Representation Theory

Lectures at the HU Berlin

Winter 2017/18

Thomas Krämer

21/2/2018

Contents

Introduction

Depending on whom you ask, the question *"What is representation theory?"* may result in very different answers:

(1) *An upgrade to linear algebra.* Technically, a representation of a group G is just an action of the group on a finite-dimensional vector space V over a field k, i.e. a homomorphism

$$\rho: \quad G \longrightarrow Gl(V) \;=\; \big\{\, g \in End_k(V) \mid \det(g) \neq 0 \,\big\}.$$

This seems a rather general, unspecific definition. Why should we care?

(2) *The study of concrete incarnations of abstract groups.* While it may be hard to study a given group abstractly, its conrete matrix representations are amenable to multilinear algebra: We can take direct sums, tensor products, symmetric and exterior powers, kernels and cokernels of morphisms, ...

$$\oplus, \quad \otimes, \quad Sym^n(-), \quad Alt^n(-), \quad \ker(-), \quad \mathrm{cok}(-), \quad \cdots$$

This gives a rich structure which contains a lot of information about the group.

(3) *A theory of everything.* Going one step further, one can say that since the notion of a representation is so unspecific, it appears naturally in almost any area of mathematics and physics. Even though it requires virtually no prerequisites, its applications are ubiquitous and unlimited in number!

Let me illustrate the last point with a few examples involving finite groups:

Example (A). The easiest nontrivial groups are finite abelian groups. Any such group is a product of cyclic groups $G = \mathbb{Z}/n\mathbb{Z}$ with $n \in \mathbb{N}$, each of whose complex representations splits into a direct sum of one-dimensional subrepresentations, the characters

$$\rho: \quad G \;=\; \mathbb{Z}/n\mathbb{Z} \;\longrightarrow\; \mathbb{C}^\times \;=\; Gl_1(\mathbb{C}), \quad (a \mod n) \;\mapsto\; \zeta^a$$

where $\zeta \in \mathbb{C}^\times$ is an n-th root of unity. The theory of such characters is completely elementary but nevertheless a crucial ingredient (besides complex analysis of course) for the famous

Theorem (Dirichlet's theorem on primes in arithmetic progressions). *For any given $a, b \in \mathbb{N}$ with $gcd(a,b) = 1$, there exist infinitely many prime numbers of the form*

$$p \;=\; an + b \quad \textit{with} \quad n \;\in\; \mathbb{N}.$$

For instance, the numbers

$$p \;=\; 3,\, 13,\, 23,\, 43,\, 53,\, 73,\, 83,\, 103,\, 113,\, \ldots$$

are all prime and this list can be extended infinitely to the right.

Example (B). Next to abelian groups, the easiest examples of groups are finite solvable groups. Here a finite group G is called *solvable* if it can be built as a successive extension of finite abelian groups, i.e. if there exists an ascending series of subgroups

$$\{1\} = G_0 \trianglelefteq G_1 \trianglelefteq G_2 \trianglelefteq \cdots \trianglelefteq G_n = G,$$

each being a normal subgroup of the following one, such that the quotients G_i/G_{i-1} are abelian for $i = 1, 2, \dots, n$. The following famous structure theorem in the classification of finite groups was proved by Burnside in the early 20th century via representation theory — more precisely, via the character theory of finite groups:

Theorem (Burnside's theorem). *Let G be any finite group whose order $|G|$ is divisible by at most two different prime numbers. Then G is solvable.*

For example, the symmetric group $\mathfrak{S}_4$ is solvable. This can of course be checked by hand easily: We have the ascending series of subgroups $\{1\} \trianglelefteq \mathfrak{V}_4 \trianglelefteq \mathfrak{A}_4 \trianglelefteq \mathfrak{S}_4$ where

$$\mathfrak{V}_4 \;=\; \{1, (12)(34), (13)(24), (1234)\} \;\simeq\; \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$$

is the *Klein'sche Vierergruppe*, and a look at the group orders immediately shows that all the successive quotients are abelian. On the other hand, the symmetric group $\mathfrak{S}_5$ is not solvable: In fact its only normal subgroup is $\mathfrak{A}_5 \trianglelefteq \mathfrak{S}_5$ and $\mathfrak{A}_5$ is known to be *simple*, i.e. nonabelian without any nontrivial normal subgroups. As the order $|\mathfrak{S}_5| = 5!$ is divisible only by the three primes $2, 3$ and 5, it follows in particular that Burnside's theorem is sharp.

A much deeper result which relies in a crucial way on character theory is the celebrated theorem of *Feit-Thompson*, which says that every group of odd order is solvable. Its proof in 1963 took no less than $2^8 - 1$ pages, covered a whole issue of the *Pacific Journal of Mathematics* and triggered the classification of finite simple groups. Needless to say, we will not be able to any of these more recent advances in this course; but we will see a proof of Burnside's theorem.

Example (C). When being introduced to complex numbers, everybody gets to know the formal identity

$$(x_1^2 + x_2^2)(y_1^2 + y_2^2) \;=\; (x_1 y_1 - x_2 y_2)^2 + (x_1 y_2 + x_2 y_1)^2$$

which expresses the multiplicativity of the absolute value. In a similar vein, in Hamilton's quaternions $\mathbb{H}$ the multiplicativity of the norm comes from the formal identity

$$\begin{aligned}(x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2) \;&=\; (x_1 y_1 - x_2 y_2 - x_3 y_3 - x_4 y_4)^2 \\ &+\; (x_1 y_2 + x_2 y_1 + x_3 y_4 - x_4 y_3)^2 \\ &+\; (x_1 y_3 + x_3 y_1 - x_2 y_4 + x_4 y_2)^2 \\ &+\; (x_1 y_4 + x_4 y_1 - x_2 y_3 - x_3 y_2)^2\end{aligned}$$

and there is a similar identity in eight variables corresponding to the multiplicativity of the norm in Cayley's octonions $\mathbb{O}$. But that's it:

Theorem (Hurwitz, 1898). *The values $n = 1, 2, 4, 8$ are the only natural numbers for which there is a formal identity*

$$(x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n) \;=\; (\ell_1(x, y))^2 + \cdots + (\ell_n(x, y))^2$$

involving bilinear functions

$$\ell_\nu(x,y) \;=\; \sum_{i,j=1}^{n} a_{ij\nu}\cdot x_i y_j \quad \textit{with coefficients} \quad a_{ij\nu}\in\mathbb{C}.$$

In fact the theorem holds with coefficients $a_{ij\nu}$ in any base field k of characteristic $char(k)\neq 2$; of course for $char(k)=2$ it trivially fails since then any sum of squares is again a square. Following an idea of Eckmann, the result of Hurwitz can also be deduced from character theory.

Example (D). In geometry, one is often interested in actions of some group G on a topological space X. In other words, to each group element $g\in G$ one assigns a continuous map

$$\rho(g): \quad X \longrightarrow X$$

such that

$$\rho(1) \;=\; id_X \quad \text{and} \quad \rho(gh)=\rho(g)\circ\rho(h) \quad \text{for all} \quad g,h\in G.$$

The induced linear maps

$$\rho(g): \quad H_i(X,\mathbb{Q}) \longrightarrow H_i(X,\mathbb{Q})$$

then define representations

$$\rho: \quad G \longrightarrow Gl(V) \quad \text{on the homology groups} \quad V \;=\; H_i(X,\mathbb{Q})$$

for each $i\in\mathbb{Z}$. These representations encode interesting information; for instance, in favorable situations the invariant part $V^G=\{v\in V \mid \rho(g)(v)=v\;\forall\, g\in G\}\subseteq V$ can be identified with $H_i(X/G,\mathbb{Q})$, the homology of the quotient space.

The goal of these lectures is to develop the basic framework to understand the above examples (and many more). In the first part of the course we will see that every representation of a given group G can be decomposed into smallest pieces that are called *irreducible* representations. The main goals of representation theory can then be described as follows:

- Classify all the irreducible representations of G.
- Given an arbitrary representation, find the rules according to which one may determine its decomposition into irreducible pieces; for example, how do tensor products, symmetric or exterior powers of a given representation decompose into irreducibles?

In the second part of the course, we will turn to finite groups and show that here the above goals are achieved by a device called *character theory.* We will then take a closer look at *symmetric groups*, where everything can be made completely explicit via the combinatorics of so-called Young diagrams, and we conclude with a brief outlook on the representation theory of compact Lie groups.

CHAPTER I

Basic notions of representation theory

This chapter provides some basic notions that are common to all incarnations of representaiton theory. In what follows we fix a group G and an arbitrary base field k over which our representations will be defined, and unless stated otherwise all vector spaces will be assumed to be finite-dimensional over k.

1. The category of representations

A representation of a group is just an action of it by linear transformations on some vector space. From a categorical point of view it is desirable to include the underlying vector space and not just its general linear group in the definition:

Definition 1.1. A *representation* of G is a pair (V, ρ), where V is a vector space and

$$\rho : \quad G \longrightarrow Gl(V) \;=\; \{ f : V \to V \mid f \text{ is } k\text{-linear and invertible} \}$$

is a homomorphism into the general linear group of the vector space. In other words, for each group element $g \in G$ we are given a linear automorphism $\rho(g) \in GL(V)$ such that

$$\rho(1) \;=\; id_V \quad \text{and} \quad \rho(gh) \;=\; \rho(g) \circ \rho(h) \quad \text{for all} \quad g, h \;\in\; G.$$

If the group action is clear from the context, we suppress it from the notation and write $V = (V, \rho)$. Similarly we put

$$gv \;:=\; \rho(g)v \;:=\; (\rho(g))(v) \quad \text{for} \quad v \;\in\; V, \quad g \;\in\; G.$$

The *dimension* of a representation (V, ρ) is defined to be the dimension $\dim(V)$ of the underlying vector space over the base field k.

Example 1.2. (a) Every 1-dimensional representation corresponds to a group homomorphism

$$\rho : \quad G \longrightarrow k^\times = Gl_1(k)$$

and conversely. In the special case of the trivial homomorphism $\rho(g) = 1$ for all g we also call this representation the *trivial representation*. For example, for the cyclic group $G = \mathbb{Z}/n\mathbb{Z}$ with $n \geq 0$, any element $\zeta \in k^\times$ with $\zeta^n = 1$ gives rise to a representation

$$\rho : \quad G \;=\; \mathbb{Z}/n\mathbb{Z} \longrightarrow k^\times \quad \text{via} \quad \rho(a \bmod n) \;=\; \zeta^a \quad \text{for} \quad a \;\in\; \mathbb{Z}.$$

and this representation is trivial iff $\zeta = 1$. For $k = \mathbb{C}$ and any $a \in \mathbb{C}$, the additive group $G = (\mathbb{R}, +)$ has the 1-dimensional representations

$$\rho : \quad G \longrightarrow \mathbb{C}^\times \quad \text{given by} \quad \rho(x) \;=\; e^{2\pi i a x}$$

and these representations are important in Fourier analysis. For $a \in \mathbb{Z}$ they can also be viewed as representations of the group $\mathbb{R}/\mathbb{Z}$. As a word of caution, outside of representation theory the word *character* is usually used for a 1-dimensional representation. We avoid this terminology since it will conflict with the notion of characters in the sense of representation theory to be introduced later.

(b) If V is a vector space, then any subgroup $G \hookrightarrow Gl(V)$ comes with a natural representation where the action is defined by the inclusion map. We also refer to this as the *natural representation of* G if the context is clear. For instance, the special linear group $Sl_n(k)$ has a natural representation of dimension n.

(c) If G is a finite group, we define the *coordinate ring of* G to be the vector space

$$k[G] \;=\; \{f : G \longrightarrow k\}$$

of all functions from the group to the base field, with pointwise addition and scalar multiplication. This vector space is equipped with two representations: The right respectively left translation of functions leads to the *right regular representation* ρ and the *left regular representation* λ, defined by

$$(\rho(g)(f))(x) \;=\; f(xg) \quad \text{and} \quad (\lambda(g)(f))(x) \;=\; f(g^{-1}x) \quad \text{for} \quad f \in k[G],\ g, x \in G.$$

These representations may be written in coordinates as follows: The characteristic functions

$$e_g(x) \;:=\; \begin{cases} 1 & \text{if } x = g \\ 0 & \text{otherwise} \end{cases}$$

form a basis of the coordinate ring indexed by the elements $g \in G$, and in this basis we have

$$\rho(h)e_g \;=\; e_{gh^{-1}} \quad \text{and} \quad \lambda(h)e_g \;=\; e_{hg} \quad \text{for} \quad g, h \;\in\; G.$$

Albeit they look different, the right and the left regular representations essentially carry the same information. In fact they are *isomorphic* in the following sense:

Definition 1.3. A *morphism* between representations (V_i, ρ_i) of G is a linear map $f : V_1 \to V_2$ such that the diagram

$$\begin{array}{ccc} V_1 & \xrightarrow{f} & V_2 \\ \downarrow{\scriptstyle \rho_1(g)} & & \downarrow{\scriptstyle \rho_2(g)} \\ V_1 & \xrightarrow{f} & V_2 \end{array}$$

commutes for all $g \in G$. We also call such a morphism a G*-equivariant map.* The set of all such forms a vector space

$$\mathrm{Hom}_G(V_1, V_2) \;=\; \mathrm{Hom}((V_1, \rho_1), (V_2, \rho_2)),$$

and the composition of morphisms

$$\mathrm{Hom}_G(V_1, V_2) \times \mathrm{Hom}_G(V_2, V_3) \;\longrightarrow\; \mathrm{Hom}_G(V_1, V_3), \quad (f, g) \;\mapsto\; g \circ f$$

is k-bilinear. So altogether the representations of G form a k-linear category that we will denote by

$$\mathrm{Rep}_k(G).$$

By an *isomorphism* of representations we mean an invertible morphism in $\mathrm{Rep}_k(G)$, i.e. a morphism whose underlying linear map is invertible.

Exercise 1.4. Show that for any finite group, the right and the left regular representations are isomorphic via the linear map $e_g \mapsto e_{g^{-1}}$.

Starting from a given set of representations and morphisms between them, we get many others from linear algebra:

DEFINITION 1.5. The *direct sum* of two representations $(V_i, \rho_i) \in \mathrm{Rep}_k(G)$ is defined to be

$$(V_1, \rho_1) \oplus (V_2, \rho_2) \;:=\; (V_1 \oplus V_2, \rho_1 \oplus \rho_2),$$

i.e. the representation whose underlying vector space is the direct sum of the two respective vector spaces, with the group action $\rho_1 \oplus \rho_2$ defined by block diagonal matrices

$$\rho_1 \oplus \rho_2 : \quad G \xrightarrow{(\rho_1,\rho_2)} Gl(V_1) \times Gl(V_2) \hookrightarrow Gl(V_1 \oplus V_2), \quad g \;\mapsto\; \begin{pmatrix} \rho_1(g) & 0 \\ 0 & \rho_2(g) \end{pmatrix}.$$

In other words,

$$((\rho_1 \oplus \rho_2)(g))(v_1, v_2) \;:=\; (\rho_1(g)(v_1), \rho_2(g)(v_2)) \quad \text{for all} \quad v_i \in V_i,\; g \in G.$$

This often allows to split up a given representation into simpler pieces:

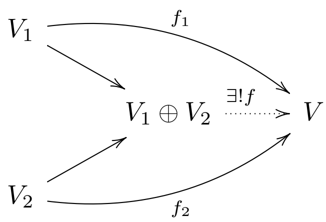
EXERCISE 1.6. Using the Jordan canonical form or otherwise, show that every finite set of commuting matrices of finite order in $Gl_d(\mathbb{C})$ may be diagonalized simultaneously. Deduce that every complex representation of a finite abelian group is isomorphic to a direct sum of 1-dimensional representations. How does this look like explicitly for the regular representation of a finite cyclic group?

The above notion of a direct sum is a special case of a general construction from category theory — it has the universal property of a *coproduct*, which by Yoneda's lemma characterizes it uniquely up to isomorphism:

LEMMA 1.7. *Let $V_1, V_2 \in \mathrm{Rep}_k(G)$. Then for every $V \in \mathrm{Rep}_k(G)$ there exists a natural isomorphism*

$$\mathrm{Hom}_G(V_1 \oplus V_2, V) \;\xrightarrow{\sim}\; \mathrm{Hom}_G(V_1, V) \oplus \mathrm{Hom}_G(V_2, V).$$

Proof. The map from left to right is defined by composition with the inclusion of the summands. So the claim follows from the diagram



by the same arguments as in linear algebra, noting that f is G-equivariant iff f_1 and f_2 both are. □

As in linear algebra, we can also form the kernel and cokernel of a morphism; for simplicity we drop the group action ρ from the notation:

DEFINITION 1.8. A *subrepresentation* of a representation V is a subspace $U \subseteq V$ that is G-stable in the sense that

$$gu \in U \quad \text{for all} \quad g \in G.$$

Clearly ρ then restricts to a natural representation on this subspace. We define the *quotient representation* to be the quotient vector space V/U with the induced group action

$$g\Big(v \bmod U\Big) \;:=\; \Big(gv \bmod U\Big) \quad \text{for } v \in V \text{ and } g \in G.$$

For example, for any representation we define the *invariants* and the *coinvariants* to be the biggest trivial sub- resp. quotient representation, and we denote these by

$$V^G \;=\; \big\{v \in V \mid gv = v \text{ for all } g \in G\big\} \quad \text{resp.} \quad V_G \;=\; V/\langle gv - v \mid v \in V, g \in G\rangle.$$

Similarly, for every morphism of representations $f : V_1 \longrightarrow V_2$ the *kernel* and the *image*

$$\begin{aligned} \ker(f) \;&=\; \big\{v \in V_1 \mid f(v) = 0\big\} \;\subseteq\; V_1 \\ \operatorname{im}(f) \;&=\; \big\{f(v) \in V_2 \mid v \in V_1\big\} \;\subseteq\; V_2 \end{aligned}$$

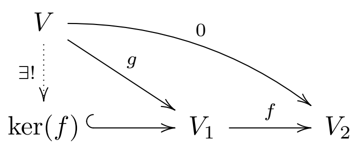
are subrepresentations of the source and target, respectively. The corresponding quotient representations

$$\operatorname{coim}(f) \;=\; V_1/\ker(f) \quad \text{and} \quad \operatorname{coker}(f) \;=\; V_2/\operatorname{im}(f)$$

are called the *coimage* and *cokernel* of the morphism. Like direct sums, the kernel and cokernel may be characterized uniquely by a universal property:

LEMMA 1.9. *Let* $V \in \operatorname{Rep}_k(G)$.

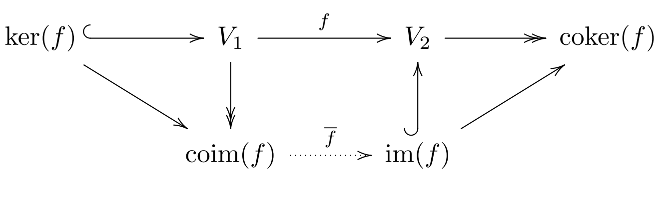
(1) *Any* $g \in \operatorname{Hom}_G(V, V_1)$ *with* $f \circ g = 0$ *factors uniquely over* $\ker(f) \subseteq V_1$ *as in the following diagram:*



(2) *Any* $h \in \operatorname{Hom}_G(V_2, V)$ *with* $h \circ f = 0$ *factors uniquely over* $V_2 \twoheadrightarrow \operatorname{coker}(f)$ *in the diagram dual to the previous one.*

Proof. Again the proof is the same as in linear algebra, noting that in our case all the involved morphisms are G-equivariant. □

COROLLARY 1.10. *For any morphism* $f \in \operatorname{Hom}_G(V_1, V_2)$*, there exists a natural factorization*



Furthermore, here the induced morphism $\overline{f}$ *is an isomorphism.*

Proof. The existence of such a factorization follows by applying the universal property of the kernel and cokernel to

$$\operatorname{coim}(f) \;=\; \operatorname{coker}\Big(\ker(f) \hookrightarrow V_1\Big) \quad \text{and} \quad \operatorname{im}(f) \;=\; \ker\Big(V_2 \twoheadrightarrow \operatorname{coker}(f)\Big).$$

We then get from linear algebra that $\overline{f}$ is an isomorphism. □

Summarizing the above discussion, in the language of category theory we have shown that $\operatorname{Rep}_k(G)$ is a *k-linear abelian category.* In particular, the usual results from homological algebra such as the 5-lemma or the snake lemma apply in this context:

DEFINITION 1.11. We say that f is a *monomorphism* if $\ker(f) = \{0\}$, resp. an *epimorphism* if $\mathrm{coker}(f) = \{0\}$. Note that f is an isomorphism in the previous sense iff it is both a monomorphism and an epimorphism. We say that a sequence of morphisms

$$\cdots \longrightarrow V_{i-1} \xrightarrow{f_{i-1}} V_i \xrightarrow{f_i} V_{i+1} \longrightarrow \cdots$$

is *exact* at the i-th position if $\mathrm{im}(f_{i-1}) = \ker(f_i)$, and by a *short exact sequence* we mean an exact sequence of the form

$$0 \longrightarrow V_1 \xrightarrow{f_1} V_2 \xrightarrow{f_2} V_3 \longrightarrow 0.$$

We say that the exact sequence *splits* if there exists an isomorphism $V_2 \simeq V_1 \oplus V_3$ such that f_1 and f_2 are the natural inclusion respectively projection maps. Unlike in linear algebra, not every exact sequence splits:

EXAMPLE 1.12. Define a representation of the additive group $\mathbb{G}_a = (k, +)$ on the vector space $V = k^2$ by

$$\rho : \quad \mathbb{G}_a \longrightarrow Gl(V) = Gl_2(k), \quad x \mapsto \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

Then we have an exact sequence

$$0 \longrightarrow U \longrightarrow V \longrightarrow W \longrightarrow 0$$

where $U = ke_1 \subset V$ denotes the 1-dimensional trivial subrepresentation which is spanned by the first standard basis vector. The action on the quotient $W = V/U$ is also trivial as $x(e_2 \bmod U) = (e_2 + xe_1 \bmod U) = (e_2 \bmod U)$ for all $x \in \mathbb{G}_a$. So the above exact sequence does not split, since otherwise the group action on the middle term would have to be trivial as well.

