{\mathfrak{h},\lambda}$ is a **Verma module** corresponding to λ if the following universal property holds:

- For every $\mathfrak{g}$-module N the map

$$\mathrm{Hom}_{\mathfrak{g}}(M, N) \to \{w \in N_{\mathfrak{h},\lambda} \mid \mathfrak{n}w = 0\}$$

given by $\phi \mapsto \phi(v)$ is a bijection.

Remark 7.17. In words, "to give a morphism from the Verma module corresponding to λ is the same as to specify an extremal vector with weight λ".

Exercise-Definition 7.18. *Show that given two Verma modules (M, v) and (M', v') corresponding to λ, there exists a unique isomorphism of $\mathfrak{g}$-modules $\epsilon : M \to M'$ satisfying $\epsilon(v) = v'$. In this sense, we can speak of the Verma module corresponding to λ, if it exists. We always denote it by (M_λ, v_λ).*

Exercise 7.9. *Given a Verma module (M_λ, v_λ), show that v_λ generates M_λ as a $\mathfrak{g}$-module (by abstract reasoning using the universal property). However, we don't know at the moment (but will know in a moment) whether $v_\lambda \neq 0$ - this is equivalent to the existence of some module admitting a non-zero extreme vector with weight λ. Once we will know that (M_λ exists and) $v_\lambda \neq 0$, we will know that M_λ is a highest weight module with highest weight λ.*

Lemma 7.19.

1. *For every $\lambda \in \mathfrak{h}^*$ there exists the Verma module corresponding to λ - we denote it always by (M_λ, v_λ).*
2. *The map*
$$\mathcal{U}(\mathfrak{n}^-) \to M_\lambda$$
given by $d \mapsto dv_\lambda$ is an isomorphism of $\mathfrak{n}^-$-modules. Equivalently, choosing an ordering $\alpha_1, \ldots, \alpha_r$ of R^+,
$$\left\{F_{\alpha_1}^{m_1} \cdot \ldots \cdot F_{\alpha_r}^{m_r} \cdot v_\lambda\right\}_{(m_1,\ldots,m_r)\in\mathbb{Z}^r_{\geq 0}}$$
is a basis for M_λ.
3. *The vector v_λ is non-zero and it generates M_λ as a $\mathfrak{g}$-module, i.e. it is a highest weight vector. Thus M_λ is a highest weight module with highest weight λ.*

Proof. Consider the left ideal I_λ in $U(\mathfrak{g})$ generated by elements X for $X \in \mathfrak{n}$ and by elements $H - \lambda(H)$ for $H \in \mathfrak{h}$. Consider the $\mathcal{U}(\mathfrak{g})$-module $M_\lambda := \mathcal{U}(\mathfrak{g})/I_\lambda$ and the element $v_\lambda \in M_\lambda$ given as $[1]$ where $[-] : \mathcal{U}(\mathfrak{g}) \to \mathcal{U}(\mathfrak{g})/I_\lambda = M_\lambda$ is the canonical projection. We leave the reader the tautological verification that (M_λ, v_λ) is indeed a Verma module corresponding to λ.

Let us leave the following as an exercise (in invertible triangular changes of basis): Choose an ordering $\alpha_1, \dots, \alpha_r$ of R^+ and choose a basis $H_1, \dots, H_k$ for $\mathfrak{h}$. Then

$$\left\{F_{\alpha_1}^{m_1} \cdot \ldots \cdot F_{\alpha_r}^{m_r} \cdot (H_1 - \lambda(H_1))^{\ell_1} \cdot \ldots \cdot (H_k - \lambda(H_k))^{\ell_k} \cdot E_{\alpha_1}^{n_1} \cdot \ldots \cdot E_{\alpha_r}^{n_r}\right\}_{\substack{(m_1,\dots,m_r)\in\mathbb{Z}_{\geq 0}^r \\ (n_1,\dots,n_r)\in\mathbb{Z}_{\geq 0}^r \\ (\ell_1,\dots,\ell_k)\in\mathbb{Z}_{\geq 0}^k}} \tag{7.1}$$

is a basis for $U(\mathfrak{g})$ (when $\lambda = 0$ this is just given by the PBW theorem, and otherwise we need to relate to that case by an invertible triangular change of basis).

Now, in fact, I_λ is the span of basis elements in (7.1) for which $(\ell_1, \dots, \ell_k, n_1, \dots, n_r) \neq 0$ (we leave this as an exercise). We thus have $U(\mathfrak{g}) = U(\mathfrak{n}^-) \oplus I_\lambda$. The rest of the claims of the Lemma follow from this.

□

Example 7.20. *Let us consider $\mathfrak{g} = \mathfrak{sl}_2$. Let us identify $\mathbb{C}$ with $\mathfrak{h}^*$, be sending $c \in \mathbb{C}$ to the functional in $\mathfrak{h}^*$ which maps H to c. The Verma module M_c has a vector v_c, and $v_c, Fv_c, F^2v_c, \dots$ forms a basis for M_c. We have $HF^nv_c = (c-2n)F^nv_c$. We have $Ev_c = 0$, and using Lemma 5.2 we obtain $EF^nv_c = n(c-(n-1))F^{n-1}v_c$ for all $n \in \mathbb{Z}_{\geq 1}$.*

Exercise 7.10. *Let us continue with Example 7.20. Show that if $c \notin \mathbb{Z}_{\geq 0}$ then M_c is an irreducible $\mathfrak{g}$-module. If $c \in \mathbb{Z}_{\geq 0}$, notice that, defining $N_c \subset M_c$ as the span of the vectors F^nv_c for $n \in \mathbb{Z}_{\geq c+1}$, N_c is a $\mathfrak{g}$-submodule of M_c, $(N_c, F^{c+1}v_c)$ is a Verma module with highest weight $-c-2$, and M_c/N_c is an irreducible $\mathfrak{g}$-module of dimension $c+1$.*

7.6 Irreducible highest weight modules - existence

Claim 7.21. *Let M be a highest weight $\mathfrak{g}$-module. Then M admits a unique irreducible quotient $\mathfrak{g}$-module. In other words, M admits a unique maximal proper $\mathfrak{g}$-submodule. The image in this irreducible quotient of a highest weight vector in M is again a highest weight vector.*

Proof. In general, given a ring R and a non-zero R-module M, let us notice that M admits a unique maximal proper R-submodule if and only if the sum of all proper R-submodules in M is not equal to M.

Given a highest weight $\mathfrak{g}$-module M with highest weight $\lambda \in \mathfrak{h}^*$, let us denote

$$M^\circ := \bigoplus_{\lambda' \in \mathrm{wt}(M) \smallsetminus \{\lambda\}} M_{\mathfrak{h},\lambda'} \subset M.$$

Then M° is a $\mathbb{C}$-linear subspace of M, and even a $\mathfrak{b}^-$-submodule (where we denote $\mathfrak{b}^- := \mathfrak{h} \oplus \mathfrak{n}^-$, this is a Lie subalgebra of $\mathfrak{g}$), but not a $\mathfrak{g}$-submodule in general. We have clearly $M^\circ \neq M$.

Now it is straight-forward that in order to prove the claim it is enough to see that every proper $\mathfrak{g}$-submodule of M is contained in M°.