EXAMPLE 1.13. For a finite set S, consider the vector space $V = \bigoplus_{s \in S} ke_s$ freely generated by this set, where we denote by $e_s \in V$ the basis vector attached to $s \in S$. Then any group action

$$G \times S \longrightarrow S, \quad (g, s) \mapsto gs$$

defines a representation

$$\rho : \quad G \longrightarrow Gl(V) \quad \text{by} \quad \rho(g)(e_s) := e_{gs} \quad \text{for} \quad s \in S,$$

the *permutation representation* corresponding to the given group action. Note that this construction generalizes the regular representation of a finite group that we introduced earlier. Again we have an exact sequence

$$0 \longrightarrow U \longrightarrow V \longrightarrow W \longrightarrow 0$$

where $U = ke$ is the trivial 1-dimensional subrepresentation that is spanned by the G-invariant vector

$$e = \sum_{s \in S} e_s,$$

but now the quotient is non-trivial and the sequence splits if $char(k) = 0$ (exercise).

We will see later that for representations of finite groups, or more generally of compact Lie groups, *every* short exact sequence splits; this will lead to a very simple classification of all representations in terms of character theory. Permutation representations will reappear frequently, in particular for symmetric groups.

2. Tensor products and multilinear algebra

The k-linear abelian category $\mathrm{Rep}_k(G)$ from the previous section comes with a very rich additional structure: The tensor product, which allows for multilinear algebra constructions such as symmetric and exterior powers.

DEFINITION 2.1. For $(V_i, \rho_i) \in \mathrm{Rep}_k(G)$, we define the *tensor product* to be the representation

$$(V_1, \rho_1) \otimes (V_2, \rho_2) \ := \ (V_1 \otimes V_2, \rho_1 \otimes \rho_2)$$

where the group action on the tensor product is defined by the Kronecker product of matrices

$$\rho_1 \otimes \rho_2 : \quad G \xrightarrow{(\rho_1, \rho_2)} Gl(V_1) \times Gl(V_2) \longrightarrow Gl(V_1 \otimes V_2),$$

in other words

$$((\rho_1 \otimes \rho_2)(g))(v_1 \otimes v_2) \ := \ \rho_1(g)(v_1) \otimes \rho_2(g)(v_2) \quad \text{for all} \quad v_i \in V_i, \ g \in G.$$

We have the usual compatibility properties:

$$V_1 \otimes (V_2 \oplus V_3) \simeq V_1 \otimes V_2 \oplus V_1 \otimes V_3, \quad V_1 \otimes V_2 \simeq V_2 \otimes V_1, \quad (V_1 \otimes V_2) \otimes V_3 \simeq V_1 \otimes (V_2 \otimes V_3),$$

etc. We also have a notion of duality:

DEFINITION 2.2. For $V \in \mathrm{Rep}_k(G)$, the *dual* or *contragredient representation* is the dual vector space

$$V^* \ = \ \mathrm{Hom}_k(V, k)$$

with the action given by

$$(gf)(v) \ = \ f(g^{-1}v) \quad \text{for} \quad f \in V^*, v \in V, g \in G.$$

Note that the definition of the group action on the dual is made precisely so that the evaluation

$$V^* \otimes V \ \longrightarrow \ k, \quad (f, v) \ \mapsto \ f(v)$$

becomes a morphism to the trivial 1-dimensional representation.

LEMMA 2.3. *The dual has the following formal properties:*

1. *Functoriality: Any morphism* $f : V_1 \to V_2$ *induces* $f^* : V_2^* \to V_1^*$.
2. *Involutivity: There is a natural isomorphism* $V^{**} \simeq V$.
3. *Compatibility:* $(V_1 \oplus V_2)^* \simeq V_1^* \oplus V_2^*$ *and* $(V_1 \otimes V_2)^* \simeq V_1^* \otimes V_2^*$.
4. *Adjunction:* $\mathrm{Hom}_G(V_1 \otimes V_2, V_3) \simeq \mathrm{Hom}_G(V_1, V_2^* \otimes V_3)$.

Proof. Exercise. □

In abstract modern language, the above essentially says that the k-linear abelian category $\mathrm{Rep}_k(G)$ is a so-called *tensor category*. We will not formalize this notion further in these lectures but only concentrate on its concrete consequences: The tensor formalism allows to carry over constructions from multilinear algebra to our present setting.

EXAMPLE 2.4. From linear algebra we have a natural isomorphism of vector spaces

$$\mathrm{Hom}(V_1, V_2) \ \simeq \ V_1^* \otimes V_2,$$

and since by the above the right hand side is endowed with a representation of G, we can view $\mathrm{Hom}(V_1, V_2)$ as an object of $\mathrm{Rep}_k(G)$ in a natural way. This object is called the *inner Hom* in the category of representations. From it the morphisms in $\mathrm{Rep}_k(G)$ may be recovered as the G-invariants

$$\mathrm{Hom}_G(V_1, V_2) \ = \ \mathrm{Hom}(V_1, V_2)^G$$

and we have natural adjunction isomorphisms

$$\mathrm{Hom}_G(V_1 \otimes V_2, V_3) \ \simeq \ \mathrm{Hom}_G(V_1, \mathrm{Hom}(V_2, V_3))$$

which explain the name *inner Hom*. Explicitly, the group action ρ on $\mathrm{Hom}(V_1, V_2)$ is defined precisely so that the following diagram commutes for all $g \in G$ and all linear maps $f \in \mathrm{Hom}(V_1, V_2)$,

$$\begin{array}{ccc} V_1 & \xrightarrow{\ f\ } & V_2 \\ {\scriptstyle \rho_1(g)}\downarrow & & \downarrow{\scriptstyle \rho_2(g)} \\ V_1 & \xrightarrow{\ \rho(g)(f)\ } & V_2 \end{array}$$

i.e. $(\rho(g)(f))(v) = g(f(g^{-1}v))$ for all $v \in V_1$. If $G \subseteq Gl_n(k)$ and $V_1 = V_2 = k^n$ is the natural representation, this is just the action by conjugation on square matrices.

Combining the tensor formalism with the abelian category structure, we get a plethora of new representations from a given $V \in \mathrm{Rep}_k(G)$ by taking subquotients of tensor powers

$$(V^* \oplus V)^{\otimes n} \quad \text{for} \quad n \ \in \ \mathbb{N}.$$

We will see this idea at work in the representation theory of symmetric groups later on, for now we restrict ourselves to the most basic example:

EXAMPLE 2.5. Let V be a vector space. For $n \in \mathbb{N}$, the symmetric group $\mathfrak{S}_n$ acts on the tensor powers

$$V^{\otimes n} \ = \ V \otimes \cdots \otimes V$$

by permutation on the factors. We define the corresponding *symmetric power* as the coinvariants

$$\begin{aligned} Sym^n(V) \quad &= \quad (V^{\otimes n})_{\mathfrak{S}_n} \\ &= \quad V^{\otimes n} / \big\langle v_1 \otimes \cdots \otimes v_n - v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)} \mid \sigma \in \mathfrak{S}_n \big\rangle. \end{aligned}$$

Similarly, we define the *exterior power* as the coinvariants for the action twisted by $sgn : \mathfrak{S}_n \to \{\pm 1\}$,

$$\begin{aligned} \wedge^n(V) \quad &= \quad (sgn \otimes V^{\otimes n})_{\mathfrak{S}_n} \\ &= \quad V^{\otimes n} / \big\langle v_1 \otimes \cdots \otimes v_n - sgn(\sigma) \cdot v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)} \mid \sigma \in \mathfrak{S}_n \big\rangle. \end{aligned}$$

So far we have viewed these as representations of the symmetric group. However, if G is any other group and $V \in \mathrm{Rep}_k(G)$, then by transport of structure we may view

$$Sym^n(V), \ \wedge^n(V) \ \in \ \mathrm{Rep}_k(G)$$

as representations of the other group as well. Note that we have deliberately defined the symmetric and exterior powers as quotient spaces, not as subspaces of the tensor powers. This distinction does not play a role if $char(k) = 0$:

EXERCISE 2.6. Let $V \in \mathrm{Rep}_k(G)$ as above. Show that if $char(k) \nmid n!$, then the epimorphisms

$$V^{\otimes n} \ \twoheadrightarrow \ Sym^n(V) \quad \text{and} \quad V^{\otimes n} \ \twoheadrightarrow \ \wedge^n(V)$$

split as the projection onto a direct summand in $\mathrm{Rep}_k(G)$. What happens if the assumption on the characteristic is dropped?

Now that we have obtained so many examples of representations, we should try to get a more structured picture towards their general classification. This is what will take up most of this course starting from the next section.

3. Semisimple representations

One of the main goals of representation theory is to decompose representations into their smallest building blocks, the analogs of elementary particles in physics:

Definition 3.1. A representation $V \in \mathrm{Rep}_k(G)$ is called *irreducible* if $V \neq \{0\}$ and if V has no subrepresentation other than itself and the zero representation.

Example 3.2. (a) Every 1-dimensional representation is irreducible, and if G is a finite abelian group, then these are the only irreducible representations. Conversely, we will see later that any finite non-abelian group G with $char(k) \nmid |G|$ has an irreducible representation of dimension > 1.

(b) If $char(k) \neq 3$, the permutation representation of $G = \mathfrak{S}_3$ on $V = \bigoplus_{i=1}^3 k \cdot e_i$ is a direct sum

$$V \;=\; V^G \oplus W$$

of two irreducible representations

$$\begin{aligned} V^G &= k \cdot (e_1 + e_2 + e_3), \\ W &= k \cdot (e_1 - e_2) \oplus k \cdot (e_2 - e_3). \end{aligned}$$

Indeed, clearly both these spaces are stable under the action of G, and the directness of their sum follows from

$$\begin{aligned} e_1 &= \tfrac{1}{3} \cdot ((e_1 + e_2 + e_3) + (e_1 - e_2) + (e_2 - e_3)), \\ e_2 &= \tfrac{1}{3} \cdot ((e_1 + e_2 + e_3) - (e_1 - e_2) + (e_2 - e_3)), \\ e_3 &= \tfrac{1}{3} \cdot ((e_1 + e_2 + e_3) - (e_1 - e_2) - 2(e_2 - e_3)), \end{aligned}$$

using that $char(k) \neq 3$. By the above the 1-dimensional direct summand V^G is irreducible. To see that also W is irreducible, note that otherwise it would contain a 1-dimensional subrepresentation. But for the subgroup $H = \{id, (12)\} \leq G$ generated by a transposition,

$$W|_H \;=\; k \cdot (e_1 - e_2) \oplus k \cdot (e_1 + e_2 - 2e_3)$$

is the direct sum of the trivial and the sign representation. Since the group action uniquely determines the subspaces underlying these two direct summands of $W|_H$ and since neither of the two is stable under the whole group $G = \mathfrak{S}_3$, it follows that W cannot contain a 1-dimensional subrepresentation.

(c) Irreducibility is not stable under field extensions: For instance the rotation representation

$$\rho: \quad G \;=\; \mathbb{R}/2\pi i\mathbb{Z} \;\longrightarrow\; Gl_2(\mathbb{R}), \quad \varphi \;\mapsto\; \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$$

is irreducible as representation over the real numbers, but its composite with the embedding $Gl_2(\mathbb{R}) \hookrightarrow Gl_2(\mathbb{C})$ is reducible as a complex representation (exercise).

If one wants to understand how more general representations are built up from irreducibles, one has to control morphisms between the latter. Here life is made easy by the following result known as Schur's lemma:

Lemma 3.3 (Schur's lemma I). *For every nonzero morphism $f \in \mathrm{Hom}_G(V, W)$ of representations the following holds:*

1. *If V is irreducible, then f is a monomorphism.*
2. *If W is irreducible, then f is an epimorphism.*

In particular, if both V and W are irreducible, then f is an isomorphism.

Proof. Both $\ker(f) \subseteq V$ and $\mathrm{im}(f) \subseteq W$ are subrepresentations. □

In view of the above lemma, when dealing with morphisms between irreducible representations it suffices to consider endomorphism rings. By the above lemma these are division algebras, so over algebraically closed fields we can put Schur's lemma in the following more precise form:

LEMMA 3.4 (Schur's Lemma II). *If k is algebraically closed and $V \in \mathrm{Rep}_k(G)$ is irreducible, then*

$$\mathrm{End}_G(V) = k \cdot id_V.$$

Proof. Since the base field is algebraically closed, every $f \in \mathrm{End}_G(V)$ has an eigenvalue $\lambda \in k$. Then the G-equivariant map $f - \lambda \cdot id_V$ is not invertible, so by the previous lemma it must be zero. □

EXERCISE 3.5. Compute $\mathrm{End}_G(V)$ for the rotation representation $V = \mathbb{R}^2$ of the circle group $G = \mathbb{R}/2\pi\mathbb{Z}$ from example 3.2(c).

The easiest way to build up more general representations from irreducible ones is to take direct sums. This leads to the notion of semisimple representations that we will focus on during most of this course:

DEFINITION 3.6. A representation $V \in \mathrm{Rep}_k(G)$ is said to be *semisimple* if it decomposes as a direct sum

$$V = \bigoplus_{i \in I} V_i$$

where the $V_i \subseteq V$ are irreducible subrepresentations. These subrepresentations are then called the *irreducible constituents* that occur in the decomposition. If all of these constituents are isomorphic, we say that V is *isotypic*.

It is natural to ask how canonical the above decompositions are, and we will discuss this in a minute. But first we need a few general remarks on the notion of semisimplicity. To begin with, semisimple representations are precisely those in which any G-stable subspace admits a G-stable complement:

LEMMA 3.7. *A representation $V \in \mathrm{Rep}_k(G)$ is semisimple if and only if for every subrepresentation $U_1 \subseteq V$ there is a subrepresentation $U_2 \subseteq V$ with $V = U_1 \oplus U_2$.*

Proof. Suppose first that $V \in \mathrm{Rep}_k(G)$ is semisimple, and let $U_1 \subseteq V$ be any subrepresentation. The subrepresentations

$$U_2 \subseteq V \quad \text{with} \quad U_1 \cap U_2 = \{0\}$$

are partially ordered with respect to inclusion, and as a candidate for the desired complement we take a maximal such subrepresentation U_2. We want to show that then $V = U_1 + U_2$ (in which case the sum will be automatically direct). Now by semisimplicity

$$V = \bigoplus_{i \in I} V_i \quad \text{with irreducible} \quad V_i \in \mathrm{Rep}_k(G).$$

For any $i \in I$ we have

$$(U_1 + U_2) \cap V_i \neq \{0\}$$

since otherwise

$$U_2 \subsetneq U_2 + V_i \quad \text{and} \quad U_1 \cap (U_2 + V_i) = \{0\},$$

which would contradict the maximality of the chosen subrepresentation U_2. So we have

$$\{0\} \neq (U_1 + U_2) \cap V_i \subseteq V_i$$

and since V_i is irreducible, it follows that the inclusion on the right hand side is in fact an equality. But then $U_1 + U_2 \supseteq V_i$ for all $i \in I$, which implies $U_1 + U_2 = V$ and so we have found the desired complement.

Conversely, suppose any subrepresentation of V admits a complement. To see that V is semisimple, pick any irreducible subrepresentation $V_1 \subseteq V$, for example a subrepresentation of smallest possible dimension > 0. By assumption we can find a subrepresentation $V_1' \subseteq V$ with

$$V \;=\; V_1 \oplus V_1',$$

and by induction on the dimension it will suffice to show that inside $V_1' \in \mathrm{Rep}_k(G)$, any subrepresentation $V_2 \subseteq V_1'$ again has a complement. For this we first choose a subrepresentation

$$V_2' \;\subseteq\; V \quad \text{with} \quad V \;=\; V_2 \oplus V_2',$$

which exists by assumption. We claim that

$$V_1' \;=\; V_2 \oplus (V_1' \cap V_2'),$$

which will finish the proof. Indeed, the directness of the sum on the right hand side is clear since

$$V_2 \cap (V_1' \cap V_2') \;\subseteq\; V_2 \cap V_2' \;=\; \{0\}.$$

Furthermore

$$V_1' \;\subseteq\; V \;=\; V_2 \oplus V_2',$$

so any $v \in V_1'$ has the form $v = v_2 + v_2'$ with $v_2 \in V_2$ and $v_2' = v - v_1' \in V_2' \cap V_1'$, and the claim follows. □

COROLLARY 3.8. *The class of all semisimple representations in* $\mathrm{Rep}_k(G)$ *is stable under*

(1) *subrepresentations and quotient representations,*

(2) *sums (not necessarily direct) inside any ambient representation.*

Proof. (1) For subrepresentations of semisimple representations this follows from the second half of the proof of the previous lemma, and the argument for quotient representations is similar (exercise).

(2) For any $V \in \mathrm{Rep}_k(G)$ and any subrepresentations $V_1, V_2 \in \mathrm{Rep}_k(G)$, the sum map

$$V_1 \oplus V_2 \;\twoheadrightarrow\; V_1 + V_2 \;\subseteq\; V$$

is G-equivariant. In particular, its image is again a representation. As a quotient representation of the direct sum $V_1 \oplus V_2 \in \mathrm{Rep}_k(G)$, it is semisimple if V_1 and V_2 are so. □

In general the decomposition of a semisimple representation as a direct sum of irreducibles and the occuring irreducible constituents are only unique up to a non-canonical isomorphism. For example, if V is the trivial representation, there is no distinguished basis of the underlying vector space. But we do have a canonical decomposition into isotypic pieces in the following sense:

DEFINITION 3.9. Let $\mathrm{Irr}_k(G) \subset \mathrm{Rep}_k(G)$ be a system of representatives for the isomorphism classes of irreducible representations. If $V \in \mathrm{Rep}_k(G)$ is semisimple, then for $\rho \in \mathrm{Irr}_k(G)$ we call

$$V_\rho \;=\; \sum_{\substack{U \subseteq V \\ U \simeq \rho}} U \;\subseteq\; V$$

the *ρ-isotypic component* of the representation V. Note that by the above this is a semisimple subrepresentation. The canonical decomposition into isotypic pieces is then given by

Lemma 3.10. *For any semisimple representation $V \in \mathrm{Rep}_k(G)$ there exists a canonical decomposition*

$$V \;=\; \bigoplus_{\rho \in \mathrm{Irr}_k(G)} V_\rho$$

into isotypic components. Furthermore, for each component there exists a canonical epimorphism

$$\mathrm{Hom}_G(\rho, V) \otimes \rho \;\twoheadrightarrow\; V_\rho,$$

and this is an isomorphism if the base field k is algebraically closed.

Proof. By semisimplicity we may find irreducible subrepresentations $V_i \subset V$ such that

$$V \;=\; \bigoplus_{i \in I} V_i,$$

although this decomposition will usually not be canonical. The index set I splits into disjoint subsets

$$I_\rho \;=\; \bigl\{\, i \in I \mid V_i \simeq \rho \,\bigr\} \quad \text{for} \quad \rho \;\in\; \mathrm{Irr}_k(G),$$

and since by Schur's lemma there is no non-trivial morphism between irreducible non-isomorphic representations, it is clear that for each ρ the inclusion $V_\rho \subseteq V$ factors over an inclusion

$$V_\rho \;\subseteq\; \bigoplus_{i \in I_\rho} V_i.$$

In fact this latter inclusion is an equality, since the reverse inclusion holds by the definition of $V_\rho \subseteq V$ as the sum of *all* subrepresentations isomorphic to ρ. Hence the first claim of the lemma follows. Writing each V_ρ non-canonically as a direct sum of copies of $\rho \in \mathrm{Irr}_k(G)$, the second claim follows from the observation that the map

$$\mathrm{Hom}_G(\rho, V) \otimes \rho \;=\; \mathrm{Hom}_G(\rho, V_\rho) \otimes \rho \;\longrightarrow\; V_\rho, \quad f \otimes v \;\mapsto\; f(v)$$

is G-equivariant and surjective by construction. Finally, if k is algebraically closed, then we have $\mathrm{End}_G(\rho) = k$ and the injectivity follows by counting dimensions. □

Note that the above decomposition into isotypic components is functorial, i.e. it is respected by any morphism between semisimple representations. Thus if we denote by

$$m_\rho(V) \;=\; \frac{\dim_k \mathrm{Hom}_G(\rho, V)}{\dim_k \mathrm{End}_G(\rho)}$$

the *multiplicity* of ρ in the semisimple representation $V \in \mathrm{Rep}_k(G)$, then for any other semisimple representation $W \in \mathrm{Rep}_k(G)$ the morphisms $f \in \mathrm{Hom}_G(V, W)$ can be written non-canonically as block diagonal matrices

$$M_f \;=\; \begin{pmatrix} \ddots & & 0 \\ & M_{f,\rho} & \\ 0 & & \ddots \end{pmatrix}$$

whose diagonal blocks

$$M_\rho \;\in\; \Bigl(\mathrm{End}_G(\rho)\Bigr)^{m_\rho(V) \times m_\rho(W)}$$

are $m_\rho(V) \times m_\rho(W)$ matrices with entries in the division algebra $\mathrm{End}_G(\rho)$. In many situations such as for finite groups or compact Lie groups, every representation is semisimple and we then have completely described the category $\mathrm{Rep}_k(G)$ in terms of irreducible representations and their endomorphisms.

4. The Jordan-Hölder theorem

Although for the rest of this course the focus will be on groups for which every representation is semisimple, let us take a brief look at what may happen in the general case.

DEFINITION 4.1. A *composition series* of a representation $V \in \mathrm{Rep}_k(G)$ is a finite series

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

of subrepresentations such that for all $i \in \{1, 2, \dots, n\}$, the quotients V_i/V_{i-1} are irreducible. We then call the quotients the *composition factors* of the series.

EXAMPLE 4.2. (a) If V is semisimple, the composition factors of any composition series are precisely the irreducible constituents. Note that here the composition factors can occur in any order, there is no canonical way to number them.

(b) If $G \subset Gl_n(k)$ is the group of upper triangular matrices, the composition factors of its standard representation on $V = k^n$ are precisely the 1-dimensional representations

$$\rho_i : \quad G \longrightarrow k^\times \quad \text{given by} \quad \begin{pmatrix} a_{11} & \cdots & * \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{pmatrix} \mapsto a_{ii}.$$

These composition factors are ordered in a natural way since none of the successive extensions split: ρ_1 is the unique non-trivial subrepresentation of V, ρ_2 is the unique non-trivial subrepresentation of the quotient V/ρ_1, and so on.

(c) Let $G = \mathfrak{S}_3$ be the symmetric group on three letters and $V = \bigoplus_{i=1}^3 k \cdot e_i$ its natural permutation representation. One may easily check that its only non-trivial subrepresentations are

$$\begin{aligned} V^G &= k \cdot (e_1 + e_2 + e_3), \\ W &= k \cdot (e_1 - e_2) \oplus k \cdot (e_2 - e_3). \end{aligned}$$

Now there are two very different cases:

- If $char(k) \neq 3$, we have seen in the previous section that $V \simeq V^G \oplus W$ is semisimple and the two summands are irreducible.
- If $char(k) = 3$, then $e_1 + e_2 + e_3 = (e_1 - e_2) - (e_2 - e_3) \in W$ and we get a composition series

  $$\{0\} \subset V^G \subset W \subset V$$

  with 1-dimensional composition factors, none of them a direct summand.

It turns out that in general, composition series always exist and the composition factors are uniquely determined:

THEOREM 4.3. *Every finite-dimensional representation $V \in \mathrm{Rep}_k(G)$ admits a composition series. Any two such composition series have the same length n, and up to isomorphism and reordering they share the same composition factors.*

Proof. To see the existence of a composition series, let $V_1 \subset V$ be any non-zero irreducible subrepresentation, for instance a non-zero subrepresentation of smallest possible dimension. Then replace V by the quotient V/V_1 and proceed inductively, using that the reduction map

$$\{\text{subrepresentations of } V \text{ containing } V_1\} \simeq \{\text{subrepresentations of } V/V_1\}$$

is a bijection. For the uniqueness part, suppose that we are given two different composition series:

$$\begin{aligned} 0 &= V_0 \subset V_1 \subset \cdots \subset V_n = V \\ 0 &= U_0 \subset U_1 \subset \cdots \subset U_m = V \end{aligned}$$

Putting

$$\begin{aligned} V_{j,i} &= V_{j-1} + V_j \cap U_i \quad \text{for} \quad i = 1, \dots, m, \\ U_{i,j} &= U_{i-1} + U_i \cap V_j \quad \text{for} \quad j = 1, \dots, n, \end{aligned}$$

we obtain

$$\begin{aligned} \cdots \subset V_{j-1} = V_{j,0} \subset V_{j,1} \subset \cdots \subset V_{j,m} = V_j \subset \cdots \\ \cdots \subset U_{i-1} = U_{i,0} \subset U_{i,1} \subset \cdots \subset U_{i,n} = U_i \subset \cdots \end{aligned}$$

as a seeming refinement of our composition series. But since V_j/V_{j-1} and U_i/U_{i-1} are irreducible, these refinements have just inserted extra copies of the same terms in our original series! In other words:

- For each $j \in \{1, \dots, m\}$ there exists a unique index $I(j) \in \{1, \dots, n\}$ such that
$$V_{j,i}/V_{j,i-1} \simeq \begin{cases} V_j/V_{j-1} & \text{if } i = I(j), \\ 0 & \text{otherwise.} \end{cases}$$
- For each $i \in \{1, \dots, n\}$ there exists a unique index $J(i) \in \{1, \dots, m\}$ such that
$$U_{i,j}/U_{i,j-1} \simeq \begin{cases} U_i/U_{i-1} & \text{if } i = I(j), \\ 0 & \text{otherwise.} \end{cases}$$

We claim that

$$V_{j,i}/V_{j,i-1} \simeq U_{i,j}/U_{i,j-1}$$

holds for all i, j. Note that this will finish the proof: Indeed, the claim in particular implies that

$$\{(I(j), j)\} = \{(i,j) \mid V_{j,i}/V_{j,i-1} \neq 0\} \stackrel{!}{=} \{(i,j) \mid U_{i,j}/V_{i,j-1} \neq 0\} = \{(i, J(i))\},$$

so the assignments $i \mapsto J(i)$ and $j \mapsto I(j)$ are mutually inverse, $m = n$ and the composition factors of the two composition series are the same up to reordering and isomorphism.

It remains to prove the above claim. Going back to the definitions, we want an isomorphism

$$(U_{i-1} + U_i \cap V_j)/(U_{i-1} + U_i \cap V_{j-1}) \simeq (V_{j-1} + V_j \cap U_i)/(V_{j-1} + V_j \cap U_{i-1}).$$

To get such an isomorphism, it suffices to note that the inclusion of $U_i \cap V_j$ in the numerator of both sides induces epimorphisms

$$\begin{aligned} p: \quad U_i \cap V_j &\twoheadrightarrow (U_{i-1} + U_i \cap V_j)/(U_{i-1} + U_i \cap V_{j-1}) \\ q: \quad U_i \cap V_j &\twoheadrightarrow (U_{i-1} + U_i \cap V_j)/(U_{i-1} + U_i \cap V_{j-1}) \end{aligned}$$

with $\ker(p) = U_i \cap V_{j-1} + U_{i-1} \cap V_j = \ker(q)$. □

5. Induction and restriction

It is often useful to study representations of a given group via the action of simpler subgroups. More generally, for any homomorphism $\varphi : H \to G$ of groups we have a restriction functor

$$\mathrm{Res}_H^G : \quad \mathrm{Rep}_k(G) \longrightarrow \mathrm{Rep}_k(H), \quad (V,\rho) \mapsto (V, \rho\circ\varphi).$$

What about the converse direction? In general of course we cannot expect the above restriction functor to be an equivalence, but it turns out that in many situations it admits an adjoint functor. We can either look for *left* or for *right* adjoints, and this duality is best seen in terms of modules over group algebras. To this end we make the following general observation:

Definition 5.1. The *group algebra* of G is the (possibly infinite-dimensional) vector space

$$k[G] \;=\; \bigoplus_{g\in G} ke_g$$

with the algebra structure given on formal basis vectors by $e_g \cdot e_h = e_{gh}$. If V is a finite-dimensional vector space over k, then we have a bijective correspondence between

- representations $\rho : G \to Gl(V)$ and
- left module structures $m : k[G] \to \mathrm{End}_k(V)$

given by

$$m\Big(\sum_{g\in G} a_g e_g\Big)(v) \;=\; \sum_{g\in G} a_g \rho(g)(v) \quad \text{resp.} \quad \rho(g)(v) \;=\; m(e_g)(v).$$

Now let $\varphi : H \to G$ be a group homomorphism such that $\varphi(H) \le G$ is a subgroup of finite index; the last condition is needed to get finite-dimensional representations but can be dropped if one allows infinite-dimensional ones. For $V \in \mathrm{Rep}_k(H)$ we define the *induced representation* as the left module

$$\mathrm{Ind}_H^G(V) \;=\; k[G] \otimes_{k[H]} V$$

under $k[G]$, so that by construction we get the following left adjointness:

Theorem 5.2 (Frobenius reciprocity I). *The induction as defined above gives a functor*

$$\mathrm{Ind}_H^G : \quad \mathrm{Rep}_k(H) \longrightarrow \mathrm{Rep}_k(G)$$

which is left adjoint to the corresponding restriction functor, i.e. we have natural isomorphisms

$$\mathrm{Hom}_G(\mathrm{Ind}_H^G(V), W) \;\simeq\; \mathrm{Hom}_H(V, \mathrm{Res}_H^G(W)) \quad \textit{for} \quad V \in \mathrm{Rep}_k(H),\; W \in \mathrm{Rep}_k(G).$$

Proof. For any finite-dimensional representation, the induced representation is again finite-dimensional because $k[G]$ is a finitely generated right module over $k[H]$ if $[G : \varphi(H)] < \infty$. The functoriality of the induced representation and the left adjointness property is clear since any ring homomorphism $\varphi : R \to S$ induces a scalar extension functor $S \otimes_R (-) : \mathrm{Mod}(R) \to \mathrm{Mod}(S)$ between the corresponding categories of left modules, with natural isomorphisms

$$\mathrm{Hom}_S(S \otimes_R V, W) \;\simeq\; \mathrm{Hom}_R(V, W)$$

for $V \in \mathrm{Mod}(R)$ and $W \in \mathrm{Mod}(S)$. □

Notice that if one allows infinite-dimensional representations, then the above result remains valid for arbitrary homomorphisms φ without any changes. There is also a right adjoint to the restriction functor, but this is more naturally defined in the following dual way and differs from the left adjoint if the finite index assumption is dropped:

DEFINITION 5.3. Let $\varphi : H \to G$ be a homomorphism with $[G : \varphi(H)] < \infty$. For $V \in \mathrm{Rep}_k(H)$ we define the *coinduced representation* by

$$\begin{aligned}\mathrm{coInd}_H^G(V) \;&=\; \mathrm{Hom}_{k[H]}(k[G], V)\\ &\simeq\; \Big\{ f : G \to V \;\big|\; h \cdot f(x) = f(\varphi(h)x)\; \forall\, h \in H, x \in G \Big\} \;\in\; \mathrm{Rep}_k(G),\end{aligned}$$

where the latter is equipped with the group action

$$(gf)(x) \;=\; f(xg) \quad \text{for} \quad f \;\in\; \mathrm{coInd}_H^G(V) \quad \text{and} \quad g, x \;\in\; G.$$

The last definition is the same as for the right regular representation, and since right and left multiplication commute with each other, it is clear that $gf \in \mathrm{coInd}_H^G(V)$ for all $f \in \mathrm{coInd}_H^G(V)$ and $g \in G$. So the above indeed defines a group action on the induced representation. We also note that any $f \in \mathrm{coInd}_H^G(V)$ is determined by its values on a set of representatives for $G/\varphi(H)$, so the dimension of the coinduced representation is finite. More precisely:

LEMMA 5.4. *In the above setting,*

$$\dim_k(\mathrm{coInd}_H^G(V)) \;=\; [G : \varphi(H)] \cdot \dim_k(V^{\ker(\varphi)}).$$

Proof. For $f \in \mathrm{coInd}_H^G(V)$, the definition of the coinduced representation implies that any $h \in \ker(\varphi)$ acts trivially on $f(x)$ for all $x \in G$. In other words, by construction

$$f : \quad G \longrightarrow V^{\ker(\varphi)} \;\subset\; V$$

takes values in the invariants under $\ker(\varphi)$. Furthermore, if we pick any system of representatives $g_1, \dots, g_n \in G$ for the finitely many left cosets $\varphi(H)\backslash G$ so that G is a disjoint union

$$G \;=\; \bigsqcup_{i=1}^{n} \varphi(H) \cdot g_i,$$

then it follows from the definition that any $f \in \mathrm{coInd}_H^G(V)$ is determined uniquely by the values

$$f(g_i) \;\in\; V^{\ker(\varphi)} \quad \text{for} \quad i \;=\; 1, 2, \dots, n.$$

Conversely, one easily checks that these values may be chosen arbitrarily. □

Notice that $V^{\ker(\varphi)} \in \mathrm{Rep}_k(H)$ because $\ker(\varphi) \trianglelefteq H$ is a *normal* subgroup. The above in fact shows

$$\mathrm{coInd}_H^G(V) \;=\; \mathrm{coInd}_{\mathrm{im}(\varphi)}^G(V^{\ker(\varphi)})$$

when $V^{\ker(\varphi)}$ is viewed as a representation of the group $H/\ker(\varphi) \simeq \mathrm{im}(\varphi)$. So in principle it suffices to consider (co-)induced representations in the special case of subgroups, which is also the most interesting one for applications.

EXAMPLE 5.5. For a finite group G, the regular representations on $k[G]$ can be obtained by induction from the trivial 1-dimensional representation of the trivial subgroup:

$$\big(k[G], \lambda\big) \;\simeq\; \mathrm{Ind}_{\{1\}}^G(\mathbf{1}) \quad \text{and} \quad \big(k[G], \rho\big) \;\simeq\; \mathrm{coInd}_{\{1\}}^G(\mathbf{1}).$$

We know from the exercises that the right and the left regular representations are isomorphic, so here the induction and coinduction are essentially the same. In fact we will see that for $[G : \varphi(H)] < \infty$ this is always the case; but first let us check that the coinduction is indeed right adjoint to the restriction functor:

THEOREM 5.6 (Frobenius reciprocity II). *Let $\varphi : H \to G$ be a homomorphism whose image is of finite index. Then the coinduction is a functor*

$$\mathrm{coInd}_H^G : \quad \mathrm{Rep}_k(H) \longrightarrow \mathrm{Rep}_k(G)$$

that is right adjoint to the corresponding restriction functor, i.e. we have natural isomorphisms

$$\mathrm{Hom}_G(U, \mathrm{coInd}_H^G(V)) \simeq \mathrm{Hom}_H(\mathrm{Res}_H^G(U), V) \quad \textit{for} \quad U \in \mathrm{Rep}_k(G),\ V \in \mathrm{Rep}_k(H).$$

Proof. The functoriality is clear if one writes $\mathrm{coInd}_H^G(-) = \mathrm{Hom}_{k[H]}(k[G], -)$, but for the sake of completeness we check by hand that everything is compatible with the group actions: Let $V, W \in \mathrm{Rep}_k(H)$. Then for $\Phi \in \mathrm{Hom}_H(V, W)$ we get a map

$$\mathrm{coInd}_H^G(\Phi) : \quad \mathrm{coInd}_H^G(V) \longrightarrow \mathrm{coInd}_H^G(W), \quad f \mapsto \Phi \circ f$$

since

$$\begin{aligned} h \cdot (\Phi \circ f)(x) &= h \cdot \Phi(f(x)) && \text{(by composition)} \\ &= \Phi(h \cdot f(x)) && \text{(as } \Phi \text{ is } H\text{-equivariant)} \\ &= \Phi(f(\varphi(h)x)) && \text{(since } f \in \mathrm{Ind}_H^G(V)) \\ &= (\Phi \circ f)(\varphi(h)x) && \text{(by composition)} \end{aligned}$$

for all $f \in \mathrm{Ind}_H^G(V)$, $h \in H$ and $x \in G$. These maps satisfy

$$\mathrm{Ind}_H^G(id) = id \quad \text{and} \quad \mathrm{Ind}_H^G(\Psi) \circ \mathrm{Ind}_H^G(\Phi) = \mathrm{Ind}_H^G(\Psi \circ \Phi)$$

when Φ, Ψ are any two composable morphisms, so $\mathrm{Ind}_H^G(-)$ is a functor.

To see that it is right adjoint to the restriction functor, take any $U \in \mathrm{Rep}_k(G)$ and $V \in \mathrm{Rep}_k(H)$. We then define maps

$$\mathrm{Hom}_G(U, \mathrm{Ind}_H^G(V)) \underset{\Psi \leftarrow\!\shortmid \Psi^\sharp}{\overset{\Phi \mapsto \Phi^\flat}{\rightleftarrows}} \mathrm{Hom}_H(\mathrm{Res}_H^G(U), V)$$

by

$$\begin{aligned} \Phi^\flat : &\quad \mathrm{Res}_H^G(U) \longrightarrow V, \quad u \mapsto \Phi(u)(1) && \text{for} \quad \Phi \in \mathrm{Hom}_G(U, \mathrm{Ind}_H^G(V)), \\ \Psi^\sharp : &\quad U \longrightarrow \mathrm{Ind}_H^G(V), \quad u \mapsto (x \mapsto \Psi(x \cdot u)) && \text{for} \quad \Psi \in \mathrm{Hom}_H(\mathrm{Res}_H^G(U), V). \end{aligned}$$

Note that

$$\begin{aligned} \Phi^\flat \in \mathrm{Hom}_H(\mathrm{Res}_H^G(U), V) &\quad \text{since} \quad \Phi^\flat(h \cdot u) = \Phi(h \cdot u)(1) = h \cdot \Phi(u)(1), \\ \Psi^\sharp \in \mathrm{Hom}_G(U, \mathrm{Ind}_H^G(V)) &\quad \text{since} \quad \Psi^\sharp(g \cdot u) = (x \mapsto \Psi(x \cdot g \cdot u)) = g \cdot \Psi^\sharp(u) \end{aligned}$$

for $u \in U$, $h \in H$ and $x, g \in G$. Finally, the assignments $\Phi \mapsto \Phi^\flat$ and $\Psi \mapsto \Psi^\sharp$ are easily checked to be inverse to each other. □

REMARK 5.7. Under our usual assumption $[G : \varphi(H)] < \infty$ one has a natural isomorphism

$$\mathrm{Ind}_H^G(V) \simeq \mathrm{coInd}_H^G(V) \quad \text{for} \quad V \in \mathrm{Rep}_k(H),$$

as one may check with the arguments of the following lemma (exercise). So for finite index subgroups the choice between the induction functor is both a left and right adjoint of the restriction functor, which is sometimes also expressed by saying that the functors Res_H^G and Ind_H^G form a *Frobenius pair*. However, in the setting of infinite-dimensional representations and arbitrary subgroups, the left and right adjoints of the restriction functor are not the same in general.