Thus, let $N \subset M$ be a proper $\mathfrak{g}$-submodule. Notice that $\lambda \notin \mathrm{wt}(N)$, because otherwise we would have $M_{\mathfrak{h},\lambda} \subset N$ then $M = N$ (as any non-zero vector in $M_{\mathfrak{h},\lambda}$ generates M as a $\mathfrak{g}$-module). Thus (recall that a $\mathfrak{g}$-submodule of a weight module is a weight module),

$$N = \bigoplus_{\lambda' \in \mathrm{wt}(M) \smallsetminus \{\lambda\}} N_{\mathfrak{h},\lambda'} \subset M^\circ.$$

□

Definition 7.22. Let $\lambda \in \mathfrak{h}^*$. We denote by L_λ the unique irreducible quotient $\mathfrak{g}$-module of M_λ.

Corollary 7.23. *Let $\lambda \in \mathfrak{h}^*$. There exists an irreducible highest weight $\mathfrak{g}$-module with highest weight λ. It is unique up to an isomorphism. Our concrete model for it is L_λ.*

7.7 When is L_λ finite-dimensional?

Definition 7.24. Let $\lambda \in \mathfrak{h}^*$.

1. We say that λ is **integral** if $\lambda(H_\alpha) \in \mathbb{Z}$ for all $\alpha \in R^s$.
2. We say that λ is **dominant** if $\mathrm{Re}(\lambda(H_\alpha)) \in \mathbb{R}_{\geq 0}$ for all $\alpha \in R^s$.

Remark 7.25. Thus, $\lambda \in \mathfrak{h}^*$ is integral and dominant if $\lambda(H_\alpha) \in \mathbb{Z}_{\geq 0}$ for all $\alpha \in R^s$.

Remark 7.26. Sometimes a different condition in the definition of dominant $\lambda \in \mathfrak{h}^*$ is more appropriate. Namely, the condition that $\lambda(H_\alpha) \notin \mathbb{Z}_{<0}$ for all $\alpha \in R^s$. But we will not use it.

Exercise 7.11. *Let $\lambda \in \mathfrak{h}^*$. Let us write*

$$\lambda(\mathrm{diag}(x_1, \ldots, x_n)) = c_1 x_1 + \ldots + c_n x_n$$

for $c_1, \ldots, c_n \in \mathbb{C}$. Denote $d_i := c_i - c_{i+1}$. Then λ is integral if and only if $d_1, \ldots, d_{n-1} \in \mathbb{Z}$. Also, λ is dominant if and only if $\mathrm{Re}(d_1), \ldots, \mathrm{Re}(d_{n-1}) \in \mathbb{R}_{\geq 0}$. Given $\lambda \in \mathfrak{h}^*$, there exists a unique dominant $\lambda' \in W\lambda$.

We want to prove the following proposition:

Proposition 7.27. *Let $\lambda \in \mathfrak{h}^*$. Then L_λ is finite-dimensional if and only if λ is integral and dominant.*

Proof (of "only if" part of Proposition 7.27). Let $\lambda \in \mathfrak{h}^*$ and suppose that L_λ is finite-dimensional. Fix $\alpha \in R^s$ and consider L_λ as a finite-dimensional $\mathfrak{g}_\alpha$-module. Taking $0 \neq v \in (L_\lambda)_{\mathfrak{h},\lambda}$, we have $H_\alpha v = \lambda(H_\alpha) v$ and $E_\alpha v = 0$. Hence, by Lemma 5.2, we obtain $\lambda(H_\alpha) \in \mathbb{Z}_{\geq 0}$. □

Definition 7.28. Let $\mathfrak{k}$ be a Lie algebra and M an $\mathfrak{k}$-module. We say that a vector $v \in M$ is **$\mathfrak{k}$-finite** if the $\mathfrak{k}$-submodule of M generated by v is finite-dimensional. We say that M is a **locally finite** $\mathfrak{k}$-module if every $v \in M$ is $\mathfrak{k}$-finite.

Lemma 7.29. *Let $\mathfrak{h}$ be a finite-dimensional Lie algebra and $\mathfrak{k} \subset \mathfrak{h}$ a Lie subalgebra. Let M be a $\mathfrak{h}$-module. The subset of M consisting of $\mathfrak{k}$-finite vectors is a $\mathfrak{h}$-submodule of M.*

Proof. Let us denote the subset of M consisting of $\mathfrak{k}$-finite vectors by N. Clearly N is a linear subspace of M. We want to see that given $X \in \mathfrak{h}$ and $v \in N$ we have $Xv \in N$. Let $L \subset M$ be the $\mathfrak{k}$-submodule generated by v. By assumption L is finite-dimensional. Consider $L' := \mathfrak{h}L$ (the linear span in M of the subset of element of the form Zw where $Z \in \mathfrak{h}$ and $w \in L$). Then clearly L' is finite-dimensional and $Xv \in L'$. So it is left to check that L' is a $\mathfrak{k}$-submodule of M. Given $Y \in \mathfrak{k}$ and $Z \in \mathfrak{h}$ and $w \in L$, we have $YZw = [Y, Z]w + ZYw \in \mathfrak{h}L + \mathfrak{h}\mathfrak{k}L \subset \mathfrak{h}L + \mathfrak{h}L = L'$. □

Lemma 7.30. *Given $\alpha \in R^s$, let M be a locally finite $\mathfrak{g}_\alpha$-module. Then the action of H_α on M is diagnolizable, with eigenvalues lying in $\mathbb{Z}$. Let $n \in \mathbb{Z}_{\geq 0}$. Then $F_\alpha^n : M_{H_\alpha,n} \to M_{H_\alpha,-n}$ and $E_\alpha^n : M_{H_\alpha,-n} \to M_{H_\alpha,n}$ are isomorphisms of vector spaces.*

Proof. The statements hold when M is finite-dimensional by Corollary 5.12 and Claim 5.13. It is easy to see that this implies the statements in general. □

Lemma 7.31. *Let M be a $\mathfrak{g}$-module. If M is locally finite as a $\mathfrak{g}_\alpha$-module, then $s_\alpha(\mathrm{wt}(M)) = \mathrm{wt}(M)$.*

Proof. Let $\lambda \in \mathrm{wt}(M)$. Let $0 \neq v \in M_{\mathfrak{h},\lambda}$. Denote $n := \lambda(H_\alpha)$. Then $n \in \mathbb{Z}$ by Lemma 7.30. Suppose that $n \geq 0$. Consider $w := F_\alpha^n v$. By Lemma 7.30 we have $w \neq 0$. Also, $w \in M_{\mathfrak{h},\lambda - n\alpha}$. But $s_\alpha(\lambda) = \lambda - \lambda(H_\alpha)\alpha = \lambda - n\alpha$ and so $M_{\mathfrak{h},s_\alpha(\lambda)} \neq 0$ i.e. $s_\alpha(\lambda) \in \mathrm{wt}(M)$. Suppose now that $n < 0$. Then similarly $w := E_\alpha^n v$ satisfies $w \neq 0$ and $w \in M_{\mathfrak{h},s_\alpha(\lambda)}$, so $s_\alpha(\lambda) \in \mathrm{wt}(M)$. □