In the study of representations, it is often important to know whether a given representation arises by induction from a subgroup. The following lemma provides an easy criterion:

LEMMA 5.8. *Let $V \in \mathrm{Rep}_k(G)$, and suppose that there exists a subgroup $H \leq G$ of finite index and a subrepresentation $U \subseteq \mathrm{Res}^G_H(V)$ such that the underlying vector spaces satisfy*

$$V \;=\; \bigoplus_{i=1}^{n} g_i^{-1} U$$

where the $g_i \in G$ are representatives for the right cosets in $G = \bigsqcup_{i=1}^n Hg_i$. Then we have

$$V \;\simeq\; \mathrm{Ind}^G_H(U).$$

Proof. We prove the lemma for the coinduced representation $\mathrm{coInd}^G_H(U)$, but taking $V = \mathrm{Ind}^G_H(U)$ we obtain *a posteriori* that coinduction and induction are the same. So consider the linear map

$$F: \quad \mathrm{coInd}^G_H(U) \;\longrightarrow\; V \quad \text{defined by} \quad F(f) \;=\; \sum_{i=1}^{n} g_i^{-1} \cdot f(g_i).$$

We first claim that this map is injective: Suppose that $f \in \ker(F)$. Since $g_i^{-1} \cdot f(g_i)$ lies in the subspace $g_i^{-1}U$ and since by assumption the sum of these subspaces is direct, it is clear that we must have $f(g_i) = 0$ for all i. By definition of $\mathrm{coInd}^G_H(U)$ it follows that

$$f(hg_i) \;=\; h \cdot f(g_i) \;=\; 0 \quad \text{for all } i \text{ and all } h \in H.$$

But then

$$f \;=\; 0 \quad \text{because} \quad G \;=\; \sqcup_{i=1}^n Hg_i.$$

So F is injective, hence an isomorphism since the source and target have the same dimension:

$$\dim_k \mathrm{Ind}^G_H(U) \;=\; [G:H] \cdot \dim_k U \;=\; \dim_k \oplus_{i=1}^n g_i^{-1} U \;=\; \dim_k V$$

It then only remains to check that F is G-equivariant.

Let $g \in G$. Since the g_i form a set of representatives for the right cosets of H, there exists a unique permutation $\sigma \in \mathfrak{S}_n$ of the index set such that for each index i we have

$$g_i g \;=\; h_i g_{\sigma(i)} \quad \text{for some unique} \quad h_i \;\in\; H.$$

For $f \in \mathrm{coInd}^G_H(U)$ we then get

$$\begin{aligned} F(g \cdot f) \;=\; \sum_{i=1}^n g_i^{-1} \cdot f(g_i g) \;&=\; \sum_{i=1}^n g_i^{-1} \cdot f(h_i g_{\sigma(i)}) \\ &=\; \sum_{i=1}^n g_i^{-1} h_i \cdot f(g_{\sigma(i)}) \\ &=\; \sum_{i=1}^n g g_{\sigma(i)}^{-1} \cdot f(g_{\sigma(i)}) \;=\; g \cdot F(f) \end{aligned}$$

and hence the G-equivariance of F follows. □

A typical application is the study of representations via their restriction to finite index normal subgroups, usually referred to as Clifford theory:

COROLLARY 5.9. *Let $V \in \mathrm{Irr}_k(G)$ be an irreducible representation and $N \trianglelefteq G$ a normal subgroup of finite index. Then $\mathrm{Res}^G_N(V)$ is semisimple, and one of the following holds:*

(1) *either the restriction $\mathrm{Res}^G_N(V)$ is isotypic,*

(2) *or there exists an intermediate group $N \hookrightarrow H \overset{\neq}{\hookrightarrow} G$ and some $U \in \mathrm{Irr}_k(H)$ such that*

$$V \simeq \mathrm{Ind}^G_H(U).$$

Proof. We first remark that for any normal subgroup $N \trianglelefteq G$ the action of G permutes the subrepresentations $W \subseteq \mathrm{Res}^G_N(V)$: For any such subrepresentation and $g \in G$ the image gW is again stable under the action of the subgroup $N \trianglelefteq G$ since

$$ngW = gg^{-1}ngW = gn^gW \subseteq gW \quad \text{for} \quad n \in N \quad \text{and} \quad n^g = g^{-1}ng \in N.$$

Clearly this action on subrepresentations preserves irreducibility. Hence if we define the socle

$$S \subseteq \mathrm{Res}^G_N(V)$$

to be the sum of all irreducible subrepresentations the restriction $\mathrm{Res}^G_N(V)$, then S is stable under G. Since V is irreducible, it follows that $S = V$ and hence $\mathrm{Res}^G_N(V)$ is a semisimple representation, being a sum of irreducible ones.

This proves the first claim. For the second claim, by semisimplicity we have a canonical decomposition

$$\mathrm{Res}^G_N(V) = \bigoplus_{\rho \in \mathrm{Irr}_k(N)} S_\rho$$

into isotypic components $S_\rho \in \mathrm{Rep}_k(N)$. The same irreducibility argument as before shows that G permutes the non-zero isotypic components transitively. If there is only a single such component, then case (1) occurs and we are done. If there are several of them, we arbitrarily fix one isotypic component $U = S_{\rho_0}$ and consider the subgroup

$$H = \{ g \in G \mid gU \subseteq U \} \overset{\neq}{\hookrightarrow} G.$$

stabilizing the underlying subspace. Notice that by construction $N \subseteq H$. We may naturally view

$$U = S_{\rho_0} \in \mathrm{Rep}_k(H) \quad \text{as a subrepresentation of} \quad \mathrm{Res}^G_H(V) = \bigoplus_\rho S_\rho,$$

and one easily checks that the previous lemma applies. □

CHAPTER II

Character theory of finite groups

In this chapter we consider in more detail the case of a *finite* group G. We always assume that the characteristic of the base field k does not divide the group order, i.e. $char(k) \nmid |G|$. In this case it turns out that every representation $V \in \mathrm{Rep}_k(G)$ is semisimple and the decomposition into irreducible constituents is governed by a powerful tool called character theory.

1. Maschke's theorem

Our assumption on the characteristic of the base field implies that $\mathrm{Rep}_k(G)$ is a semisimple category:

THEOREM 1.1 (Maschke). *If $char(k) \nmid |G|$, every $V \in \mathrm{Rep}_k(G)$ is semisimple.*

Proof. We must show that any subrepresentation $U \subseteq V$ splits off as a direct summand in $\mathrm{Rep}_k(G)$. So for any such subrepresentation we want a G-equivariant projection

$$p \ \in \ \mathrm{End}_G(V) \quad \text{with} \quad \mathrm{im}(p) \ = \ U \quad \text{and} \quad p|_U \ = \ id.$$

To get such an equivariant projection, we use an averaging trick: Pick any k-linear projection $\tilde{p} \in \mathrm{End}_k(V)$ with $\mathrm{im}(\tilde{p}) = U$ and $\tilde{p}|_U = id$, and consider the sum over all its conjugates,

$$p \ \in \ \mathrm{End}_k(V) \quad \text{defined by} \quad p(v) \ = \ \frac{1}{|G|} \sum_{g \in G} g \cdot \tilde{p}(g^{-1}v).$$

Then one easily checks that all the requirements are met. □

REMARK 1.2. The converse to Maschke's theorem also holds: If $char(k) \mid |G|$, there exists a representation $V \in \mathrm{Rep}_k(G)$ which is not semisimple. Here one may take $V = k[G]$ to be the regular representation: Its invariants are 1-dimensional and spanned by the sum $e = \sum_{g \in G} e_g$ of all basis vectors; but for $char(k) \mid |G|$ we have an inclusion

$$V^G \ = \ k \cdot e \ \subseteq \ W \ = \ \Big\{ \sum_{g \in G} a_g e_g \in V \mid \sum_{g \in G} a_g = 0 \Big\}$$

and V/W is also a 1-dimensional trivial representation. So the trivial representation occurs twice as a composition factor, but only once as a subrepresentation.

The theory of representations over fields whose characteristic divides the group order is also called *modular representation theory*. The failure of semisimplicity makes it an interesting but rather hard subject, so from now on we will always assume that $char(k) \nmid |G|$. In this case the representation theory is much nicer but still has many nontrivial applications as we will see later. For the most basic example, the regular representation, the decomposition into irreducible constituents takes the following form:

PROPOSITION 1.3. *If k is algebraically closed, then the regular representation splits as*

$$k[G] \quad \simeq \bigoplus_{\rho \in \mathrm{Irr}_k(G)} \rho^{\oplus d_\rho} \quad \textit{where} \quad d_\rho \;=\; \dim_k \rho \;\geq\; 1.$$

Proof. To see how $V = k[G]$ decomposes, it suffices by Maschke's theorem to compute the multiplicities

$$\begin{aligned} m_\rho(V) \;&=\; \dim_k \mathrm{Hom}_G(\rho, V) \\ &=\; \dim_k \mathrm{Hom}_G(\rho, \mathrm{Ind}_{\{1\}}^G(\mathbf{1})) \\ &=\; \dim_k \mathrm{Hom}_k(\mathrm{Res}_{\{1\}}^G(\rho), \mathbf{1}) \;=\; \dim_k \rho \end{aligned}$$

for $\rho \in \mathrm{Irr}_k(G)$, using Frobenius reciprocity and the observation that the regular representation is induced from the 1-dimensional trivial representation $\mathbf{1}$. □

Note that this provides a way to determine all possible representations, as *every* irreducible representation $\rho \in \mathrm{Irr}_k(G)$ enters in the regular representation with positive multiplicity! We also record the following important numerical

COROLLARY 1.4. *If k is algebraically closed, then $|G| = \sum_{\rho \in \mathrm{Irr}_k(G)} (\dim_k \rho)^2$.*

As a toy application, this formula allows for a slick proof of the following result without using the Jordan canonical form:

COROLLARY 1.5. *Let k be algebraically closed. Then the finite group G is abelian iff all irreducible representations $\rho \in \mathrm{Irr}_k(G)$ are 1-dimensional.*

Proof. For a finite abelian group, the 1-dimensional representations again form a finite abelian group

$$\hat{G} \;=\; \mathrm{Hom}(G, k^\times)$$

of the same order. On the other hand, the group orders are related by the previous corollary:

$$|G| \;\;=\;\; |\hat{G}| \;\;+ \sum_{\substack{\rho \in \mathrm{Irr}(G) \\ \dim_k \rho > 1}} (\dim_k \rho)^2$$

Hence the second summand on the right hand side must be empty, which means that every irreducible representation is 1-dimensional. Conversely, if this latter property holds, then the regular representation is a direct sum of 1-dimensional representations and it follows by faithfulness that G is abelian. □

For very small groups, the above results about the regular representation are already enough to determine explicitly all irreducible representations. We here restrict ourselves to a very simple example:

EXAMPLE 1.6. Over any algebraically closed field k with $char(k) \neq 2, 3$, the only irreducible representations of $G = \mathfrak{S}_3$ are

- the trivial representation $\mathbf{1}$,
- the sign representation $sgn : \mathfrak{S}_3 \to \{\pm 1\}$,
- the natural representation on $W = \{(x, y, z) \in k^3 \mid x + y + z = 0\}$.

Indeed, we know that all these are irreducible and pairwise non-isomorphic, and the squares of their dimensions sum up to $1^2 + 1^2 + 2^2 = 6 = |\mathfrak{S}_3|$. Note that since any other representation is a direct sum of copies of the above three ones, we get identities such as

$$W^* \simeq W, \quad W \otimes sgn \simeq W, \quad W \otimes W \simeq W \otimes W^* \simeq \mathbf{1} \oplus sgn \oplus W \quad \text{etc.}$$

Can we do such things more systematically? And is there a direct way to determine how many irreducible representations there are for a given group? To answer these questions, we need to gain a better understanding of what Maschke's theorem really says about the structure of the group algebra.

2. The structure of semisimple algebras

Although we have proved Maschke's theorem by a simple averaging argument in a few lines, the appropriate framework to understand its implications is the general structure theory of semisimple algebras. Let A be a finite-dimensional algebra[1] over a field k, and denote by

$$\mathrm{Mod}(A)$$

the category of its left modules that are finite-dimensional over k. If $A = k[G]$ is the group algebra of a finite group, we recover the category of representations from above, but in fact all notions from the previous chapter carry over to arbitrary finite-dimensional algebras:

Remark 2.1. (a) A module $V \in \mathrm{Mod}(A)$ is called *irreducible* if it is non-zero and has no submodules other than zero and itself. With the same proof as in Schur's lemma, one sees that any non-zero morphism with irreducible source resp. target is a mono- resp. epimorphism. In particular, if k is algebraically closed, it follows that

$$\mathrm{End}_A(V) \;=\; k \cdot id_V \quad \text{for all irreducible} \quad V \;\in\; \mathrm{Mod}(A).$$

(b) A module $V \in \mathrm{Mod}(A)$ is called *semisimple* if it is a direct sum of irreducible submodules. As in the previous chapter this is the case iff every submodule of V splits off as a direct summand, and the class of semisimple modules is stable under arbitrary sums, submodules and quotient modules.

(c) The algebra A is called *semisimple* if it is so when considered as a left module over itself. For example, applying Maschke's theorem to the regular representation we get that for $char(k) \nmid |G|$ the group algebra $A = k[G]$ is semisimple.

For representations of finite groups, we have seen that the regular representation plays an important role since it contains all the irreducible representations. Again this generalizes to modules over arbitrary algebras, and if we consider quotients rather than submodules, we do not even need semisimplicity:

Lemma 2.2. *Any irreducible* $V \in \mathrm{Mod}(A)$ *is isomorphic to a quotient of* A.

Proof. Let $V \in \mathrm{Mod}(A)$ be irreducible. For any non-zero vector $v \in V \setminus \{0\}$, the scalar multiplication

$$m_v : \quad A \longrightarrow V \quad \text{given by} \quad a \mapsto a \cdot v$$

is a homomorphism of left modules, and it is non-zero because $m_v(1) = v \neq 0$. So by Schur's lemma it must be an epimorphism. □

In particular, there are only finitely many isomorphism classes of irreducible modules in $\mathrm{Mod}(A)$. We fix a representative set $\mathrm{Irr}(A)$ for these; then as in the previous chapter any semisimple $V \in \mathrm{Mod}(A)$ decomposes as the direct sum of its isotypic pieces

$$V \;=\; \bigoplus_{\rho \in \mathrm{Irr}(A)} V_\rho \quad \text{where} \quad V_\rho \;\simeq\; \rho^{\oplus m_\rho} \quad \text{for certain} \quad m_\rho \;\in\; \mathbb{N}_0.$$

If A is semisimple, we may apply this to $V = A$ considered as a left module over itself and obtain the following

[1] By an algebra we always mean an associative algebra A with a unit element $1 \in A$.

LEMMA 2.3. *Any semisimple algebra A, considered as a left module over itself, has the decomposition*

$$A = \bigoplus_{\rho \in \mathrm{Irr}(A)} A_\rho \quad \textit{where} \quad A_\rho \simeq \rho^{\oplus m_\rho} \quad \textit{with} \quad m_\rho = \frac{\dim_k \rho}{\dim_k \mathrm{End}_A(\rho)} \geq 1.$$

Proof. The argument is the same as for the left regular representation of a finite group as $\dim_k \mathrm{Hom}_A(A_\rho, \rho) = \dim_k \mathrm{Hom}_A(A, \rho) = \dim_k \mathrm{Hom}_k(k, \rho) = \dim_k \rho$. □

In fact it turns out that the above decomposition into isotypic components is not just a decomposition as a left module but as an algebra, and as such it is as fine as possible. To formulate the result concisely we recall the following

DEFINITION 2.4. A *two-sided ideal* of A is an additive subgroup $\mathfrak{a} \trianglelefteq A$ which is stable under the multiplication by scalars from both left and right in the sense that $ax \in \mathfrak{a}$ and $xa \in \mathfrak{a}$ for all $a \in A$, $x \in \mathfrak{a}$. Note that this property is the same as being a submodule for both the left and right module structures. An algebra A is called *simple* if it has no two-sided ideals different from zero and itself.

In these terms, the decomposition into isotypic components results in the main structure theorem for semisimple algebras:

THEOREM 2.5 (Weddeburn). *Any semisimple algebra A is a product of simple subalgebras. Explicitly, in the decomposition*

$$A = \bigoplus_{\rho \in \mathrm{Irr}(A)} A_\rho$$

each $A_\rho \trianglelefteq A$ is a simple subalgebra and we have $A_\rho \cdot A_\sigma = \{0\}$ for all $\sigma \neq \rho$.

Proof. To start with, note that by definition each isotypic component $A_\rho \subseteq A$ is a left submodule. It is also a right submodule: For any $a \in A$ the right multiplication map $x \mapsto xa$ is an endomorphism of A as a left module, hence by naturality of the isotypic decomposition it preserves the isotypic components. So each $A_\rho \trianglelefteq A$ is a two-sided ideal. Hence

$$A_\rho \cdot A_\sigma \subseteq A_\rho \cap A_\sigma = \{0\} \quad \text{for all} \quad \sigma \neq \rho,$$

because any two different isotypic components intersect trivially. In other words, it follows that the isotypic decomposition is in fact a decomposition as a product of subalgebras. It remains to see that these subalgebras $A_\rho \trianglelefteq A$ are simple. Suppose we have a two-sided ideal

$$\mathfrak{a} \trianglelefteq A_\rho \quad \text{with} \quad \mathfrak{a} \neq A_\rho.$$

Since the ideal is not equal to the whole isotypic component, by semisimplicity we can find an irreducible submodule

$$V \subseteq A_\rho \quad \text{with} \quad V \not\hookrightarrow \mathfrak{a}.$$

Then $\mathfrak{a} \cap V = \{0\}$ by irreducibility, and using the right module property of $\mathfrak{a}$ we get as above

$$\mathfrak{a} \cdot V \subseteq \mathfrak{a} \cap V = \{0\}.$$

Since $V \simeq \rho$, it follows that

$$\mathfrak{a} \subseteq \mathrm{Ann}(\rho) = \{a \in A \mid a \mapsto 0 \in \mathrm{End}_A(\rho)\}.$$

But by construction the isotypic component A_ρ is isomorphic to a direct sum of copies of ρ, hence we get that

$$\mathfrak{a} \subseteq \mathrm{Ann}(A_\rho)$$

On the other hand also

$$\mathfrak{a} \;\subseteq\; A_\rho \;\subseteq\; \mathrm{Ann}(A_\sigma) \quad \text{for all} \quad \sigma \;\neq\; \rho$$

by what we have seen earlier. It then follows that $\mathfrak{a} \subseteq \mathrm{Ann}(A) = \{0\}$. □

For a complete structure theorem it now only remains to describe the occuring simple subalgebras A_ρ. By Schur's lemma this is particularly easy over algebraically closed fields, in general we must consider the division algebras $D = \mathrm{End}_A(\rho)$:

Corollary 2.6. *If A is semisimple, then any irreducible module $\rho \in \mathrm{Irr}(A)$ satisfies*

$$A_\rho \;=\; \bigcap_{\sigma\neq\rho} \mathrm{Ann}(\sigma) \;\simeq\; \mathrm{End}_D(\rho),$$

where $D = \mathrm{End}_A(\rho)$ and

$$\mathrm{End}_D(\rho) \;=\; \{f \in \mathrm{End}_k(\rho) \mid f\circ g = g\circ f \text{ for all } g \in D\} \;\simeq\; \mathrm{Mat}_{m_\rho\times m_\rho}(D^{op}).$$

Proof. Since A_ρ is a simple algebra, the structure map $A \to \mathrm{End}_k(\rho)$ restricts to an embedding

$$\begin{array}{ccc} A_\rho & \hookrightarrow & \mathrm{End}_k(\rho) \\ & \searrow & \nearrow \\ & \mathrm{End}_D(\rho) & \end{array}$$

and by definition the image is contained in the subalgebra $\mathrm{End}_D(\rho) \subseteq \mathrm{End}_k(\rho)$ of endomorphisms commuting with $D = \mathrm{End}_A(\rho)$. But for the dimensions one computes

$$\begin{aligned} \dim_k A_\rho \;&=\; m_\rho \cdot \dim_k \rho \\ &=\; m_\rho^2 \cdot \dim_k D \\ &=\; \dim_k \mathrm{End}_D(\rho)\cdot \dim_k D \;=\; \dim_k \mathrm{End}_D(\rho) \end{aligned}$$

by lemma 2.3 and the remark below. □

Remark 2.7. If D is a division algebra, then the endomorphisms of D as a left module over itself again form a division algebra $\mathrm{End}_D(D)$. For the latter we have an isomorphism

$$D^{op} \;\simeq\; \mathrm{End}_D(D), \quad a \;\mapsto\; (x \mapsto xa)$$

with the opposite algebra

$$D^{op} \;=\; (D, +, \cdot_{op}) \quad \text{whose multiplication is defined by} \quad x \cdot_{op} y \;=\; y \cdot x.$$

For a finite-dimensional vector space $V \simeq D^n$ over the division algebra D this shows that

$$\mathrm{End}_D(V) \;\simeq\; \mathrm{Mat}_{n\times n}(\mathrm{End}_D(D)) \;\simeq\; \mathrm{Mat}_{n\times n}(D^{op})$$

is isomorphic to a matrix algebra over the opposite division algebra.

For k algebraically closed, this allows to determine the number of irreducible modules in terms of the center $Z(A) = \{z \in A \mid az = za \,\text{for all}\, a \in A\}$:

Corollary 2.8. *The center of any semisimple algebra A is the sum of those of its simple subalgebras,*

$$Z(A) \;=\; \bigoplus_{\rho\in\mathrm{Irr}(A)} Z(A_\rho).$$

If k is algebraically closed, then each summand on the right is 1-dimensional and we then get

$$|\mathrm{Irr}(A)| \;=\; \dim_k Z(A).$$

Proof. We know that any semisimple algebra decomposes as an algebra into the direct sum of its isotypic components. Hence any element $z \in A$ can be decomposed in the form

$$z \;=\; \sum_{\rho \in \mathrm{Irr}(A)} z_\rho \quad \text{with} \quad z_\rho \;\in\; A_\rho.$$

This decomposition is compatible with the algebra structure, hence $(za)_\rho = z_\rho a_\rho$ and $(az)_\rho = a_\rho z_\rho$ for all $a \in A$. Thus the center of the algebra decomposes into a direct sum as claimed:

$$z \;\in\; Z(A) \quad \Longleftrightarrow \quad z_\rho \;\in\; Z(A_\rho) \quad \text{for all} \quad \rho \;\in\; \mathrm{Irr}(A).$$

Notice that under the inclusion map from the previous corollary, the center $Z(A_\rho)$ maps into the subalgebra $\mathrm{End}_A(\rho) \subseteq \mathrm{End}_k(\rho)$. If k is algebraically closed, then by Schur's lemma

$$\dim_k Z(A_\rho) \;=\; \dim_k A_\rho \cap \mathrm{End}_A(\rho) \;\leq\; \dim_k \mathrm{End}_A(\rho) \;=\; 1$$

and the claim follows. □

Example 2.9. Suppose that the field k is algebraically closed, and let G be a finite group with $char(k) \nmid |G|$. Then for the group algebra $A = k[G]$ the center is given by

$$Z(A) \;=\; \Big\{ \sum_{g \in G} a_g e_g \;\Big|\; a_{h^{-1}gh} = a_g \text{ for all } h \in G \Big\},$$

hence the number of irreducible representations is equal to the number of conjugacy classes in the group:

$$|\mathrm{Irr}_k(G)| \;=\; |\,\mathrm{Cl}(G)| \quad \text{for the set} \quad \mathrm{Cl}(G) \;=\; \big\{\text{conjugacy classes of } G\big\}.$$

In particular, this number only depends on the group but not on the base field over which the representations are defined. In general no canonical bijection between the sets $\mathrm{Irr}_k(G)$ and $\mathrm{Cl}(G)$ is known, although we will see later that if $G = \mathfrak{S}_d$ is a symmetric group, such a bijection can be constructed explicitly via partitions.

3. Characters and orthogonality

We now always assume that the base field k is algebraically closed. Let G be a finite group with $char(k) \nmid |G|$. We have seen that every $V \in \mathrm{Rep}_k(G)$ is semisimple and hence decomposes as a direct sum

$$V \;\simeq\; \bigoplus_{\rho \in \mathrm{Irr}_k(G)} \rho^{\oplus m_\rho} \quad \text{with multiplicities} \quad m_\rho \;=\; m_\rho(V) \;\in\; \mathbb{N}_0,$$

but how can we compute these multiplicities for a given representation? To get an idea, let us take a look at the case where $\rho = \mathbf{1}$ is the 1-dimensional trivial representation. Here

$$m_{\mathbf{1}}(V) \;=\; \dim_k(V^G)$$

is just the dimension of the subspace of invariants under the group action. We can then use the averaging argument from Maschke's theorem: The endomorphism given by

$$p \;=\; \frac{1}{|G|} \sum_{g \in G} \rho(g) \;\in\; \mathrm{End}_k(V)$$

is a projector onto the subspace of invariants. In particular, on a numerical level this gives

$$\dim_k(V^G) \cdot 1_k \;=\; tr(p) \;=\; \frac{1}{|G|} \sum_{g \in G} tr(\rho(g))$$

where 1_k denotes the identity element of the field k. If $char(k) = 0$, then $\mathbb{Z} \subset k$ and it follows from the above equation that the multiplicity of the trivial representation can be recovered if we know the traces of the action of all group elements.

DEFINITION 3.1. The *character* of a representation $V = (V, \rho) \in \mathrm{Rep}_k(G)$ is the function

$$\chi_V : \quad G \longrightarrow k, \quad g \mapsto \chi_V(g) \;=\; tr(\rho(g)).$$

Since the trace of a matrix is invariant under conjugation, this character descends to a function on the set of conjugacy classes which we also denote by $\chi_V : \mathrm{Cl}(G) \longrightarrow k$.

For $\dim_k V = 1$ the character χ_V takes values in the multiplicative group $k^\times$ and is the group homomorphism given by the representation: This is what is usually called a character outside of representation theory. For $\dim_k V > 1$, however, the characters in the above sense are no longer group homomorphisms. The motivation for looking at them is that we want to extract the relevant information about a representation with as little redundancy as possible:

(1) Scalars $\chi_V(g) \in k$ are easier to work with than matrices $\rho(g) \in Gl(V)$,
(2) Nevertheless, we will see that the knowledge of all these scalars determines the representation $\rho : G \to Gl(V)$ up to isomorphism if $char(k) = 0$, for instance

$$\dim_k V \;=\; \chi_V(1) \quad \text{for the neutral element} \quad 1 \;\in\; G.$$

Roughly speaking, the passage from a representation to its character is like the one from a matrix to its characteristic polynomial:

EXAMPLE 3.2. Let $(V, \rho) \in \mathrm{Rep}_k(G)$. If k is algebraically closed, then for $g \in G$ let $\lambda_1, \dots, \lambda_n \in k$ be the eigenvalues of the corresponding endomorphism, so that the characteristic polynomial takes the form

$$p_{\rho(g)}(t) \;:=\; \det\bigl(t \cdot id_V - \rho(g)\bigr) \;=\; \prod_{i=1}^{n} (t - \lambda_i) \;\in\; k[t].$$

From the character $\chi_V : G \longrightarrow k$ of the representation we then recover the power sums

$$p_\nu(\lambda_1, \dots, \lambda_n) \;:=\; \lambda_1^\nu + \cdots + \lambda_n^\nu \;=\; \chi_V(g^\nu) \quad \text{for all} \quad \nu \;\in\; \mathbb{Z}.$$

If $char(k) = 0$, these power sums determine the coefficients of the characteristic polynomial, i.e. the elementary symmetric functions

$$e_\nu(\lambda_1, \dots, \lambda_n) \;:=\; \sum_{i_1 < \cdots < i_\nu} \lambda_{i_1} \cdots \lambda_{i_\nu},$$

by the formulae

$$\begin{aligned}
e_1 &= p_1, \\
e_2 &= \tfrac{1}{2!}\bigl(p_1^2 - p_2\bigr), \\
e_3 &= \tfrac{1}{3!}\bigl(p_1^3 - 3p_1p_2 + 2p_3\bigr), \\
e_4 &= \tfrac{1}{4!}\bigl(p_1^4 - 6p_1^2p_2 + 3p_2^2 + 8p_1p_3 - 6p_4\bigr), \\
&\;\vdots \\
e_\nu &= \tfrac{1}{\nu}\sum_{i=1}^{\nu} (-1)^{i+1}\, p_i\, e_{\nu-i} \;=\; (-1)^\nu \sum_{m_1 + 2m_2 + \cdots + \nu m_\nu = \nu} \; \prod_{i=1}^{\nu} \frac{1}{m_i!}\Bigl(-\frac{p_i}{i}\Bigr)^{m_i}
\end{aligned}$$

In other words, knowing the character of a representation amounts to the same as knowing the eigenvalues of the action of all group elements.

Lemma 3.3. *For $V, W \in \mathrm{Rep}_k(G)$, the characters of the dual, direct sum and tensor product are given by*

$$\chi_{V^*}(g) \;=\; \chi(g^{-1}), \quad \chi_{V\oplus W}(g) \;=\; \chi_V(g) + \chi_V(g), \quad \chi_{V\otimes W}(g) \;=\; \chi_V(g)\cdot\chi_V(g).$$

If $char(k) \neq 2$, then for the symmetric and exterior square one has

$$\chi_{Sym^2(V)}(g) \;=\; \tfrac{1}{2}\big((\chi_V(g))^2 + \chi_V(g^2)\big),$$
$$\chi_{Alt^2(V)}(g) \;=\; \tfrac{1}{2}\big((\chi_V(g))^2 - \chi_V(g^2)\big).$$

Proof. The claim for the dual representation is clear since the action on the dual is given by the inverse of the transpose matrix. Furthermore, if $\lambda_1, \dots, \lambda_n$ resp. $\mu_1, \dots, \mu_m$ denote the eigenvalues of g acting on V resp. W, then clearly the eigenvalues

- on $V \oplus W$ are $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_m$,
- on $V \otimes W$ are the products $\lambda_i \mu_j$ with $1 \leq i \leq n$ and $1 \leq j \leq m$,
- on $Sym^2(V)$ are the products $\lambda_i \lambda_j$ with $1 \leq i \leq j \leq n$,
- on $Alt^2(V)$ are the products $\lambda_i \lambda_j$ with $1 \leq i < j \leq n$,

so the result follows by taking the sum of the eigenvalues in each case. □

As an exercise you may try to generalize the above to a formula for the character of $Sym^n(V)$ and $Alt^n(V)$ for $n > 2$. In general, the data of the characters is usually represented in the form of a *character table*, by which we understand the square matrix of size $|\mathrm{Irr}_k(G)| = |\,\mathrm{Cl}(G)|$ whose columns correspond to the conjugacy classes and whose rows give the character values of the irreducible representations on these conjugacy classes. Let us compute a simple example by hand:

Example 3.4. We have seen earlier that the symmetric group $\mathfrak{S}_3$ has precisely three irreducible representations $\mathbf{1}$, *sgn* and $W = \{(x, y, z) \in \mathbb{C}^3 \mid x + y + z = 0\}$ over $k = \mathbb{C}$. Since the conjugacy classes in the symmetric group are given by the cycle types of permutations, we may take $(1), (12)$ and (123) as representatives for the classes and obtain the following character table:

	(1)	(12)	(123)
$\chi_{\mathbf{1}}$	1	1	1
χ_{sgn}	1	-1	1
χ_W	2	0	-1

For the last two entries, note that in the basis $(1, -1, 0)$, $(0, 1, -1)$ of W the group action is given by matrices

$$\rho((12)) \;=\; \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \rho((123)) \;=\; \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}.$$

Note that the rows of the above character table are linearly independent! Hence by the formula for the character of a direct sum, the multiplicities of the irreducible representations in any semisimple representation $V \in \mathrm{Rep}_{\mathbb{C}}(G)$ can be read off from the character

$$\chi_V \;=\; \sum_{\rho \in \mathrm{Irr}_{\mathbb{C}}(G)} m_\rho(V) \cdot \chi_\rho.$$

We will see below that this is true in general, but let us take a brief look at some easy examples: From the formulae in the previous lemma and the above character table one computes

	(1)	(12)	(123)
$\chi_{W\otimes sgn}$	2	0	−1
$\chi_{W\otimes W}$	4	0	1
$\chi_{Sym^2(W)}$	3	1	0
$\chi_{Alt^2(W)}$	1	−1	1
⋮	⋮	⋮	⋮

and writing the rows of the above table as linear combinations of those from the character table we recover

$$W \otimes sgn \simeq W, \quad W \otimes W \simeq \mathbf{1} \oplus sgn \oplus W, \quad Sym^2(W) \simeq W \oplus \mathbf{1}, \quad Alt^2(W) \simeq sgn.$$

Thus we have succeeded in putting our previous *ad hoc* computations for $G = \mathfrak{S}_3$ in a systematic framework. In order to generalize this to arbitrary finite groups, we only need to show that the rows of the character table are always linearly independent. For this we make the following

Definition 3.5. By a *class function* we mean a function $f : G \to k$ that is invariant under conjugation. The vector space

$$\mathscr{C}(G) = \{f : G \to k \mid f(gxg^{-1}) = f(x) \text{for all} g \in G\} = \{f : \mathrm{Cl}(G) \to k\}$$

of such class functions is equipped with the symmetric bilinear form

$$\langle \cdot, \cdot \rangle : \quad \mathscr{C}(G) \times \mathscr{C}(G) \longrightarrow k, \quad \langle f, h \rangle \;=\; \frac{1}{|G|} \sum_{g \in G} f(g) h(g^{-1}),$$

which is easily seen to be *nondegenerate* in the sense that it induces an *iso*morphism of vector spaces

$$\mathscr{C}(G) \xrightarrow{\sim} \mathscr{C}(G)^* \;=\; \mathrm{Hom}_k(\mathscr{C}(G), k), \quad f \;\mapsto\; \langle f, - \rangle.$$

For characters of representations this bilinear form has a very concrete meaning, where again the equality below has to be read in the field k but for $char(k) = 0$ remains true as an equality of natural numbers:

Theorem 3.6 (Multiplicity formula). *For any* $V, W \in \mathrm{Rep}_k(G)$,

$$\dim_k \mathrm{Hom}_G(V, W) \;=\; \langle \chi_V, \chi_W \rangle.$$

Proof. By adjunction we may identify the equivariant morphisms from V to W with the invariants in $V^* \otimes W = \mathrm{Hom}_k(V, W)$. Hence the formula for the character of duals and tensor products together with our computation of invariants from the beginning of this section gives

$$\begin{aligned} \dim_k \mathrm{Hom}_G(V, W) \;=\; \dim_k (V^* \otimes W)^G \;&=\; \frac{1}{|G|} \sum_{g \in G} \chi_{V^* \otimes W}(g) \\ &=\; \frac{1}{|G|} \sum_{g \in G} \chi_V(g^{-1}) \cdot \chi_W(g) \;=\; \langle \chi_V, \chi_W \rangle \end{aligned}$$

as required. □

As a special case of the above, it follows that the rows of the character table are indeed linearly independent:

Corollary 3.7 (First orthogonality relation). *The $\chi_\rho \in \mathscr{C}(G)$ with $\rho \in \mathrm{Irr}_k(G)$ form an orthonormal basis for the vector space of class functions in the sense that for all $\rho, \sigma \in \mathrm{Irr}_k(G)$,*

$$\langle \chi_\rho, \chi_\sigma \rangle = \begin{cases} 1 & \text{if } \sigma = \rho, \\ 0 & \text{otherwise.} \end{cases}$$

Hence

$$f = \sum_{\rho \in \mathrm{Irr}_k(G)} \langle f, \chi_\rho \rangle \cdot \chi_\rho \quad \text{for all} \quad f \in \mathscr{C}(G).$$

Proof. The orthonormality follows directly from the previous theorem and Schur's lemma since k is algebraically closed. In particular, the characters of the irreducible representations are linearly independent class functions. Hence they form a basis since

$$\dim_k \mathscr{C}(G) = |\operatorname{Cl}(G)| = |\mathrm{Irr}_k(G)|$$

as we have seen earlier. In particular, every class function $f \in \mathscr{C}(G)$ is a linear combination of these characters and the coefficients can be recovered by taking scalar products with the orthonormal basis vectors. □

We can now see that at least for $char(k) = 0$, the *character* of a representation indeed *characterizes* it up to isomorphism. Note that this reduces the question of the simultaneous conjugacy of two sets of matrices to the much simpler question of an equality of functions:

Corollary 3.8. *For $char(k) = 0$, two representations $V, W \in \mathrm{Rep}_k(G)$ are isomorphic iff they have the same character:*

$$W \simeq V \iff \chi_W = \chi_V.$$

Furthermore,

$$V \text{ is irreducible} \iff \langle \chi_V, \chi_V \rangle = 1.$$

Proof. The orthogonality relation allows to find the multiplicity of $\rho \in \mathrm{Irr}_k(G)$ in V as

$$m_\rho(V) = \langle \chi_\rho, \chi_V \rangle.$$

So the multiplicities only depends on the character and we are done. □

Since the number of rows and columns of the character table is the same due to the identity $|\mathrm{Irr}_k(G)| = |\operatorname{Cl}(G)|$, the orthogonality of the rows also implies the one of the columns. There is a twist here since the columns do not label group elements but conjugacy classes

$$\operatorname{Cl}(g) := \{ hgh^{-1} \mid h \in G \} \subseteq G$$

for $g \in G$, so that in

$$\langle f, g \rangle = \frac{1}{|G|} \sum_{g \in G} f(g) h(g^{-1}) = \frac{1}{|G|} \sum_{x \in \operatorname{Cl}(G)} |x| \cdot f(x) h(x^{-1}),$$

each summand on the right is weighted with the size of the corresponding conjugacy classes. Since

$$|G| = |\operatorname{Cl}(g)| \cdot |Z_G(g)|$$

for the centralizer

$$Z_G(g) = \{ h \in G \mid hg = gh \},$$

we obtain

Corollary 3.9 (Second orthogonality relation). *For $g, h \in G$ one has in k the identity*

$$\sum_{\rho \in \mathrm{Irr}_k(G)} \chi_\rho(g)\chi_\rho(h^{-1}) \;=\; \begin{cases} |Z_G(g)| & \text{if } h \in \mathrm{Cl}(g), \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Consider the class function $f \in \mathscr{C}(G)$ which is the characteristic function of the class $\mathrm{Cl}(g^{-1})$,

$$f(x) \;=\; \begin{cases} 1 & \text{if } x \in \mathrm{Cl}(g^{-1}), \\ 0 & \text{otherwise.} \end{cases}$$

For $\rho \in \mathrm{Irr}_k(G)$ then

$$\langle f, \chi_\rho \rangle \;=\; \frac{1}{|G|}\sum_{h \in G} f(h)\chi_\rho(h^{-1}) \;=\; \frac{|\,\mathrm{Cl}(g)|}{|G|} \cdot \chi_\rho(g) \;=\; \frac{1}{|Z_G(g)|} \cdot \chi_\rho(g),$$

so by orthonormality

$$f \;=\; \frac{1}{|Z_G(g)|} \sum_{\rho \in \mathrm{Irr}_k(G)} \chi_\rho(g) \cdot \chi_\rho$$

and the result follows by evaluating this class function at $x = h^{-1}$. □

Remark 3.10. If $k = \mathbb{C}$, the eigenvalues of any endomorphism of finite order are complex roots of unity. In particular, their inverse is equal to its complex conjugate and

$$\chi_V(g^{-1}) \;=\; \overline{\chi_V(g)} \quad \text{for all} \quad V \in \mathrm{Rep}_{\mathbb{C}}(G), \quad g \in G.$$

We may then everywhere in the above replace the nondegenerate bilinear form $\langle \cdot, \cdot \rangle$ by the pairing

$$(\cdot, \cdot): \quad \mathscr{C}(G) \times \mathscr{C}(G) \;\longrightarrow\; \mathbb{C}, \quad (f, h) \;=\; \frac{1}{|G|}\sum_{g \in G} f(g)\overline{h(g)},$$

which is a Hermitian scalar product in the usual sense. We can hence view $\mathscr{C}(G)$ as a finite-dimensional Hilbert space, so that the above becomes a toy example of harmonic analysis; we discuss this for finite abelian groups in the next section.