Lemma 7.32. *Let $\lambda \in \mathfrak{h}^*$ and let $\alpha \in R^s$ be such that $n := \lambda(H_\alpha) \in \mathbb{Z}_{\geq 0}$. Then $F_\alpha^{n+1} v_\lambda \in M_\lambda$ is a non-zero extremal vector with weight $\lambda - (n+1)\alpha$.*

Proof. Clearly, $F_\alpha^{n+1} v_\lambda$ is a weight vector with weight $\lambda - (n+1)\alpha$. By Lemma 7.19 it is non-zero. To see that it is extreme, by Lemma 7.11 it is enough to check that $E_\beta F_\alpha^{n+1} v_\lambda = 0$ for all $\beta \in R^s$. Let us first consider $\beta \neq \alpha$. Since in that case $\beta - \alpha \notin R$, we have $[E_\beta, F_\alpha] \in \mathfrak{g}_{\mathfrak{h},\beta-\alpha} = 0$ i.e. $[E_\beta, F_\alpha] = 0$. Working inside $\mathcal{U}(\mathfrak{g})$ this means that $E_\beta F_\alpha = F_\alpha E_\beta$. Iterating, this gives also $E_\beta F_\alpha^m = F_\alpha^m E_\beta$ for every $m \in \mathbb{Z}_{\geq 0}$. Thus

$$E_\beta F_\alpha^{n+1} v_\lambda = F_\alpha^{n+1} E_\beta v_\lambda = 0.$$

Now, we consider $\beta = \alpha$. We have

$$[E_\alpha, F_\alpha] = H_\alpha$$

i.e. (working in $\mathcal{U}(\mathfrak{g})$)

$$E_\alpha F_\alpha = F_\alpha E_\alpha + H_\alpha.$$

Assuming

$$E_\alpha (F_\alpha)^m v_\lambda = c_m \cdot (F_\alpha)^{m-1} v_\lambda$$

for some $c_m \in \mathbb{C}$ we find

$$E_\alpha (F_\alpha)^{m+1} v_\lambda = F_\alpha E_\alpha (F_\alpha)^m v_\lambda + H_\alpha (F_\alpha)^m v_\lambda = F_\alpha (c_m \cdot (F_\alpha)^{m-1} v_\lambda) + (\lambda - m\alpha)(H_\alpha) \cdot (F_\alpha)^m v_\lambda =$$

$$= (c_m + \lambda(H_\alpha) - 2m)(F_\alpha)^m v_\lambda.$$

Thus we get the recursive relation

$$c_{m+1} = \lambda(H_\alpha) + c_m - 2m,$$

where $c_0 = 0$. One deduces

$$c_m = m \cdot \lambda(H_\alpha) - 2\frac{m(m-1)}{2} = m\,(\lambda(H_\alpha) - (m-1)) = m(n - (m-1))$$

for all $m \in \mathbb{Z}_{\geq 0}$. Thus, $c_{n+1} = 0$ and therefore we have $E_\alpha F_\alpha^{n+1} v_\lambda = 0$. □

Corollary 7.33. *Let $\lambda \in \mathfrak{h}^*$ and let $\alpha \in R^s$ be such that $n := \lambda(H_\alpha) \in \mathbb{Z}_{\geq 0}$. Then, denoting by $[-] : M_\lambda \to L_\lambda$ the canonical projection, we have $F_\alpha^{n+1}[v_\lambda] = 0$.*

Proof. Denote by $N \subset M_\lambda$ the $\mathfrak{g}$-submodule generated by $F_\alpha^{n+1} v_\lambda$. Then N is a highest weight $\mathfrak{g}$-module, with highest weight $\lambda - (n+1)\alpha$. By Lemma 7.13 we have $\mathrm{wt}(N) \subset \{\mu \in \mathfrak{h}^* \mid \mu \leq \lambda - (n+1)\alpha\}$ and in particular $\lambda \notin \mathrm{wt}(N)$. Therefore $v_\lambda \notin N$ and so $N \neq M_\lambda$. Thus, since the kernel of the projection $[-] : M_\lambda \to L_\lambda$ is the unique maximal proper $\mathfrak{g}$-submodule of M_λ, we must have N sitting in this kernel. Since $F_\alpha^{n+1} v_\lambda \in N$, we obtain $0 = [F_\alpha^{n+1} v_\lambda] = F_\alpha^{n+1}[v_\lambda]$. □

Definition 7.34. Let us call a subset $S \subset \mathfrak{h}^*$ **conical** if there exists a finite subset $S_0 \subset \mathfrak{h}^*$ such that $S \subset S_0 - \sum_{\alpha \in R^s} \mathbb{Z}_{\geq 0} \cdot \alpha$.

Exercise 7.12. *Let V be a highest weight $\mathfrak{g}$-module. Then $\mathrm{wt}(V) \subset \mathfrak{h}^*$ is conical.*

Lemma 7.35. *A conical W-invariant subset of $\mathfrak{h}^*$ is finite.*

Proof. We will show that given a conical subset $S \subset \mathfrak{h}^*$, the subset $S_0 \subset S$ consisting of dominant elements is finite. Then, if S is W-invariant, we have $S = WS_0$ (by Exercise 7.11) and therefore S is finite as well. It is enough to fix $\lambda \in \mathfrak{h}^*$ and show that the subset of $\lambda - \sum_{\alpha \in R^s} \mathbb{Z}_{\geq 0} \cdot \alpha$ consisting of dominant elements is finite. Denote $H := \frac{1}{2} \sum_{\alpha \in R^+} H_\alpha$. Notice that if $\mu \in \mathfrak{h}^*$ is dominant we have $\mathrm{Re}(\mu(H)) \in \mathbb{R}_{\geq 0}$. Also, notice that, given $\alpha \in R^s$, we have $\alpha(H) = 1$. Therefore, considering collections $(m_\alpha)_{\alpha \in R^s}$ of elements in $\mathbb{Z}_{\geq 0}$, only for finitely many of them we have $\mathrm{Re}((\lambda - \sum_{\alpha \in R^s} m_\alpha \cdot \alpha)(H)) \in \mathbb{R}_{\geq 0}$ and therefore only for finitely many of them can $\lambda - \sum_{\alpha \in R^s} m_\alpha \cdot \alpha$ be dominant. □

Proof (of "if" part of Proposition 7.27). Let $\lambda \in \mathfrak{h}^*$ and suppose that λ is integral and dominant. We would like to see that, given $\alpha \in R^s$, L_λ is a locally finite $\mathfrak{g}_\alpha$-module. Given that, by Lemma 7.31 we obtain that $\mathrm{wt}(V)$ is s_α-invariant, and since $\{s_\alpha\}_{\alpha \in R^s}$ generates W, we deduce that $\mathrm{wt}(V)$ is W-invariant. Then Lemma 7.35 shows that $\mathrm{wt}(V)$ is finite, and hence V is finite-dimensional, as desired.

By Lemma 7.29 it is enough to check that v_λ is a $\mathfrak{g}_\alpha$-finite vector. Since we have $E_\alpha v_\lambda = 0$ and $H_\alpha v_\lambda = \lambda(H_\alpha) v_\lambda$, by what we have learned about highest weight modules, applied to $\mathfrak{g}_\alpha$, the $\mathfrak{g}_\alpha$-submodule of L_λ generated by v_λ coincides with the span of $\{F_\alpha^m v_\lambda\}_{m \in \mathbb{Z}_{\geq 0}}$. But, by Corollary 7.33 we have $F_\alpha^{n+1} v_\lambda = 0$ and therefore the $\mathfrak{g}_\alpha$-submodule of L_λ generated by v_λ is spanned by $\{F_\alpha^m v_\lambda\}_{0 \leq m \leq n}$, and so is finite-dimensional. So v_λ is a $\mathfrak{g}_\alpha$-finite vector. □

8 Formal character

8.1 Convolution

We denote by $\mathrm{Fun}(\mathfrak{h}^*)$ the vector space of functions on $\mathfrak{h}^*$. Given $\phi \in \mathrm{Fun}(\mathfrak{h}^*)$ we denote by $\mathrm{supp}(\phi) \subset \mathfrak{h}^*$ the support, i.e. the subset consisting of λ for which $\phi(\lambda) \neq 0$. Given $\mu \in \mathfrak{h}^*$ let us define $e^\mu \in \mathrm{Fun}(\mathfrak{h}^*)$ by $e^\mu(\mu) = 1$ and $e^\mu(\lambda) = 0$ if $\lambda \neq \mu$.

Given $\phi, \psi \in \mathrm{Fun}(\mathfrak{h}^*)$, let us say that ϕ and ψ are **convolutionable** if for every $\lambda \in \mathfrak{h}^*$ there exists finitely many $\mu \in \mathfrak{h}^*$ for which $\phi(\lambda + \mu) \neq 0$ and $\psi(-\mu) \neq 0$. Given two convolutionable $\phi, \psi \in \mathrm{Fun}(\mathfrak{h}^*)$ define the **convolution** $\phi \star \psi \in \mathrm{Fun}(\mathfrak{h}^*)$ by

$$(\phi \star \psi)(\lambda) := \sum_{\mu \in \mathfrak{h}^*} \phi(\lambda + \mu)\psi(-\mu).$$

For example, e^μ and every other function are convolutionable, and we have $(e^\mu \star \phi)(\lambda) = \phi(\lambda - \mu)$, i.e. convolution by e^μ shifts functions by μ. We have $e^{\mu_1} \star e^{\mu_2} = e^{\mu_1 + \mu_2}$.

We denote by $\mathrm{Fun}_{fin}(\mathfrak{h}^*) \subset \mathrm{Fun}_{con}(\mathfrak{h}^*) \subset \mathrm{Fun}(\mathfrak{h}^*)$ the linear subspaces consisting of functions ϕ for which $\mathrm{supp}(\phi)$ is, respectively, finite or conical.

Exercise 8.1.

1. *Check that the convolution of two convolutionable functions in* $\mathrm{Fun}(\mathfrak{h}^*)$ *is well-defined.*

2. *Check that convolution is* $\mathbb{C}$*-bilinear, associative and commutative whenever defined, and that* e^0 *is a neutral element with respect to convolution.*

3. *Check that every two functions in* $\mathrm{Fun}_{con}(\mathfrak{h}^*)$ *are convolutionable. Deduce that* $\mathrm{Fun}_{con}(\mathfrak{h}^*)$ *is a commutative unital* $\mathbb{C}$*-algebra with respect to convolution.*

8.2 Formal character

Definition 8.1. Let us say that a $\mathfrak{g}$-module M is **conical** if M is a weight module, $\mathrm{wt}(M) \subset \mathfrak{h}^*$ is conical, and $M_{\mathfrak{h},\lambda}$ is finite-dimensional for all $\lambda \in \mathfrak{h}^*$.

Example 8.2. *Finite-dimensional $\mathfrak{g}$-modules are conical, as well as highest weight $\mathfrak{g}$-modules.*

Definition 8.3. Let M be a conical $\mathfrak{g}$-module. We define its **formal character**

$$\mathrm{fch}_M \in \mathrm{Fun}_{con}(\mathfrak{h}^*)$$

by

$$\mathrm{fch}_M(\lambda) := \dim_{\mathbb{C}} M_{\mathfrak{h},\lambda}.$$

Our goal, Weyl's character formula, is a formula for fch_{L_λ} for a f.d. L_λ. We have a linear W-action on $\mathfrak{h}$, which induces a linear W-action on $\mathfrak{h}^*$, which in its turn induces a W-action on $\mathrm{Fun}(\mathfrak{h}^*)$. Notice that this W-action preserves $\mathrm{Fun}_{fin}(\mathfrak{h}^*)$, but does not preserve $\mathrm{Fun}_{con}(\mathfrak{h}^*)$.

Claim 8.4. *Let M be a finite-dimensional $\mathfrak{g}$-module. Then $\mathrm{fch}_M \in \mathrm{Fun}_{fin}(\mathfrak{h}^*)$ is W-invariant.*

Proof. Using the $\mathrm{SU}(n)$-action on M corresponding to the $\mathfrak{g}$-action, the claim is clear by what we saw about finite-dimensional $\mathrm{SU}(n)$-representations. However, we can also give an algebraic proof. Namely, fixing $\alpha \in R^s$, it is enough to check that fch_M is s_α-invariant. In other words, we want to check that $\dim M_{\mathfrak{h},\lambda} = \dim M_{\mathfrak{h},s_\alpha\lambda}$ for all $\lambda \in \mathfrak{h}^*$. By symmetry, it is enough to check that $\dim M_{\mathfrak{h},\lambda} \leq \dim M_{\mathfrak{h},s_\alpha\lambda}$ for all $\lambda \in \mathfrak{h}^*$. Fix $\lambda \in \mathfrak{h}^*$. If $\lambda \notin \mathrm{wt}(M)$ the claim is clear, so we assume $\lambda \in \mathrm{wt}(M)$. We again consider the $\mathfrak{g}_\alpha$-action on M. If Since $\lambda(H_\alpha)$ is an eigenvalue of H_α acting on M, by what we saw on finite-dimensional $\mathfrak{sl}_2$-modules, we have $\lambda(H_\alpha) \in \mathbb{Z}$. We assume $\lambda(H_\alpha) \geq 0$, as the other case is analogous. Denote $n := \lambda(H_\alpha)$. We have $\lambda - n\alpha = s_\alpha(\lambda)$ and thus the action of F_α^n on M sends $M_{\mathfrak{h},\lambda}$ into $M_{\mathfrak{h},s_\alpha(\lambda)}$. Recall that we saw that the action of F_α^n on M, restricted to $M_{H_\alpha,\lambda(H_\alpha)}$, is injective. Since $M_{\mathfrak{h},\lambda}$ is contained in $M_{H_\alpha,\lambda(H_\alpha)}$, we obtain that the action of F_α^n on M, restricted to $M_{\mathfrak{h},\lambda}$, is injective. Hence $\dim M_{\mathfrak{h},\lambda} \leq \dim M_{\mathfrak{h},s_\alpha\lambda}$ as desired. □

Now, let us study fch_{M_λ}.