4. Harmonic analysis on finite abelian groups

If G is a finite abelian group, then so is its *Pontryagin dual* $\widehat{G} = \mathrm{Hom}(G, \mathbb{C}^\times)$ with the pointwise multiplication

$$(\chi_1 \cdot \chi_2)(g) \;=\; \chi_1(g) \cdot \chi_2(g) \quad \text{for} \quad \chi_1, \chi_2 \;\in\; \widehat{G} \quad \text{and} \quad g \;\in\; G.$$

In fact this dual is non-canonically isomorphic to the original group: Since taking duals is compatible with direct products (exercise), this is easily reduced to the case $G = \mathbb{Z}/n\mathbb{Z}$, and in this case any primitive n-th root of unity $\zeta \in \mathbb{C}$ gives an isomorphism

$$G \;\xrightarrow{\sim}\; \widehat{G} \;=\; \mathrm{Hom}(G, \mathbb{C}^\times) \quad \text{via} \quad 1 \;\mapsto\; (a \mapsto \zeta^n).$$

While this isomorphism depends on the choice of a root of unity, the double dual is *canonically* isomorphic to the original group:

Lemma 4.1 (Pontryagin duality for finite abelian groups). *For the double dual we have a canonical isomorphism*

$$ev: \quad G \;\xrightarrow{\sim}\; \widehat{\widehat{G}} \;=\; \mathrm{Hom}(\widehat{G}, \mathbb{C}^\times), \quad g \;\mapsto\; (\chi \mapsto \chi(g)).$$

Proof. The evaluation map ev is a homomorphism of finite abelian groups. We know that its source and target are non-canonically isomorphic, so we only need to see that ev is injective. This boils down to the statement that for any $g \in G \setminus \{1\}$ there exists a character $\chi : G \longrightarrow \mathbb{C}^\times$ with $\chi(g) \neq 1$. If $G = \mathbb{Z}/n\mathbb{Z}$ is a finite cyclic group, then this statement immediately follows from the description of characters given above, and the general case easily reduces to the cyclic one. □

The above is an elementary special case of the more general *Pontryagin duality* for locally compact abelian groups and their continuous characters, and our previous orthogonality relations for characters will then translate into Fourier analysis. To make this analogy more vivid, let $\mathscr{L}^2(G) = \{f : G \longrightarrow \mathbb{C}\}$ denote the vector space of all complex valued functions on the group, equipped with the Hermitian scalar product

$$(f_1, f_2)_G \;=\; \frac{1}{|G|}\sum_{g\in G} f_1(g)\overline{f_2(g)} \quad \text{for} \quad f_i : G \longrightarrow \mathbb{C}.$$

Note that the Hermitian scalar product for the dual group restricts on $G \overset{ev}{\hookrightarrow} \mathscr{L}^2(\widehat{G})$ via ev to

$$(g_1, g_2)_{\widehat{G}} \;=\; \frac{1}{|G|}\sum_{\chi\in\widehat{G}} \chi(g_1)\overline{\chi(g_2)} \quad \text{for} \quad g \;\in\; G.$$

So the second orthogonality relation is just the first orthogonality relation for the dual group:

Lemma 4.2 (Harmonic analysis on finite abelian groups). *In the above setting,*

(1) *we have the orthogonality relations*

$$\begin{aligned}(\chi_1, \chi_2)_G \;&=\; \delta_{\chi_1,\chi_2} \quad \textit{for} \quad \chi_i \;\in\; \widehat{G},\\ (g_1, g_2)_{\widehat{G}} \;&=\; \delta_{g_1,g_2} \quad \textit{for} \quad g_i \;\in\; G.\end{aligned}$$

(2) *any function $f : G \longrightarrow \mathbb{C}$ admits a Fourier development*

$$f \;=\; \sum_{\chi\in\widehat{G}} \hat{f}(\chi)\cdot\chi \quad \textit{where} \quad \hat{f}(\chi) \;=\; (f,\chi) \;=\; \frac{1}{|G|}\sum_{g\in G} f(g)\overline{\chi(g)}.$$

and the corresponding dual statement holds for the functions $\widehat{G} \longrightarrow \mathbb{C}$.

(3) *the Plancherel formula holds: Viewing the Fourier coefficients as functions on the dual group we get an isometry*

$$\mathscr{L}^2(G) \;\xrightarrow{\sim}\; \mathscr{L}^2(\widehat{G}), \quad f \;\mapsto\; \sqrt{|G|}\cdot\hat{f} \;=\; (\chi \mapsto \sqrt{|G|}\cdot\hat{f}(\chi)).$$

Proof. Essentially everything has been proven in the previous section, except for the Plancherel formula. To show that the assignment $f \mapsto \hat{f}$ is an isometry up to a factor $\sqrt{|G|}$, we must see

$$(f_1, f_2)_G \;=\; |G|\cdot(\hat{f}_1, \hat{f}_2)_{\widehat{G}} \quad \text{for all} \quad f_1, f_2 \;\in\; \mathscr{L}^2(G).$$

It suffices to check this on the orthonormal basis of characters. But if $f_i \in \widehat{G}$ is a character, one easily sees that the Fourier coefficient $\hat{f}_i \in \mathscr{L}^2(\widehat{G})$ is the function given by

$$\hat{f}_i(\chi) \;=\; \delta_{f_1,\chi} \;=\; \begin{cases} 1 & \text{if } \chi = f_i,\\ 0 & \text{otherwise.}\end{cases}$$

Hence $(\hat{f}_1, \hat{f}_2) = \frac{1}{|G|}\cdot\delta_{f_1,f_2}$ and the claim follows. □

Note that the above statement is entirely parallel to the Fourier transform for periodic functions on the real line when one takes $G = \mathbb{R}/\mathbb{Z}$ to be the circle and considers the Hilbert space $\mathscr{L}^2(G)$ of square integrable functions $f : G \to \mathbb{C}$ with respect to the Lebesgue measure. Replacing finite sums by integrals, one uses the scalar product

$$(f,h)_G \;=\; \int_G f(x)\overline{h(x)}dx \quad \text{for} \quad f,h \;\in\; \mathscr{L}^2(G).$$

Here the Pontryagin dual is

$$\widehat{G} \;=\; \mathrm{Hom}(G,\mathbb{C}^\times) \;\xrightarrow{\sim}\; \mathbb{Z} \quad \text{via} \quad e_n \;=\; (x \mapsto e^{2\pi i n x}) \;\mapsto\; n.$$

For $f \in \mathscr{L}^2(G)$ we have the Fourier development

$$f \;=\; \sum_{n\in\mathbb{Z}} \hat{f}(n)\cdot e_n \quad \text{with} \quad \hat{f}(n) \;=\; (f,e_n) \;=\; \int_G f(x)e_n(-x)dx$$

and the usual Plancherel formula says that the assignment

$$\mathscr{L}^2(G) \;\xrightarrow{\sim}\; \mathscr{L}^2(\widehat{G}) \;=\; \ell^2(\mathbb{Z}) \;=\; \{(c_n)_{n\in\mathbb{Z}} \mid c_n \in \mathbb{C}, \sum_{n\in\mathbb{Z}} |c_n|^2 < \infty\}, \quad f \;\mapsto\; \hat{f}$$

is an isometry. The area of harmonic analysis on locally compact abelian groups is a common generalization of the cases of finite abelian groups and the circle.

Even in the completely elementary setting of finite abelian groups, the above has nontrivial applications. We would like to discuss two such applications. The first application is a formula by Dedekind for *circulants*, i.e. determinants of square matrices all of whose rows are cyclic permutations of the first row: How does the determinant

$$D(X_0,\dots,X_{n-1}) \;=\; \det\begin{pmatrix} X_0 & X_1 & \dots & X_{n-1} \\ X_{n-1} & X_0 & \dots & X_{n-2} \\ \vdots & & & \vdots \\ X_1 & X_2 & \dots & X_0 \end{pmatrix} \;\in\; \mathbb{C}[X_0,X_1,\dots,X_{n-1}]$$

factor as a complex polynomial in n variables? This problem naturally arises in number theory if one tries to find the discriminant of the set of Galois conjugates of some algebraic number. Back to our general question, for $n = 2,3$ one computes by hand that

$$\begin{aligned} D(X_0,X_1) \;&=\; X_0^2 - X_1^2 \;=\; (X_0+X_1)(X_0-X_1), \\ D(X_0,X_1,X_2) \;&=\; X_0^3 + X_1^3 + X_2^3 - 3X_0X_1X_2 \;=\; \prod_{\zeta^3=1} (X_0 + \zeta X_1 + \zeta^2 X_2), \end{aligned}$$

so these determinants split into linear factors. It turns out that this is a general phenomenon which has a conceptual explanation via Fourier inversion without any messy computations:

THEOREM 4.3 (Dedekind's formula for abelian group determinants). *Let G be a finite abelian group and take a set of formal variables X_g indexed by the group elements $g \in G$. Then*

$$\det(X_{gh^{-1}})_{g,h\in G} \;=\; \prod_{\chi\in\widehat{G}} \Bigl(\sum_{g\in G} \chi(g)X_g\Bigr)$$

Proof. Both sides of the identity are polynomials with complex coefficients, so it suffices to show that the identity holds when we specialize the indeterminates X_g

to arbitrary values $a_g \in \mathbb{C}$. Now the specialized matrix $(a_{gh^{-1}})_{g,h\in G}$ describes the left translation

$$\lambda(a): \quad \mathbb{C}[G] \longrightarrow \mathbb{C}[G] \quad \text{by} \quad a \;=\; \sum_{g\in G} a_g e_g \;\in\; \mathbb{C}[G]$$

with respect to the standard basis, since

$$\lambda(a)(e_h) \;=\; \sum_{g\in G} a_g e_{gh} \;=\; \sum_{g\in G} a_{gh^{-1}} e_g.$$

So our task is to compute the determinant of the endomorphism $\lambda(a)$.

For this we replace the standard basis of the regular representation by the basis of the vectors

$$v_\chi \;=\; \sum_{g\in G} \chi(g)\cdot e_g \quad \text{for} \quad \chi \;\in\; \widehat{G}.$$

To see that these vectors form a basis, note that there are precisely $|\widehat{G}| = |G|$ such vectors; and they are linearly independent because a linear combination of them can vanish only if the corresponding linear combination of characters vanishes, which by Fourier inversion can happen only if all coefficients are zero. Hence we indeed have a basis, and this is a basis of eigenvectors for the action of $\lambda(a)$ since

$$\begin{aligned}
\lambda(a)(v_\chi) \;&=\; \Big(\sum_{g\in G} a_g e_g\Big)\Big(\sum_{h\in G} \chi(h) e_h\Big) \\
&=\; \sum_{g\in G}\Big(\sum_{h\in G} a_g \chi(h)\Big) e_{gh} \\
&=\; \sum_{g\in G} a_g \chi(g^{-1}) \sum_{h\in G} \chi(gh) e_{gh} \;=\; \sum_{g\in G} a_g \overline{\chi(g)}\cdot v_\chi.
\end{aligned}$$

Taking the product of all eigenvalues one obtains

$$\det(a_{gh^{-1}})_{g,h\in G} \;=\; \prod_{\chi\in\widehat{G}} \Big(\sum_{g\in G} a_g \overline{\chi(g)}\Big) \;=\; \prod_{\chi\in\widehat{G}} \Big(\sum_{g\in G} a_g \chi(g)\Big)$$

because with χ also the conjugate $\overline{\chi}$ runs over all complex characters of G. Hence the claim follows. □

In the case of finite nonabelian groups, the determinant in the above theorem has irreducible factors of degree > 1, and the study of its factorization has led Frobenius to the general notion of characters and the discovery of representation theory. Before we come back to the latter, we briefly mention a second application of harmonic analysis for finite abelian groups:

Theorem 4.4 (Dirichlet's theorem on primes in arithmetic progressions). *For any natural numbers $a, m \in \mathbb{N}$ with $\gcd(a,m) = 1$, there are infinitely many prime numbers*

$$p \;\equiv a \mod m.$$

Idea of the proof. One way to show there are infinitely many prime numbers uses the Euler product expansion

$$\zeta(s) \;:=\; \sum_{n\geq 1} n^{-s} \;=\; \prod_{p \text{ prime}} \frac{1}{1-p^{-s}}$$

of the Riemann zeta function. The left hand side is a holomorphic function on the half plane $\{s \in \mathbb{C} \mid \mathrm{Re}(s) > 1\}$ but has a pole at $s = 1$, hence the right hand side must have a pole there as well and therefore the product must involve infinitely many factors. If one wants to avoid considering infinite products, one may take the logarithm and use

$$\log \zeta(s) \;=\; \sum_{p \text{ prime}} p^{-s} + O(1) \quad \text{for} \quad s \downarrow 1,$$

which is obtained from the previous identity via $\log \frac{1}{1-x} = \sum_{n\geq 1} \frac{x^n}{n}$.

If we want to adapt this reasoning to primes in an arithmetic progression, we run into the problem how to control the left hand side if on the right hand side we only consider primes $p \equiv a \bmod m$. The idea is to single out $(a \bmod m) \in G = (\mathbb{Z}/m\mathbb{Z})^\times$ by harmonic analysis: For any character

$$\chi \in \widehat{G} = \mathrm{Hom}(G, \mathbb{C}^\times)$$

we define the corresponding *Dirichlet character* to be the function $\chi : \mathbb{Z} \longrightarrow \mathbb{C}$ given by

$$\chi(n) \;=\; \begin{cases} \chi(n \bmod m) & \text{if } \gcd(m,n) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

The second orthogonality relation gives

$$\sum_{\chi \in \widehat{G}} \overline{\chi(a)}\chi(n) \;=\; \begin{cases} \varphi(m) & \text{if } n \equiv a \mod m, \\ 0 & \text{otherwise,} \end{cases}$$

where $\varphi(m) = |(\mathbb{Z}/m\mathbb{Z})^\times|$ denotes Euler's totient function. We therefore obtain that

$$\sum_{\substack{p \equiv a \bmod m \\ p \text{ prime}}} p^{-s} \;=\; \frac{1}{\varphi(m)} \sum_{\chi \in \widehat{G}} \overline{\chi(a)} \sum_{p \text{ prime}} \chi(p) \cdot p^{-s}$$

where the inner sum on the right hand side no longer involves any congruence condition on the prime numbers over which we sum! It remains to show the right hand side diverges when $s \downarrow 1$. For this we express it via generalizations of $\zeta(s)$, the *Dirichlet L-functions*

$$L(s,\chi) \;=\; \sum_{n \geq 1} \chi(n) \cdot n^{-s} \;=\; \prod_{p \text{ prime}} \frac{1}{1 - \chi(p)p^{-s}}.$$

Similarly as for $\zeta(s)$ one shows

$$\log L(s,\chi) \;=\; \sum_{p \text{ prime}} \chi(p) \cdot p^{-s} + O(s) \quad \text{for} \quad s \downarrow 1.$$

Delicate estimates from complex analysis, which are really the main part of the proof besides character theory and for which we refer to any book on analytic number theory, show that for each non-trivial character χ the logarithm on the left hand side converges to a finite number as $s \downarrow 1$. On the other hand, for the trivial character $\chi = \mathbf{1}$ the left hand side diverges because $L(s,\mathbf{1}) - \zeta(s)$ is a finite sum of terms coming from primes $p \mid m$. □

5. How to read character tables

From now on we will always assume that the base field k is algebraically closed with $char(k) = 0$. In fact we could as well work with the complex numbers or its subfield

$$\overline{\mathbb{Q}} \;=\; \big\{ a \in \mathbb{C} \mid \exists f(x) \in \mathbb{Q}[x] \setminus \{0\} \text{ with } f(a) \;=\; 0 \big\} \;\subset\; \mathbb{C}$$

of algebraic numbers:

Lemma 5.1. *For any algebraically closed field k with char$(k) = 0$,*

(1) *the functor* $(-) \otimes_{\overline{\mathbb{Q}}} k : \mathrm{Rep}_{\overline{\mathbb{Q}}}(G) \xrightarrow{\sim} \mathrm{Rep}_k(G)$ *is an equivalence.*
(2) *the character values of any* $V \in \mathrm{Rep}_k(G)$ *are algebraic integers.*

Proof. We know that over any algebraically closed field of characteristic zero, a representation is irreducible iff its character χ satisfies $\langle \chi, \chi \rangle = 1$. As this condition is stable under field extensions, the functor

$$(-) \otimes_{\overline{\mathbb{Q}}} k : \quad \mathrm{Rep}_{\overline{\mathbb{Q}}}(G) \longrightarrow \mathrm{Rep}_k(G)$$

sends irreducibles to irreducibles. The number of irreducibles is independent of the base field since

$$|\mathrm{Irr}_{\overline{\mathbb{Q}}}(G)| \;=\; |\,\mathrm{Cl}(G)| \;=\; |\mathrm{Irr}_k(G)|.$$

Furthermore

$$\begin{aligned} V \;\simeq\; W \quad &\iff \quad \chi_V \;=\; \chi_W \\ &\iff \quad \chi_{V \otimes_{\overline{\mathbb{Q}}} k} \;=\; \chi_{W \otimes_{\overline{\mathbb{Q}}} k} \\ &\iff \quad V \otimes_{\overline{\mathbb{Q}}} k \;\simeq\; W \otimes_{\overline{\mathbb{Q}}} k \end{aligned}$$

for $V, W \in \mathrm{Rep}_{\overline{\mathbb{Q}}}(G)$, hence the base extension functor induces a bijection between the sets of isomorphism classes of irreducible representations. By semisimplicity it is then an equivalence of categories.

The statement about character values follows from the fact that the eigenvalues of any endomorphism of finite order are roots of unity, hence algebraic integers, and the sum of algebraic integers is again an algebraic integer. □

The same argument with character values also shows that we may read off from the character table which group elements act trivially on a representation:

Corollary 5.2. *For any representation* $(V, \rho) \in \mathrm{Rep}_{\mathbb{C}}(G)$ *and* $g \in G$*, we have the equivalence*

$$\rho(g) = id_V \quad \iff \quad \chi_V(g) = \dim_{\mathbb{C}} V$$

Proof. Being of finite order, the matrix of the endomorphism $\rho(g)$ is conjugate to a diagonal matrix whose eigenvalues $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ are roots of unity. Now clearly $\chi_V(g) = \lambda_1 + \cdots + \lambda_n$ is equal to the dimension n iff all the eigenvalues are one, and the claim follows. □

This being said, let us see what information we can read off easily from the character table. By the second orthogonality relation the columns of the character table determine the size of the centralizer of the elements $g \in G$ in the corresponding conjugacy class:

$$|Z_G(g)| \;=\; \sum_{\chi \in \mathrm{Irr}_k(G)} \chi(g)\chi(g^{-1}).$$

In particular, the column of the trivial conjugacy class $\{1\}$ determines the group order $|G| = |Z_G(1)|$, hence also the size

$$|\,\mathrm{Cl}(g)| \;=\; \frac{|G|}{|Z_G(g)|}$$

of the individual conjugacy classes. Often these numerical values are included above the character table to make it easier to compute scalar products with the rows. The above corollary also allows to determine all normal subgroups $N \trianglelefteq G$ as unions of conjugacy classes:

LEMMA 5.3. *For any character χ put* $\ker(\chi) = \{g \in G \mid \chi(g) = \chi(1)\}$.

(1) *For any* $V \in \mathrm{Rep}_k(G)$,

$$\ker(\chi_V) \;=\; \bigcap \left\{ \ker(\chi_\rho) \mid m_\rho(V) > 0 \right\}.$$

(2) *Any normal subgroup* $N \trianglelefteq G$ *has the form*

$$N \;=\; \bigcap_{i=1}^{n} \ker(\chi_{\rho_i}) \quad \textit{for suitable} \quad \rho_i \;\in\; \mathrm{Irr}_k(G).$$

(3) *The group G is simple iff* $\ker(\chi_\rho) = \{1\}$ *for all non-trivial* $\rho \in \mathrm{Irr}_k(G)$.