Definition 8.5. Define the **Kostant function** $K \in \mathrm{Fun}_{con}(\mathfrak{h}^*)$ by

$$K(\lambda) = \left| \left\{ m : R^+ \to \mathbb{Z}_{\geq 0} \mid \lambda = - \sum_{\alpha \in R^+} m(\alpha) \cdot \alpha \right\} \right|.$$

In other words, $K(\lambda)$ is the number of ways that λ can be written as a $\mathbb{Z}_{\geq 0}$-linear combination of negative roots. We can also write:

$$K = \sum_{m : R^+ \to \mathbb{Z}_{\geq 0}} e^{-\sum_{\alpha \in R^+} m(\alpha) \cdot \alpha}.$$

Exercise 8.2. *We have*

$$K = \prod_{\alpha \in R^+} \left(e^0 + e^{-\alpha} + e^{-2\alpha} + \ldots\right)$$

(an expression similar to the ***Euler product*** *one encounters when studying L-functions) where the product is the convolution product* $\star$*. Here* $\left(e^0 + e^{-\alpha} + e^{-2\alpha} + \ldots\right)$ *has the obvious meaning - it is the function which is equal to* 1 *on one of the elements* $0, -\alpha, -2\alpha, \ldots$ *and to* 0 *elsewhere.*

Claim 8.6. *Let* $\lambda \in \mathfrak{h}^*$*. We have*

$$\mathrm{fch}_{M_\lambda} = e^\lambda \star K.$$

Proof. Recall that a basis for M_λ is given by

$$\{F^m \cdot v_\lambda\}_{m : R^+ \to \mathbb{Z}_{\geq 0}},$$

where the notation F^m is as follows. We choose an ordering of R^+, and then define F^m to be $\prod_{\alpha \in R^+} F_\alpha^{m(\alpha)}$, where the product is taken in the order we chose (this depends on the order, but we fix it). Therefore

$$\mathrm{fch}_{M_\lambda} = \sum_{m : R^+ \to \mathbb{Z}_{\geq 0}} e^{\lambda - \sum_{\alpha \in R^+} m(\alpha) \cdot \alpha} = e^\lambda \star \sum_{m : R^+ \to \mathbb{Z}_{\geq 0}} e^{-\sum_{\alpha \in R^+} m(\alpha) \cdot \alpha} = e^\lambda \star K.$$

□

Definition 8.7. We define

$$D := \prod_{\alpha \in R^+} (e^{\alpha/2} - e^{-\alpha/2}) \in \mathrm{Fun}_{fin}(\mathfrak{h}^*),$$

where the product is the convolution product $\star$.

Definition 8.8. We define

$$\rho := \frac{1}{2} \sum_{\alpha \in R^+} \alpha \in \mathfrak{h}^*.$$

Claim 8.9. *Let* $\lambda \in \mathfrak{h}^*$*. We have*

$$K \star D = e^\rho,$$

and so

$$\mathrm{fch}_{M_\lambda} \star (D \star e^{-(\lambda+\rho)}) = e^0.$$

In other words, fch_{M_λ}*, which lies in* $\mathrm{Fun}_{con}(\mathfrak{h}^*)$*, is the inverse with respect to the convolution product* $\star$ *of an element in* $\mathrm{Fun}_{fin}(\mathfrak{h}^*)$*, namely of* $D \star e^{-(\lambda+\rho)}$*.*

Proof. Notice that

$$(e^0 + e^{-\alpha} + e^{-2\alpha} + \ldots) \star (e^{\alpha/2} - e^{-\alpha/2}) = e^{\alpha/2}.$$

Therefore

$$K \star D = \prod_{\alpha \in R^+} e^{\alpha/2} = e^\rho.$$

□

8.3 Expressing the formal character of an irreducible module in terms of formal characters of Verma modules, given a fact

Definition 8.10 (Dot action)**.** We define a new action of W on $\mathfrak{h}^*$ by $w \bullet \lambda := w(\lambda + \rho) - \rho$. Thus, it is no more linear.

Definition 8.11. A **subquotient** of a module is a quotient module of a submodule or, which is the same, a submodule of a quotient module.

The following fact we will explain later.

Fact 8.12. *Let M be a highest weight $\mathfrak{g}$-module with highest weight $\lambda \in \mathfrak{h}^*$. Then $\mathrm{ext}(N) \subset W \bullet \lambda$ for any subquotient $\mathfrak{g}$-module N of M.*

We will prove Weyl's character formula, granted this fact.

Exercise 8.3. *A non-empty conical subset in $\mathfrak{h}^*$ has maximal elements w.r.t. our partial order $\leq$. As a corollary, given a non-zero conical $\mathfrak{g}$-module M we have $\mathrm{ext}(M) \neq \emptyset$.*

Claim 8.13. *Let $S \subset \mathfrak{h}^*$ be a finite subset. Let M be a conical $\mathfrak{g}$-module such that $\mathrm{ext}(N) \subset S$ for any subquotient $\mathfrak{g}$-module N of M. Then there exists a collection $(n_\mu)_{\mu \in S} \subset \mathbb{Z}_{\geq 0}$ such that*

$$\mathrm{fch}_M = \sum_{\mu \in S} n_\mu \cdot \mathrm{fch}_{L_\mu}.$$

Proof. The proof will be by induction on

$$n(M) := \sum_{\mu \in S} \dim_{\mathbb{C}} M_{\mathfrak{h},\mu}.$$

If $n(M) = 0$ then we obtain $\mathrm{ext}(M) = \emptyset$ and hence by Exercise 8.3 we obtain $M = 0$, and the claim is clear. Assume now $n(M) > 0$, and so $M \neq 0$ and hence $\mathrm{ext}(M) \neq \emptyset$. If M is irreducible, then take $\mu \in \mathrm{ext}(M)$ and take an extremal vector $0 \neq v \in M_{\mathfrak{h},\mu}$ Then v generates M as a $\mathfrak{g}$-module (since M is irreducible) and therefore we see that M is a highest weight module with highest weight μ, i.e. M is isomorphic to L_μ, and hence the claim is clear in this case. If M is not irreducible, let $N \subset M$ be a submodule with $N \neq 0$ and $N \neq M$. Then $n(M) = n(N) + n(M/N)$ and we deduce that $n(N) < n(M)$ and $n(M/N) < n(M)$, and therefore we can write $\mathrm{fch}(N)$ and $\mathrm{fch}(M/N)$ as desired, by induction. Since

$$\mathrm{fch}(M) = \mathrm{fch}(N) + \mathrm{fch}(M/N),$$

the claim is clear. □

Claim 8.14. *The elements*

$$(\mathrm{fch}_{L_\mu})_{\mu \in \mathfrak{h}^*}$$

in $\mathrm{Fun}_{con}(\mathfrak{h}^*)$ are linearly independent.

Proof. Suppose given a finite subset $S \subset \mathfrak{h}^*$ and scalars $(c_\mu)_{\mu \in S}$ such that $\sum_{\mu \in S} c_\mu \cdot \mathrm{fch}_{L_\mu} = 0$. Let us order $S = \{\mu_1, \ldots, \mu_n\}$ in such a way so that $\mu_i < \mu_j$ implies $i > j$. Fix $1 \le i \le n$ and suppose that we have already showed that $c_{\mu_j} = 0$ for $j < i$ - we want to show that $c_{\mu_i} = 0$ as well. Plugging in μ_i in our relation, we obtain then

$$\sum_{i \le j \le n} c_{\mu_j} \cdot \dim_{\mathbb{C}} (L_{\mu_j})_{\mathfrak{h}, \mu_i} = 0.$$

However, $\mathrm{wt}(L_{\mu_j}) \subset \{\mu \in \mathfrak{h}^* \mid \mu \le \mu_j\}$ and therefore $\mu_i \notin \mathrm{wt}(L_{\mu_j})$ for $j > i$. Therefore the equality becomes just $c_{\mu_i} \cdot \dim_{\mathbb{C}} (L_{\mu_i})_{\mathfrak{h}, \mu_i} = 0$, and since $\dim_{\mathbb{C}} (L_{\mu_i})_{\mathfrak{h}, \mu_i} = 1$ we obtain $c_{\mu_i} = 0$. □

Claim 8.15. *Let $\lambda \in \mathfrak{h}^*$.*

1. *We can write*
$$\mathrm{fch}_{M_\lambda} = \sum_{\mu \in W \bullet \lambda} n_\mu \cdot \mathrm{fch}_{L_\mu}$$
for $n_\mu \in \mathbb{Z}_{\ge 0}$. We have $n_\mu = 0$ unless $\mu \le \lambda$ and $n_\lambda = 1$.

2. *There exist integers $m_\mu \in \mathbb{Z}$ for $\mu \in W \bullet \lambda$ such that*
$$\mathrm{fch}_{L_\lambda} = \sum_{\mu \in W \bullet \lambda} m_\mu \cdot \mathrm{fch}_{M_\mu}.$$
We have $m_\mu = 0$ unless $\mu \le \lambda$ and $m_\lambda = 1$.

Proof.

1. Let us denote by $K \subset M_\lambda$ the kernel of the canonical projection $M_\lambda \to L_\lambda$. Then
$$\mathrm{fch}_{M_\lambda} = \mathrm{fch}_K + \mathrm{fch}_{L_\lambda}.$$
Let us set
$$S := (W \bullet \lambda) \cap \{\mu \in \mathfrak{h}^* \mid \mu < \lambda\}.$$
Clearly, taking Fact 8.12 into consideration, we have $\mathrm{ext}(N) \subset S$ for any subquotient $\mathfrak{g}$-module N on K. Therefore Claim 8.13 shows that
$$\mathrm{fch}_K \in \sum_{\mu \in S} \mathbb{Z}_{\ge 0} \cdot \mathrm{fch}_{L_\mu}$$
and the claim is clear.

2. This part follows from the previous part, by inverting a triangular matrix with 1's on the diagonal (as an exercise, formulate the precise claim and proof).

□

8.4 Proof of Weyl's character formula (given the fact)

We are now ready to prove Weyl's character formula (conditional on Fact 8.12). Let $\lambda \in \mathfrak{h}^*$ be integral and dominant, so that L_λ is finite-dimensional.

Lemma 8.16. *Let $\lambda \in \mathfrak{h}^*$ be integral and dominant. Then*

$$\{w \in W \mid w \bullet \lambda = \lambda\} = \{1\},$$

or in other words

$$\{w \in W \mid w(\lambda + \rho) = \lambda + \rho\} = \{1\}.$$

Proof. Let us write $\lambda(\mathrm{diag}(x_1, \ldots, x_n)) = c_1 x_1 + \ldots + c_n x_n$. That λ is integral and dominant means that $c_i - c_{i+1} \in \mathbb{Z}_{\geq 0}$ for all $1 \leq i < n$. We have

$$\rho(\mathrm{diag}(x_1, \ldots, x_n)) = \frac{n-1}{2} x_1 + \frac{n-3}{2} x_2 + \ldots + \frac{-(n-1)}{2} x_n.$$

Therefore, writing $(\lambda + \rho)(\mathrm{diag}(x_1, \ldots, x_n)) = d_1 x_1 + \ldots + d_n x_n$ we have $d_i - d_{i+1} = (c_i - c_{i+1}) + 1$ for all $1 \leq i < n$ and thus $d_i - d_{i+1} \in \mathbb{Z}_{\geq 1}$ for all $1 \leq i < n$. Thus, we have $d_1 > d_2 > \ldots > d_n$. Clearly, any non-trivial reordering of $(d_1, \ldots, d_n)$ does not satisfy this monotonicity property, and thus can't be defining the same functional on $\mathfrak{h}^*$ as $(d_1, \ldots, d_n)$. □

From this lemma we have a bijection $W \to W \bullet \lambda$ given by $w \mapsto w \bullet \lambda$. By Claim 8.15 we have $m : W \to \mathbb{Z}$ so that

$$\mathrm{fch}_{L_\lambda} = \sum_{w \in W} m(w) \cdot \mathrm{fch}_{M_{w \bullet \lambda}}.$$

By 8.6 we get

$$\mathrm{fch}_{L_\lambda} = \left(\sum_{w \in W} m(w) \cdot e^{w \bullet \lambda} \right) \star K.$$

By Claim 8.9 we get

$$\mathrm{fch}_{L_\lambda} \star D = \left(\sum_{w \in W} m(w) \cdot e^{w \bullet \lambda} \right) \star e^{\rho} = \sum_{w \in W} m(w) \cdot e^{w(\lambda + \rho)}. \tag{8.1}$$

We will use now the following lemma:

Lemma 8.17. *For every $w \in W$ we have $wD = \mathrm{sgn}(w) \cdot D$.*

Proof. Since $\{s_\alpha\}_{\alpha \in R^s}$ generate the group W, it is enough to show the equality for w being s_α for some $\alpha \in R^s$, i.e. to show that $s_\alpha D = -D$. We have

$$s_\alpha D = s_\alpha \left(\prod_{\beta \in R^+} (e^{\beta/2} - e^{-\beta/2}) \right) = \prod_{\beta \in R^+} (e^{s_\alpha(\beta)/2} - e^{-s_\alpha(\beta)/2}) =$$

$$= (e^{s_\alpha(\alpha)/2} - e^{-s_\alpha(\alpha)/2}) \cdot \prod_{\alpha \neq \beta \in R^+} (e^{s_\alpha(\beta)/2} - e^{-s_\alpha(\beta)/2}) =$$

$$= (e^{-\alpha/2} - e^{\alpha/2}) \cdot \prod_{\alpha \neq \beta \in R^+} (e^{\beta/2} - e^{-\beta/2}) = - \prod_{\beta \in R^+} (e^{\beta/2} - e^{-\beta/2}) = -D.$$