Proof. The inclusion $\subseteq$ in (1) follows from $|\chi_\rho(g)| \leq \dim_k(\rho)$ and the previous corollary if one compares

$$\chi_V(g) = \sum_\rho m_\rho(V) \cdot \chi_\rho(g) \quad \text{with} \quad \sum_\rho m_\rho(V) \dim_k(\rho),$$

and the converse is trivial. For (2) one then only needs to realize $N \trianglelefteq G$ as the kernel of some representation

$$\rho : \quad G \longrightarrow Gl(V),$$

for which one may take the permutation representation on $V = k[G/N]$. Part (3) is immediate from part (2). □

Note that since we know the sizes of all conjugacy classes, we then also obtain the order

$$|N| \;=\; \sum_{\mathrm{Cl}(g) \subseteq N} |\,\mathrm{Cl}(g)|$$

when $N \trianglelefteq G$ is written as a union of conjugacy classes. While there is no known way to determine the character table of N from the one of G, the situation for G/N is better:

LEMMA 5.4. *For $N \trianglelefteq G$, the quotient map $p : G \twoheadrightarrow G/N$ gives an equivalence of categories*

$$p^* : \quad \mathrm{Rep}_k(G/N) \xrightarrow{\sim} \left\{ V \in \mathrm{Rep}_k(G) \mid N \trianglelefteq \ker(\chi_V) \right\} \;\subseteq\; \mathrm{Rep}_k(G)$$

sending irreducibles to irreducibles, and a surjective map $p_* : \mathrm{Cl}(G) \twoheadrightarrow \mathrm{Cl}(G/N)$ *such that*

$$p_*(\mathrm{Cl}(g_1)) = p_*(\mathrm{Cl}(g_2)) \quad \Longleftrightarrow \quad \chi_V(g_1) = \chi_V(g_2) \;\; \textit{for all} \;\; V \in \mathrm{im}(p^*).$$

So starting from the character table of the group G, the character table of G/N is obtained by

- *deleting all rows corresponding to characters χ with* $N \not\trianglelefteq \ker(\chi)$,
- *deleting any multiple occurance of rows in the remaining table.*

Proof. The surjectivity of the quotient homomorphism $p : G \twoheadrightarrow G/N$ implies that the restriction functor $p^* : \mathrm{Rep}_k(G/N) \to \mathrm{Rep}_k(G)$ is fully faithful in the sense that

$$\mathrm{Hom}_G(p^*V, p^*W) \;\simeq\; \mathrm{Hom}_{G/N}(V, W) \quad \text{for all} \quad V, W \;\in\; \mathrm{Rep}_k(G/N).$$

Obviously the essential image of this restriction functor is the full subcategory of all $V \in \mathrm{Rep}_k(G)$ on which N acts trivially, which is equivalent to $N \trianglelefteq \ker(\chi_V)$ by what we have seen above. It is also clear that the induced map on conjugacy classes is a surjection $p_* : \mathrm{Cl}(G) \twoheadrightarrow \mathrm{Cl}(G/N)$. Finally, since the characters of G/N form a basis for the vector space of class functions $\mathscr{C}(G/N)$, we have

$$\begin{aligned} p_*(\mathrm{Cl}(g_1)) \;=\; p_*(\mathrm{Cl}(g_2)) \quad &\Longleftrightarrow \quad \chi_W(p(g_1)) \;=\; \chi_W(p(g_2)) \;\;\forall\, W \in \mathrm{Rep}_k(G/N) \\ &\Longleftrightarrow \quad \chi_{p^*W}(g_1) \;=\; \chi_{p^*W}(g_2) \;\;\forall\, W \in \mathrm{Rep}_k(G/N) \end{aligned}$$

and hence the claim follows. □

As an application we get that the solvability of a group can also be read off from the character table. Recall that a group G is called *solvable* if it admits an ascending chain of subgroups

$$\{1\} = G_0 \trianglelefteq G_1 \trianglelefteq G_2 \trianglelefteq \cdots \trianglelefteq G_n = G,$$

each being a normal subgroup of the following one, such that the quotients G_i/G_{i-1} are abelian for $i = 1, 2, \ldots, n$. Note that this is equivalent to the requirement that its *derived series*

$$G \trianglerighteq G^{(1)} \trianglerighteq G^{(2)} \trianglerighteq \cdots \quad \text{terminates at some finite level} \quad G^{(n)} \;=\; \{1\},$$

where the derived groups are defined recursively by

$$G^{(1)} \;=\; G' \;=\; [G,G] \quad \text{and} \quad G^{(n+1)} \;=\; [G^{(n)}, G^{(n)}].$$

Note that each term of the derived series is actually a *characteristic subgroup* in the sense that it is stable under any automorphism of the ambient group G, hence in particular normal in this group. So from the previous lemma we get

Corollary 5.5. *For any finite group G,*

(1) *the derived group is given by $G' = \cap\{\ker(\chi_\rho) \mid \dim_k(\rho) = 1\}$.*

(2) *its index is given by $[G : G'] = |\{\rho \in \mathrm{Irr}_k(G) \mid \dim_k(\rho) = 1\}|$.*

(3) *whether or not G is solvable may be read off from the character table.*

Proof. If an irreducible representation $\rho \in \mathrm{Irr}_k(G)$ satisfies $G' \leq \ker(\chi_\rho)$, then it factors over $G^{ab} = G/G'$ and is therefore 1-dimensional, being an irreducible representation of a finite abelian group. Conversely, since the group $Gl_1(k) = k^\times$ is abelian, any 1-dimensional representation factors over the abelianization, so it is trivial on the derived group. Hence (1) follows. For (2) note that

$$|\{\rho \in \mathrm{Irr}_k(G) \mid G' \leq \ker(\chi_\rho)\}| \;=\; |\mathrm{Irr}_k(G^{ab})| \;=\; |\,\mathrm{Cl}(G^{ab})| \;=\; |G^{ab}| \;=\; [G : G'].$$

Part (3) is then a consequence of the fact that a finite group is solvable iff it admits an ascending series of subgroups

$$\{1\} = G_0 \trianglelefteq G_1 \trianglelefteq G_2 \trianglelefteq \cdots \trianglelefteq G_n = G \quad \text{with} \quad G_i \trianglelefteq G \quad \text{for all } i$$

such that the successive quotients are of prime power order; indeed this can be seen by refining the derived series and noting that every group of prime power order is solvable (exercise). Since the normal subgroups, the inclusions between them and their sizes can be read off from the character table, so can the solvability. □

Above we have seen that the group elements $g \in G$ which act trivially on a given representation $(V, \rho) \in \mathrm{Rep}_k(G)$ can be characterized by $\chi_\rho(g) = \chi_\rho(1)$. Relaxing this condition, we call

$$Z(\chi_\rho) \;=\; \{g \in G \mid |\chi_\rho(g)| = \chi_\rho(1)\}$$

the *centralizer* of the corresponding character. It has the following meaning:

LEMMA 5.6. *With notations as above,*

(1) $Z(\chi_\rho) = \{g \in G \mid \rho(g) = \lambda \cdot id_V \text{ for some } \lambda \in k^\times\}$.

(2) $Z(G) = \bigcap \{Z(\chi_\rho) \mid \rho \in \mathrm{Irr}_k(G)\}$.

Proof. We may assume $k \subseteq \mathbb{C}$. If $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ are the eigenvalues of $\rho(g)$, then we know

$$|\chi_\rho(g)| \;=\; |\lambda_1 + \cdots + \lambda_n| \;\leq\; |\lambda_1| + \cdots + |\lambda_n| \;=\; n \;=\; \chi_\rho(1),$$

with equality holding iff all the eigenvalues are equal to the same number λ. But since $\rho(g)$ is diagonalizable, this occurs iff $\rho(g) = \lambda \cdot id_V$ and hence the first claim follows. For the second claim, note that by Schur's lemma the center $Z(G)$ acts by a scalar on every irreducible representation, so $\subseteq$ is clear. The reverse inclusion $\supseteq$ follows from the observation that if

$$[\rho(g), \rho(h)] = 1 \quad \text{for all } h \in G \text{ and all } \rho \in \mathrm{Irr}_k(G),$$

then this commutation relation holds for arbitrary representations $\rho \in \mathrm{Rep}_k(G)$, whence $g \in Z(G)$ since we may take ρ to be any faithful representation. □

In particular, the center of a group may be determined from the character table. Recalling that a group is called *nilpotent* if the *upper central series*

$$1 = Z_0 \trianglelefteq Z_1 \trianglelefteq \cdots \quad \text{defined by} \quad Z_{i+1} = \{g \in G \mid [g,h] =\in Z_i \forall h \in G\}$$

terminates at some level with the whole group $Z_n = G$, we obtain

COROLLARY 5.7. *Nilpotency of a finite group is detected by its character table.*

Proof. From the character table we may determine the center $Z(G)$, hence the character table of $G/Z(G)$, then use the latter to determine the center $Z(G/Z(G))$ and carry on like this to compute the upper central series. □

REMARK 5.8. In general the character table of a finite group does *not* determine the group up to isomorphism. For example, it turns out that the dihedral and the quaternion group

$$D \;=\; \langle a, b \mid a^4 = b^2 = abab = 1 \rangle \quad \text{and} \quad Q \;=\; \langle a, b \mid a^4 = 1, bab = a, a^2 = b^2 \rangle$$

share the same character table:

g	1	a	a^2	b	ab
$\lvert \mathrm{Cl}(g) \rvert$	1	2	1	2	2
$\lvert Z_G(g) \rvert$	8	4	8	4	4
$\mathbf{1}$	1	1	1	1	1
χ_2	1	1	1	−1	−1
χ_3	1	−1	1	1	−1
χ_4	1	−1	1	−1	1
χ_5	2	0	−2	0	0

Indeed, one easily sees that both groups have precisely five conjugacy classes and that these are represented by the given elements. It is also clear from the given presentation that there are four 1-dimensional representations given by $a \mapsto \pm 1$ and $b \mapsto \pm 1$. The character χ_5 of the remaining representation is then determined by the second orthogonality relation. To see that $D \not\simeq Q$ let us determine their maximal normal subgroups from the table: The maximal normal subgroups are the kernels

$$N_i \;=\; \ker(\chi_i) \quad \text{for} \quad i \;\in\; \{2, 3, 4\}.$$

One checks that

(1) for the dihedral group $N_2 \simeq \mathbb{Z}/4\mathbb{Z}$ but $N_3 \simeq N_4 \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$,
(2) for the quaternion group $N_2 \simeq N_3 \simeq N_4 \simeq \mathbb{Z}/4\mathbb{Z}$.

Hence the two groups are not isomorphic.

For the computation of character tables one usually starts with a few irreducible representations that are easy to describe, then obtains as many further ones by multilinear algebra, induction or restriction from smaller/bigger groups and *ad hoc* arguments, and then hopes to complete the character table via the orthogonality relations. Let us illustrate this with a simple example:

LEMMA 5.9. *The character table for the symmetric group $G = \mathfrak{S}_5$ looks as follows:*

g	1	(12)	(12)(34)	(123)	(123)(45)	(1234)	(12345)
$\lvert \mathrm{Cl}(g)\rvert$	1	10	15	20	20	30	24
$\lvert Z_G(g)\rvert$	120	12	8	6	6	4	5
$\mathbf{1}$	1	1	1	1	1	1	1
sgn	1	-1	1	1	-1	-1	1
U	4	2	0	1	-1	0	1
U'	4	-2	0	1	1	0	-1
V	5	1	1	-1	1	-1	0
V'	5	-1	1	-1	-1	1	0
W	6	0	-2	0	0	0	1

Proof. The conjugacy classes in the symmetric group $\mathfrak{S}_d$ are given by the cycle types of permutations. Furthermore, if for each $i \in \mathbb{N}$ we denote by $n_i \in \mathbb{N}_0$ the number of i-cycles in $g \in \mathfrak{S}_d$, then

$$|\operatorname{Cl}(g)| \;=\; \frac{d!}{\prod_i i^{n_i} \cdot n_i!}$$

since the denominator gives the number of multiple counts of the same permutation among the $d!$ fillings of the given cycle type with the labels $1, \dots, d$. This explains the listed conjugacy classes and their sizes.

The values of $\mathbf{1}$ and sgn are clear. We also know that the standard permutation representation decomposes in the form $\mathbb{C}^5 \simeq \mathbf{1} \oplus U$ for some $U \in \operatorname{Rep}_{\mathbb{C}}(G)$, hence we get

$$\chi_U(g) \;=\; \chi_{\mathbb{C}^5}(g) - 1 \;=\; |\operatorname{Fix}(g)| - 1.$$

Computing

$$\langle \chi_U, \chi_U \rangle \;=\; \cdots \;=\; 1,$$

we see that the underlying representation is irreducible and obtain the third row of the character table. We then get another irreducible representation $U' = U \otimes sgn$ for free: The tensor product of any irreducible representation with a 1-dimensional representation is again irreducible, and in our case $U' \not\simeq U$ as one sees from the character

$$\chi_{U'}(g) \;=\; \chi_U(g) \cdot sgn(g).$$

We can get further irreducible characters by multilinear algebra: For $W = Alt^2(U)$ one has

$$\chi_W(g) \;=\; \tfrac{1}{2}\big(\chi_U(g))^2 - \chi_U(g^2)\big)$$

and again

$$\langle \chi_W, \chi_W \rangle \;=\; \cdots \;=\; 1,$$

so the underlying representation is irreducible and we obtain the last row of the character table. Since there are precisely 7 conjugacy classes in $\mathfrak{S}_5$, we need to find two more representations

$$V, V' \in \operatorname{Rep}_{\mathbb{C}}(G).$$

Since $G^{ab} \simeq \mathbb{Z}/2\mathbb{Z}$, there can be no more 1-dimensional representations, so the missing two representations are of dimension > 1. But by the orthogonality relation for the first column,

$$(\dim_{\mathbb{C}} V)^2 + (\dim_{\mathbb{C}} V')^2 \;=\; 5! - 1^2 - 1^2 - 4^2 - 4^2 - 6^2 \;=\; 50$$

and therefore

$$\dim_{\mathbb{C}} V \;=\; \dim_{\mathbb{C}} V' \;=\; 5$$

is the only possibility. We can now successively fill in the remaining entries by using the second orthogonality relation. □

As a sanity check for the results from this section, you may read off from the above table that the centre $Z(\mathfrak{S}_5) = \{1\}$ is trivial and $\mathfrak{A}_5 = \ker(sgn) \trianglelefteq \mathfrak{S}_5$ is the only normal subgroup. The above character table also helps if we want to determine the character table of the alternating group:

Lemma 5.10. *Put* $a_{\pm} = \frac{1 \pm \sqrt{5}}{2} \in \mathbb{C}$. *Then the character table of* $G = \mathfrak{A}_5$ *looks as follows:*

g	1	(12)(34)	(123)	(12345)	(21345)
$\lvert \operatorname{Cl}(g)\rvert$	1	15	20	12	12
$\lvert Z_G(g)\rvert$	60	4	3	5	5
$\mathbf{1}$	1	1	1	1	1
U	4	0	1	-1	-1
V	5	1	-1	0	0
W_1	3	-1	0	a_+	a_-
W_2	3	-1	0	a_-	a_+

Proof. A conjugacy class of even permutations in $\mathfrak{S}_d$ either coincides with a conjugacy class in $\mathfrak{A}_d$ or splits in two distinct conjugacy classes in the latter. Such a splitting happens iff in the corresponding cycle type all cycle lengths are *pairwise distinct* and *odd*, including cycles of length one (exercise). For $\mathfrak{A}_5 \trianglelefteq \mathfrak{S}_5$ this happens precisely for the class of (12345).

This being said, we may carry over the character values for U and V from the previous table and compute

$$\langle \chi_U, \chi_U \rangle \;=\; \langle \chi_V, \chi_V \rangle \;=\; 1$$

also for the scalar product with respect to the alternating group. Therefore the restrictions $U|_{\mathfrak{A}_5}$ and $V|_{\mathfrak{A}_5}$ remain irreducible and we obtain the second and third rows of the character table. The remaining two rows are then again obtained from the orthogonality relations. □

The fact that the irreducible permutation representation of the symmetric group remains irreducible when restricted to the alternating group is no accident, as the following characterization of double transitivity shows:

Lemma 5.11. *An action of a finite group* G *on a finite set* X *with* $|X| \geq 3$ *is doubly transitive iff*

$$W_X \;=\; \Big\{ (a_x)_{x\in X} \;\Big|\; a_x \in k, \; \sum_{x\in X} a_x \;=\; 0 \Big\} \;\in\; \operatorname{Rep}_k(G) \quad \textit{is irreducible.}$$

Proof. We can assume that the action is transitive, since otherwise clearly both statements fail: For

$$V_X \;=\; \Big\{ (a_x)_{x\in X} \;\Big|\; a_x \in k \Big\} \;\simeq\; W_X \oplus \mathbf{1}$$

one easily checks

$$\langle \chi_{V_X}, \mathbf{1}\rangle \;=\; \dim_k (V_X)^G \;=\; \text{number of } G\text{-orbits on } X,$$

so for a non-transitive G-operation one would have $(W_X)^G \neq \{0\}$. Now we apply the same ideas to the permutation action on the set

$$Y \;=\; \big\{(x_1, x_2) \in X \times X \mid x_1 \neq x_2\big\}.$$

By definition the group action on X is doubly transitive iff the one on Y is transitive, so we want to know whether

$$\langle \chi_{V_Y}, \mathbf{1}\rangle \;=\; 1.$$

Now $X \times X \simeq X \sqcup Y$ as a G-set, hence

$$V_X \oplus V_Y \;\simeq\; V_{X\times X} \;\simeq\; V_X \otimes V_X \;\simeq\; V_X \otimes V_X^*$$

where the last fact comes from the observation that the permutation representation is isomorphic to its dual (exercise). Therefore

$$\begin{aligned}\sum_{\rho\in \mathrm{Irr}_k(G)} (m_\rho(V_X))^2 \;&=\; \langle \chi_{V_X}, \chi_{V_X}\rangle \\ &=\; \langle \chi_{V_X\otimes V_X^*}, \mathbf{1}\rangle \;=\; \langle \chi_{V_X}, \mathbf{1}\rangle + \langle \chi_{V_Y}, \mathbf{1}\rangle \;=\; 1 + \langle \chi_{V_Y}, \mathbf{1}\rangle\end{aligned}$$

and the claim follows. □

6. Integrality of characters

We know that the character of any complex representation of a finite group G takes values in the subring

$$\overline{\mathbb{Z}} \;=\; \{a \in \mathbb{C} \mid f(a) = 0 \;\text{ for some } \textit{monic} \text{ polynomial } \; f(x) \in \mathbb{Z}[x]\}$$

of algebraic integers. In fact a much stronger property holds:

Proposition 6.1. *If $\chi = \chi_V$ is the character of an irreducible $V \in \mathrm{Irr}_{\mathbb{C}}(G)$, then*

$$\frac{\chi(g)\cdot |\operatorname{Cl}(g)|}{\chi(1)} \;\in\; \overline{\mathbb{Z}} \quad \textit{for all} \quad g \;\in\; G.$$

Proof. It follows from Schur's lemma that each element $z \in Z(\mathbb{C}[G])$ of the center of the group algebra acts on the given irreducible representation by a scalar multiple $\omega_\chi(z)\cdot id_V$ of the identity. Assigning this scalar to each element we get an algebra homomorphism

$$\omega_\chi : \quad Z(\mathbb{C}[G]) \;\longrightarrow\; k,$$

the *central character* of the representation. More explicitly, the center has a basis consisting of the vectors

$$e_K \;=\; \sum_{x\in K} e_x \quad \text{for the conjugacy classes} \quad K \;\in\; \operatorname{Cl}(G),$$

and

$$\omega_\chi(e_K) \;=\; \frac{\chi(g)\cdot |\operatorname{Cl}(g)|}{\chi(1)} \quad \text{for} \quad K \;=\; \operatorname{Cl}(g).$$

So the claim amounts to the statement that the values of the central character ω_χ on each of these basis vectors are algebraic integers. For this we note that the product of any two basis vectors is again an integral linear combination of basis vectors: If $K_i = \mathrm{Cl}(g_i) \in \mathrm{Cl}(G)$ for $i = 1, \dots, n$ are the distinct conjugacy classes, then

$$e_{K_i} \cdot e_{K_j} \;=\; \sum_{\nu=1}^{n} a_{ij\nu} \cdot e_{K_\nu} \quad \text{with} \quad a_{ij\nu} \;=\; |\{(x,y) \in K_i \times K_j \mid xy = g_\nu\}| \;\in\; \mathbb{N}_0.$$

Hence the subalgebra

$$R \;=\; \langle e_K \mid K \in \mathrm{Cl}(G) \rangle \;\subset\; \mathbb{C}$$

generated by the basis vectors is a finitely generated module over $\mathbb{Z}$, i.e. the ring extension $\mathbb{Z} \subseteq R$ is *finite*. Now it is a general fact in commutative algebra that any finite ring extension is integral.

In our present case, writing $R = \mathbb{Z}r_1 + \cdots + \mathbb{Z}r_n$ for suitable $r_i \in R$, any $r \in R$ satisfies

$$r \cdot r_i \;=\; \sum_{j=1}^{n} a_{ij} \cdot r_j \quad \text{for some matrix} \quad A \;=\; (a_{ij}) \;\in\; \mathrm{Mat}_{n \times n}(\mathbb{Z}).$$

Then we obtain the eigenvalue $A \cdot v = r \cdot v$ for the vector $v = (r_1, \dots, r_n)^t \;\in\; R^n$ so that

$$\ker(A - r \cdot id) \;\neq\; \{0\},$$

whence $f(r) = 0$ for the monic polynomial $f(x) = \det(x \cdot id - A) \in \mathbb{Z}[x]$. □

Corollary 6.2. *If χ is the character of an irreducible representation, then for any $g \in G$ the complex number*

$$\alpha \;=\; \frac{\chi(g)}{\chi(1)}$$

has absolute value $|\alpha| \leq 1$. Furthermore, we have $|\alpha| \in \{0, 1\}$ iff $\alpha \in \overline{\mathbb{Z}}$.