□

We continue. Note that we should be careful that the W-action on $\text{Fun}(\mathfrak{h}^*)$ does not preserve $\text{Fun}_{con}(\mathfrak{h}^*)$. However, it is straightforward that if $\phi, \psi \in \text{Fun}(\mathfrak{h}^*)$ are convolutionable, then $w\phi, w\psi$ are also convolutionable, and we have $w(\phi \star \psi) = (w\phi) \star (w\psi)$. We now see what we get when we apply some $w \in W$ to both sides of (8.1). The left side satisfies $w(\text{LEFT}) = \text{sgn}(w) \cdot (\text{LEFT})$. Hence also for the right side we have $w(\text{RIGHT}) = \text{sgn}(w) \cdot (\text{RIGHT})$. This gives:

$$\sum_{w' \in W} m(w') \cdot e^{ww'(\lambda+\rho)} = \text{sgn}(w) \cdot \sum_{w' \in W} m(w') \cdot e^{w'(\lambda+\rho)}.$$

Comparing coefficients we obtain $m(ww') = \text{sgn}(w) m(w')$ for all $w, w' \in W$. Recall that $m(1) = 1$ and therefore we deduce $m(w) = \text{sgn}(w)$ for all $w \in W$. We have obtained:

Theorem 8.18 (Weyl's character formula)**.** *Let $\lambda \in \mathfrak{h}^*$ be integral and dominant. Then*

$$\text{fch}(L_\lambda) \star \prod_{\alpha \in R^+} (e^{\alpha/2} - e^{-\alpha/2}) = \sum_{w \in W} \text{sgn}(w) \cdot e^{w(\lambda+\rho)}.$$

Exercise 8.4. *Understand how all what we saw regarding L_λ's recovers the classification of irreducible finite-dimensional* SU(n)*-representations and their Weyl character formula.*

8.5 The center of the universal enveloping algebra

Definition 8.19. We denote by $\mathcal{Z}(\mathfrak{g})$ the center of $\mathcal{U}(\mathfrak{g})$. It is a commutative $\mathbb{C}$-algebra.

Remark 8.20. It will turn out that $\mathcal{Z}(\mathfrak{g})$ is not too small, and useful. We could not quite grasp it when looking "inside" $\mathfrak{g}$ itself, and had to consider $\mathcal{U}(\mathfrak{g})$.

Definition 8.21. Let A be a commutative $\mathbb{C}$-algebra. We denote by $\text{Sp}(A)$ the set of morphisms of $\mathbb{C}$-algebras from A to $\mathbb{C}$, and call it the **spectrum** of A. We also refer to elements of $\text{Sp}(A)$ as **characters** of A.

Exercise-Definition 8.22. *Let $\zeta \in \text{Sp}(A)$. Let M be a $\mathfrak{g}$-module. We say that M has **infinitesimal character** ζ if for all $D \in \mathcal{Z}(\mathfrak{g})$ and $v \in M$ we have $Dv = \zeta(D) \cdot v$. If $M \neq 0$ then clearly M has at most one infinitesimal character, and if it has an infinitesimal character we denote it by ζ_M.*

Exercise 8.5. *Let M be a $\mathfrak{g}$-module. If every $D \in \mathcal{Z}(\mathfrak{g})$ acts on M by a scalar then M has infinitesimal character.*

Lemma 8.23. *A highest weight $\mathfrak{g}$-module has infinitesimal character.*

Proof. Let M be a highest weight $\mathfrak{g}$-module, and let $v \in M$ be a highest weight vector, with weight $\lambda \in \mathfrak{h}^*$. Let $D \in \mathcal{Z}(\mathfrak{g})$. Then, since $DX = XD$ for all $X \in \mathfrak{g}$, Dv is also a vector of weight λ. Since $M_{\mathfrak{h},\lambda}$ is one-dimensional, we must have $Dv = cv$ for some $c \in \mathbb{C}$. Now, if we consider the subspace $N \subset M$ consisting of w for which $Dw = cw$, notice that N is a $\mathfrak{g}$-submodule of M, and it contains v. Since v generates M as a $\mathfrak{g}$-module, we obtain $N = M$. Thus$D$ acts on M by a scalar. By Exercise 8.5 we obtain that M has infinitesimal character. $\square$

Let us denote by $\mathrm{Fun}(\mathfrak{h}^*)$ the $\mathbb{C}$-algebra of $\mathbb{C}$-valued functions on $\mathfrak{h}^*$ (with multiplication being pointwise). We define a map

$$\mathcal{S} : \mathcal{Z}(\mathfrak{g}) \to \mathrm{Fun}(\mathfrak{h}^*)$$

by

$$D \mapsto \zeta_{M_\lambda}(D).$$

It is clearly a $\mathbb{C}$-algebra morphism. Let us denote by $\mathrm{Pol}(\mathfrak{h}^*) \subset \mathrm{Fun}(\mathfrak{h}^*)$ the $\mathbb{C}$-subalgebra consisting of polynomial functions.

Lemma 8.24. *The image of $\mathcal{S}$ lies in* $\mathrm{Pol}(\mathfrak{h}^*)$.

Proof. Notice that from the PBW theorem we see that, inside $\mathcal{U}(\mathfrak{g})$, we can write

$$\mathcal{U}(\mathfrak{g}) = \mathcal{U}(\mathfrak{h}) + \mathcal{U}(\mathfrak{g})\mathfrak{n} + \mathfrak{n}^-\mathcal{U}(\mathfrak{g}).$$

Let us also choose a basis $H_1, \ldots, H_r$ for $\mathfrak{h}$. Given $D \in \mathcal{Z}(\mathfrak{g})$ we can write

$$D \in \sum_i c_i \cdot H_1^{m_i^1} \cdot \ldots \cdot H_r^{m_i^r} + \mathcal{U}(\mathfrak{g})\mathfrak{n} + \mathfrak{n}^-\mathcal{U}(\mathfrak{g})$$

where $c_i \in \mathbb{C}$ and $m_i^p \in \mathbb{Z}_{\geq 0}$. Next, recall that we have a decomposition

$$M_\lambda = \mathbb{C} \cdot v_\lambda \oplus M_\lambda^\circ$$

where

$$M_\lambda^\circ := \sum_{\mu \in \mathrm{wt}(M_\lambda) \smallsetminus \{\lambda\}} (M_\lambda)_{\mathfrak{h},\mu},$$

and notice that

$$\mathfrak{n}^-\mathcal{U}(\mathfrak{g})v_\lambda \subset M_\lambda^\circ.$$

Also, notice that $\mathcal{U}(\mathfrak{g})\mathfrak{n}v_\lambda = 0$. Hence we obtain

$$Dv_\lambda = \sum_i c_i \cdot H_1^{m_i^1} \cdot \ldots \cdot H_r^{m_i^r} v_\lambda = \left(\sum_i c_i \cdot \lambda(H_1)^{m_i^1} \cdot \lambda(H_r)^{m_i^r} \right) \cdot v_\lambda.$$

In other words, we obtain

$$\mathcal{S}(D)(\lambda) = \sum_i c_i \cdot \lambda(H_1)^{m_i^1} \cdot \lambda(H_r)^{m_i^r},$$

and this is indeed polynomial in λ. $\square$

Recall that W acts linearly on $\mathfrak{h}$, and therefore acts linearly on $\mathfrak{h}^*$. Recall in addition that we defined the **dot action** of W on $\mathfrak{h}^*$ by $w \bullet \lambda := w(\lambda + \rho) - \rho$. It is no more linear (just affine). We obtain an induced dot action on $\mathrm{Fun}(\mathfrak{h}^*)$ and on its subalgebra $\mathrm{Pol}(\mathfrak{h}^*)$. This is an action by algebra automorphisms. We denote by $\mathrm{Pol}(\mathfrak{h}^*)^{W\bullet} \subset \mathrm{Pol}(\mathfrak{h}^*)$ the subalgebra of W-invariants with respect to this dot action.

Claim 8.25. *The image of* $\mathcal{S}$ *lies in* $\mathrm{Pol}(\mathfrak{h}^*)^{W\bullet}$.

Proof (of Claim 8.25). Since $\{s_\alpha\}_{\alpha \in R^s}$ generates W, it is enough to check that, fixing $D \in \mathcal{Z}(\mathfrak{g})$ and $\alpha \in R^s$, we have $s_\alpha \bullet \mathcal{S}(D) = \mathcal{S}(D)$, i.e. $\mathcal{S}(X)(s_\alpha \bullet \lambda) = \mathcal{S}(X)(\lambda)$ for all $\lambda \in \mathfrak{h}^*$. Let us first assume that λ is such that $\lambda(H_\alpha) \in \mathbb{Z}_{\geq 0}$. In Lemma 7.32 that, denoting $n := \lambda(H_\alpha)$, we have that $F_\alpha^{n+1} v_\lambda \in M_\lambda$ is a non-zero extremal vector with weight $\lambda - (n+1)\alpha$. Therefore we obtain a non-zero morphism of $\mathfrak{g}$-modules $M_{\lambda-(n+1)\alpha} \to M_\lambda$. Denoting by N the (non-zero) image of this morphism, it is both a quotient module of the source and a submodule of the target. Therefore $\mathcal{Z}(\mathfrak{g})$ acts on it both by $\zeta_{M_{\lambda-(n+1)\alpha}}$ and by ζ_{M_λ}. Thus we must have $\zeta_{M_{\lambda-(n+1)\alpha}} = \zeta_{M_\lambda}$, and so $\mathcal{S}(D)(\lambda - (n+1)\alpha) = \mathcal{S}(D)(\lambda)$. Now, notice that

$$s_\alpha \bullet \lambda = ((\lambda + \rho) - (\lambda + \rho)(H_\alpha)\alpha) - \rho = \lambda - (n+1)\alpha$$

(where we have used $\rho(H_\alpha) = 1$ - check this). Hence the required.

Thus, we have obtained $\mathcal{S}(D)(s_\alpha \bullet \lambda) = \mathcal{S}(D)(\lambda)$ for all $\lambda \in \mathfrak{h}^*$ satisfying $\lambda(H_\alpha) \in \mathbb{Z}_{\geq 0}$. Since $\mathcal{S}(D)$ is a polynomial on $\mathfrak{h}^*$, this implies the equality for all $\lambda \in \mathfrak{h}^*$ - by an easy lemma: Let V be a f.d. $\mathbb{C}$-vector space and let f be a polynomial on V. Let $0 \neq \ell \in V^*$. If $f(v) = 0$ for all $v \in V$ satisfying $\ell(v) \in \mathbb{Z}_{\geq 0}$, then $f = 0$. □

The following is a fundamental theorem:

Theorem 8.26 (Harish-Chandra)**.** *The* $\mathbb{C}$*-algebra morphism*

$$\mathcal{S} : \mathcal{Z}(\mathfrak{g}) \to \mathrm{Pol}(\mathfrak{h}^*)^{W\bullet}$$

is an isomorphism.

Proof. Omitted. □

Let us now see how this material establishes Fact 8.12. Let $\lambda \in \mathfrak{h}^*$ and let N be a non-zero subquotient of M_λ. We want to see that $\mathrm{ext}(N) \subset W \bullet \lambda$. Let $\mu \in \mathrm{ext}(N)$. We obtain a non-zero morphism of $\mathfrak{g}$-modules $M_\mu \to N$. Now, since N is a subquotient of M_λ, $\mathcal{Z}(\mathfrak{g})$ acts on N by ζ_{M_λ}. Since we have a non-zero morphism of $\mathfrak{g}$-modules $M_\mu \to N$, reasoning as above we see that the operators in $\mathcal{Z}(\mathfrak{g})$ act on M_μ by the same scalars with which they act on N, and hence $\zeta_{M_\mu} = \zeta_N = \zeta_{M_\lambda}$. In other words, for any $D \in \mathcal{Z}(\mathfrak{g})$ we have $\mathcal{S}(D)(\lambda) = \mathcal{S}(D)(\mu)$. By Theorem 8.26 we see that all the polynomials on $\mathfrak{h}^*$ which are invariant under the dot-action of W have the same values on λ and on

μ. It is left to see that this implies that $\mu \in W \bullet \lambda$. Indeed, let us suppose that $\mu \notin W \bullet \lambda$ and show that then there exists a polynomial $f \in \mathrm{Pol}(\mathfrak{h}^*)$ which is invariant under the dot-action of W and such that $f(\mu) \neq f(\lambda)$. Clearly, for any disjoint finite subsets $S, T \subset \mathfrak{h}^*$ we can find a polynomial $f_0 \in \mathrm{Pol}(\mathfrak{h}^*)$ such that $f_0(\nu) = 0$ for all $\nu \in S$ and $f_0(\nu) \neq 0$ for all $\nu \in T$. We apply this to $S := W \bullet \lambda$ and $T := W \bullet \nu$ and then define $f \in \mathrm{Pol}(\mathfrak{h}^*)$ by $f(\nu) := \prod_{w \in W} f_0(w \bullet \nu)$ for $\nu \in \mathfrak{h}^*$. Then clearly f is invariant under the dot-action of W, and we have $f(\lambda) = 0$ and $f(\mu) \neq 0$, so that $f(\mu) \neq f(\lambda)$, as desired.

Example 8.27. *Let us see what Harish-Chandra's theorem says for* $\mathfrak{g} := \mathfrak{sl}_2$. *We identify* $\mathfrak{h}^*$ *with* $\mathbb{C}$ *by sending* λ *to* $\lambda(H)$. *This gives an identification of* $\mathrm{Pol}(\mathfrak{h}^*)$ *with the algebra of polynomials* $\mathbb{C}[z]$. *The dot-action of the non-trivial element* $s \in W$ *is given by* $s \bullet c = -2 - c$. *Then* $\mathrm{Pol}(\mathfrak{h}^*)^{W\bullet} \subset \mathrm{Pol}(\mathfrak{h}^*)$ *gets identified with* $\mathbb{C}[(z+1)^2] \subset \mathbb{C}[z]$. *Recall the Casimir element*

$$C = \frac{1}{8}(H^2 + 2H + 4FE) \in \mathcal{Z}(\mathfrak{g}).$$

We have

$$Cv_\lambda = \left(\frac{1}{8}(\lambda(H)^2 + 2\lambda(H))\right) \cdot v_\lambda.$$

Therefore, under our identification,

$$\mathcal{S}(C) = \frac{1}{8}(z^2 + 2z) = \frac{1}{8}((z+1)^2 - 1).$$

Thus we see that

$$1, C, C^2, \ldots$$

is a basis for $\mathcal{Z}(\mathfrak{g})$.