Proof. If $n = ord(g)$, then the eigenvalues of the action of g are n-th roots of unity. In particular, it follows that $\alpha \in \mathbb{Q}(\zeta_n)$ where $\zeta_n \in \mathbb{C}$ is a primitive n-th root of unity, and

$$|\sigma(\alpha)| \;\leq\; \frac{1 + \cdots + 1}{\chi(1)} \;=\; 1 \quad \text{for all} \quad \sigma \;\in\; Gal(\mathbb{Q}(\zeta_n)/\mathbb{Q})$$

since the Galois operation permutes the roots of unity. Then the norm of α also satisfies

$$|N_{\mathbb{Q}(\zeta_n)/\mathbb{Q}}(\alpha)| \;\leq\; 1.$$

If $\alpha \in \overline{\mathbb{Z}}$, then

$$N_{\mathbb{Q}(\zeta_n)/\mathbb{Q}}(\alpha) \;\in\; \overline{\mathbb{Z}} \cap \mathbb{Q} \;=\; \mathbb{Z} \quad \text{and hence} \quad |\alpha| \;\in\; \{0, 1\}$$

by the previous estimate. Conversely, for $|\alpha| = 0$ obviously $\alpha = 0 \in \overline{\mathbb{Z}}$, whereas for $|\alpha| = 1$ all the eigenvalues of g in the given representation are the same root of unity. But then α is equal to this root of unity and hence an algebraic integer. □

Theorem 6.3. *If a finite group G contains a conjugacy class of prime power order $|\,\mathrm{Cl}(g)| = p^r > 1$, then it cannot be simple.*

Proof. After division by the prime p, the second orthogonality relation for $g \neq 1$ reads

$$\frac{1}{p} \;+\; \sum_{\chi \neq \mathbf{1}} \frac{\chi(g)\chi(1)}{p} \;=\; 0$$

where the sum runs over all non-trivial irreducible characters. In particular this gives

$$\frac{\chi(1)\chi(g)}{p} \;\notin\; \overline{\mathbb{Z}} \quad \text{for some irreducible character} \quad \chi \;\neq\; \mathbf{1}.$$

Then $p \nmid \chi(1)$, and since by assumption $|\operatorname{Cl}(g)| = p^r$ we get from Bézout's identity that

$$a \cdot |\operatorname{Cl}(g)| \; + \; b \cdot \chi(1) \; = \; 1 \quad \text{for some} \quad a, b \; \in \; \mathbb{Z}.$$

Multiplying by $\chi(g)/\chi(1)$ we get

$$a \cdot \frac{|\operatorname{Cl}(g)| \cdot \chi(1)}{\chi(g)} \; + \; b \cdot \chi(g) \; = \; \frac{\chi(g)}{\chi(1)}.$$

The left hand side is an algebraic integer by proposition 6.1. The right hand side is nonzero since $\chi(g)\chi(1)/p \notin \overline{\mathbb{Z}}$, hence by the previous corollary we get that the right hand side must have absolute value one:

$$\left| \frac{\chi(g)}{\chi(1)} \right| \; = \; 1.$$

By what we have seen in the previous section, this implies that $g \in Z(\chi)$. But the centralizer $Z(\chi) \trianglelefteq G$ is a *normal* subgroup. Hence if G is simple, then from $g \neq 1$ we get that the centralizer $Z(\chi) = G$. But again by simplicity of the group we have $\ker(\chi) \neq \{1\}$ so that the character χ is faithful. Altogether then $Z(G) = G$ and hence the group G is abelian. But then it does not contain any conjugacy class with more than one element, a contradiction. □

Corollary 6.4 (Burnside). *Any group of order $p^a q^b$ with primes p, q is solvable.*

Proof. Suppose that G is a group of order $p^a q^b$ and let $N \trianglelefteq G$ be any maximal proper normal subgroup. If $N \neq \{1\}$, we are done by induction on the order $|G|$ since

$$\{1\} \longrightarrow N \longrightarrow G \longrightarrow G/N \longrightarrow \{1\}$$

is an exact sequence and any extension of solvable groups is solvable. So $N = \{1\}$, i.e. G is simple. Take a Sylow subgroup

$$P \; \leq \; G \quad \text{with} \quad |P| \; = \; p^a > 1.$$

Any p-group has non-trivial center, so there exists $1 \neq g \in Z(P)$. Now $P \leq Z_G(g)$ because $g \in Z(P)$. But then

$$|\operatorname{Cl}(g)| \; = \; \frac{|G|}{|Z_G(g)|} \quad \text{divides} \quad \frac{|G|}{[G:P]} \; = \; q^b$$

hence $|\operatorname{Cl}(g)|$ is a prime power. Since G is simple, this forces $|\operatorname{Cl}(g)| = 1$ by the previous theorem. But then $g \in Z(G)$, which again is impossible since the center of a simple group must be trivial. □

As another nice application of the integrality result from the beginning of this section, we get the following

Theorem 6.5. *For any $V \in \operatorname{Irr}_{\mathbb{C}}(G)$, the dimension $\dim_{\mathbb{C}} V$ divides $|G|$.*

Proof. Let $\chi = \chi_V$ and let $g_1, \dots, g_n \in G$ denote a representative system for the conjugacy classes of G. Recalling that the central character ω_χ is given by the formula

$$\omega_\chi(e_{K_i}) \; = \; \frac{|K_i|\chi(g_i)}{\chi(1)} \quad \text{for} \quad K_i \; = \; \operatorname{Cl}(g_i),$$

one computes

$$|G| \; = \; \sum_{g \in G} \chi(g)\chi(g^{-1}) \; = \; \sum_{i=1}^{n} |K_i| \cdot \chi(g_i)\chi(g_i^{-1}) \; = \; \chi(1) \sum_{i=1}^{n} \omega_\chi(e_{K_i})\chi(g_i^{-1}).$$

For the right hand side we know that all the $\omega_\chi(e_{K_i})$ and $\chi(g_i^{-1})$ are algebraic integers, hence $|G| \cdot \chi(1)^{-1} \in \overline{\mathbb{Z}} \cap \mathbb{Q} = \mathbb{Z}$ and we are done. □

CHAPTER III

Representations of the symmetric group

We now specialize to symmetric groups. Here everything about the irreducible complex representations can be made completely explicit in terms of combinatorial data encoded in *Young tableaux* and *symmetric functions*. At the same time, we will later see that there is a close connection with the representation theory of the general linear group known as *Schur-Weyl-duality*.

1. Young tableaux and irreducibles

We have seen that any irreducible representation $V \in \mathrm{Irr}_{\mathbb{C}}(G)$ can be embedded as a direct summand in the left regular representation on the algebra $\mathscr{A} = \mathbb{C}[G]$, and any G-equivariant projection

$$p: \quad \mathscr{A} \twoheadrightarrow V \subset \mathscr{A}$$

must then be given by right multiplication with the element $c = p(1) \in \mathscr{A}$, which is a primitive idempotent in the following sense:

REMARK 1.1. Any element $e \in \mathscr{A} \setminus \{0\}$ with $e^2 = e$ is called *idempotent*. Such an idempotent is called *primitive* if it cannot be written as the sum $e = e_1 + e_2$ of two idempotents $e_1, e_2 \in \mathscr{A} \setminus \{0\}$ with $e_1 \cdot e_2 = 0$. Recalling that the left ideals of the group algebra are precisely its subrepresentations, we obtain a bijective correspondence

$$\left\{ \begin{matrix} \text{idempotents} \\ 0 \neq e \in \mathscr{A} \end{matrix} \right\} \longleftrightarrow \left\{ \begin{matrix} \text{left ideals} \\ \mathscr{A} \cdot e \neq 0 \end{matrix} \right\} \longleftrightarrow \left\{ \begin{matrix} \text{subrepresentations of} \\ \text{the left regular one} \end{matrix} \right\}$$

where primitive idempotents correspond to the irreducible subrepresentations. So in order to determine all the irreducible representations we "only" need to find the idempotents in the group algebra.

This is not feasible in general, but it can be carried out explicitly for symmetric groups. So put

$$G = \mathfrak{S}_d \quad \text{and} \quad \mathscr{A} = \mathbb{C}[\mathfrak{S}_d] \quad \text{for} \quad d \in \mathbb{N}.$$

To see what the primitive idempotents may look like in this case, let us begin with the following

EXAMPLE 1.2. With notations as above, there are two cases where $\mathscr{A} \cdot e = \mathbb{C} \cdot e$ has dimension one:

- For $e = \sum_{g \in G} e_g$ we get the trivial representation $\mathscr{A} \cdot e = \mathbb{C} \cdot e \simeq \mathbf{1}$.
- For $e = \sum_{g \in G} sgn(g) e_g$ we get the sign representation $\mathscr{A} \cdot e = \mathbb{C} \cdot e \simeq sgn$.

It turns out that every primitive idempotent can be obtained by combining these two constructions. For this we introduce the following notions:

Definition 1.3. Let $\lambda = (\lambda_1, \dots, \lambda_\ell)$ be a partition of d, i.e. a finite sequence of integers

$$\lambda_1 \geq \cdots \geq \lambda_\ell > 0 \quad \text{with} \quad \lambda_1 + \cdots + \lambda_n = d.$$

We also call $n = \ell(\lambda)$ the *length* of the partition and $d = \deg(\lambda)$ its *degree*. We will represent such a partition by its associated *Young diagram*, which is the left-aligned diagram consisting of

- λ_1 boxes in the top row,
- λ_2 boxes in the second row,
- $\vdots$
- λ_n boxes in the bottom row.

By a *Young tableau of shape* λ we mean an assignment of the numbers $1, \dots, \deg(\lambda)$ to the boxes of the Young diagram without repetitions. For example, the following is a Young tableau of shape $(4, 3, 1)$:

4	3	5	8
2	1	6	
7			

By abuse of notation we will often simply write λ for a Young tableau, the filling of the boxes being understood. Via the filling we attach to each row and each column a subset of $\{1, 2, \dots, d\}$, and we define the *row* resp. *column stabilizer* to be the subgroups

$$\begin{aligned} P_\lambda &= \{ g \in \mathfrak{S}_d \mid g \text{ preserves the total contents of each row } \}, \\ Q_\lambda &= \{ g \in \mathfrak{S}_d \mid g \text{ preserves the total contents of each column } \}. \end{aligned}$$

Inside $\mathscr{A} = \mathbb{C}[\mathfrak{S}_d]$ we define the *row symmetrizer* and the *column antisymmetrizer* by

$$a_\lambda = \sum_{g \in P_\lambda} e_g \quad \text{and} \quad b_\lambda = \sum_{g \in Q_\lambda} sgn(g) \cdot e_g,$$

and we define the *Young symmetrizer* to be their product $c_\lambda = a_\lambda b_\lambda \in \mathscr{A}$.

Remark 1.4. For any Young tableau of shape $\lambda = (d)$ resp. $\mu = (1, 1, \dots, 1)$ we have

$$c_{(d)} = \sum_{g \in \mathfrak{S}_d} e_g \quad \text{and} \quad c_{(1,1,\dots,1)} = \sum_{g \in \mathfrak{S}_d} sgn(g) \cdot e_g,$$

which are precisely the primitive idempotents corresponding to the trivial and the sign representation. For more general partitions, different fillings of the same Young tableau will usually lead to different Young symmetrizers, but any two such will be conjugate via a permutation $g \in \mathfrak{S}_d$. Hence for the classification of irreducible representations up to isomorphism, the choice of the filling does not matter and we fix any Young tableau of given shape in what follows.

The main ingredient for the classification of irreducible representations will be the following result; since its proof is slightly technical, the reader is invited to first consult the following theorem 1.6 and its proof for motivation.

Proposition 1.5. *Let $\mathscr{A} = \mathbb{C}[\mathfrak{S}_d]$, and fix Young tableaux of shape λ, μ.*

(1) *Up to a scalar the Young symmetrizer is the unique element $c_\lambda \in \mathscr{A}$ such that $p \cdot c_\lambda \cdot q = sgn(q) \cdot c_\lambda$ for all $p \in P_\lambda$, $q \in Q_\lambda$.*

(2) *We have*

$$c_\lambda \cdot \mathscr{A} \cdot c_\mu \;=\; \begin{cases} \mathbb{C} \cdot c_\lambda & \text{if } \lambda = \mu, \\ \{0\} & \text{if } \lambda \neq \mu. \end{cases}$$

Proof. (1) The definitions easily imply $p \cdot c_\lambda = c_\lambda$ and $c_\lambda \cdot q = sgn(q) \cdot c_\lambda$ for all $p \in P_\lambda$ and $q \in Q_\lambda$, so it only remains to show that these conditions determine the Young symmetrizer uniquely up to a scalar. Suppose that we are given $n_g \in \mathbb{C}$ such that the element

$$c \;=\; \sum_{g \in \mathfrak{S}_d} n_g \cdot e_g \;\in\; \mathscr{A}$$

also satisfies them, i.e.

$$n_{pgq} \;=\; sgn(q) \cdot n_g \quad \text{for all} \quad g \in \mathfrak{S}_d,\, p \in P_\lambda,\, q \in Q_\lambda.$$

We must see

$$n_g \;=\; 0 \quad \text{for all} \quad g \;\notin\; P_\lambda \cdot Q_\lambda.$$

This is now a bit messy:

First of all, we claim that for any group element $g \notin P_\lambda \cdot Q_\lambda$ there exist distinct indices $i \neq j$ which are simultaneously

- in the same row of the tableau λ, and
- in the same column of the tableau $\lambda' = g\lambda$ obtained from λ by applying the permutation $g \in \mathfrak{S}_d$ to the entries.

Indeed, if this were not true, all the entries of the first row of λ were in different columns of λ', hence we could find an element of Q'_λ which when applied to λ' moves all these entries to the first row. Proceeding inductively we then obtain an equality

$$p\lambda \;=\; q'\lambda' \quad \text{for suitable} \quad p \in P_\lambda \quad \text{and} \quad q' \in Q_{\lambda'} \;=\; gQg^{-1}.$$

Since a permutation is determined uniquely by what it does to a tableau, $p\lambda = q'g\lambda$ gives $p = q'g$ and so

$$g \;=\; p \cdot q \quad \text{for the element} \quad q \;=\; g^{-1} \cdot (q')^{-1} \cdot g \;\in\; Q_\lambda,$$

a contradiction. This proves our claim.

For any $i \neq j$ as in this claim, consider the transposition $p = (ij) \in P_\lambda$. We have

$$g \;=\; p \cdot g \cdot q \quad \text{for the odd permutation} \quad q \;=\; g^{-1} \cdot p \cdot g \;\in\; Q_\lambda$$

hence it follows that $n_g = n_{pgq} = sgn(q) \cdot n_g = -n_g$ and hence $n_g = 0$ as required.

(2) The containment $c_\lambda \cdot \mathscr{A} \cdot c_\lambda \subseteq \mathbb{C} \cdot c_\lambda$ follows from the uniqueness in (1) by computing

$$p \cdot c_\lambda a c_\lambda \cdot q = sgn(q) \cdot c_\lambda a c_\lambda \quad \text{for} \quad a \in \mathscr{A},\; p \in P_\lambda,\; q \in Q_\lambda.$$

To see that this containment is in fact an equality, it suffices to observe that $c_\lambda^2 \neq 0$, for which one may use that $P_\lambda \cap Q_\lambda = \{1\}$ is trivial so that any element of $P_\lambda \cdot Q_\lambda$ admits a unique representation as a product $p \cdot q$ with $p \in P_\lambda$, $q \in Q_\lambda$.

It remains to check the vanishing $c_\lambda \cdot \mathscr{A} \cdot c_\mu = \{0\}$ for all partitions $\mu \neq \lambda$. Via the anti-involution

$$\iota : \quad \mathscr{A} \;\longrightarrow\; \mathscr{A}, \quad e_g \;\mapsto\; e_{g^{-1}}$$

and similar arguments to those below, this can be reduced to the case that in the lexicographic order $\lambda > \mu$ (exercise). It suffices to show $a_\lambda \cdot a \cdot b_\mu = 0$ for all $a \in \mathscr{A}$ since one can then specialize to $a \in b_\lambda \cdot \mathscr{A} \cdot a_\lambda$. Furthermore, by linearity it will be enough to treat the case of a standard basis vector $a = e_g$ for $g \in \mathfrak{S}_d$. Using the conjugation action $e_g \cdot b_\mu \cdot e_{g^{-1}} = b_{g\mu}$, we can even assume $g = 1$ and only need to show

$$a_\lambda \cdot b_\mu \;=\; 0 \quad \text{if} \quad \lambda \;>\; \mu.$$

To check this remaining vanishing statement, we use an argument similar to the one in the first part: Under the condition $\lambda > \mu$ one can find distinct indices $i \neq j$ that lie simultaneously

- in the same row of λ and
- in the same column of μ.

We leave this as an exercise to the reader. For the transposition $g = (ij)$ one then computes

$$a_\lambda \cdot g \;=\; a_\lambda \quad \text{and} \quad g \cdot b_\mu \;=\; -b_\mu$$

which together gives

$$a_\lambda \cdot b_\mu \;=\; a_\lambda \cdot g \cdot g \cdot b_\mu \;=\; -a_\lambda \cdot b_\mu$$

and hence $a_\lambda \cdot b_\mu = 0$ as required. □

After these preparations, we can now easily list all irreducible representations of the symmetric group:

Theorem 1.6. *For any partition λ of degree $d = \deg(\lambda)$, fix a corresponding Young tableau. Then*

$$V_\lambda \;=\; \mathscr{A} \cdot c_\lambda \;\in\; \mathrm{Rep}_{\mathbb{C}}(\mathfrak{S}_d)$$

is irreducible, and every irreducible representation of the symmetric group $\mathfrak{S}_d$ is isomorphic to such a representation for a unique partition λ.

Proof. To see V_λ is irreducible, let $W \subseteq V_\lambda$ be an irreducible subrepresentation, then we have

$$c_\lambda \cdot W \;\subseteq\; c_\lambda \cdot V_\lambda \;\subseteq\; \mathbb{C} \cdot c_\lambda$$

where the second inclusion comes from the previous proposition. Since the right hand side is 1-dimensional, we either have $c_\lambda W = \mathbb{C} \cdot c_\lambda$ or $c_\lambda W = \{0\}$. In the former case it follows that

$$V_\lambda = \mathscr{A} \cdot c_\lambda \subseteq \mathscr{A} \cdot W = W$$

and we are done. In the latter case

$$W \cdot W \subseteq V_\lambda \cdot W = \mathscr{A} \cdot c_\lambda \cdot W = \mathscr{A} \cdot \{0\} = \{0\}.$$

But then, since the projection $p = \mathscr{A} \twoheadrightarrow W$ is given by right multiplication with an idempotent $e \in W$ we get

$$W = \mathrm{im}(p) = \mathrm{im}(p \circ p) \subseteq W \cdot W = \{0\}.$$

Summing up, V_λ has no non-trivial subrepresentations and is hence irreducible.

It remains to see that $V_\lambda \not\simeq V_\mu$ if $\lambda \neq \mu$ are distinct partitions. By symmetry we can assume $\lambda > \mu$ in the lexicographic ordering. But in that case the previous proposition shows

$$\begin{aligned} c_\lambda \cdot V_\mu \;&=\; c_\lambda \cdot \mathscr{A} \cdot c_\mu \;=\; \{0\}, \\ c_\lambda \cdot V_\lambda \;&=\; c_\lambda \cdot \mathscr{A} \cdot c_\lambda \;\ni\; c_\lambda^2 \;\neq\; 0, \end{aligned}$$

where the last inequality comes from the fact that c_λ is a non-trivial projector. □

Notice that the above gives an explicit bijection between conjugacy classes and irreducible representations. It also explains why all character tables of symmetric groups we have seen so far had integral entries:

COROLLARY 1.7. *Any representation of the symmetric group is defined over* $\mathbb{Q}$, *i.e. the functor* $\mathrm{Rep}_{\mathbb{Q}}(\mathfrak{S}_d) \longrightarrow \mathrm{Rep}_{\mathbb{C}}(\mathfrak{S}_d)$ *is an equivalence of categories.*

Proof. We have $V_\lambda = (\mathbb{Q}[\mathfrak{S}_d] \cdot c_\lambda) \otimes_{\mathbb{Q}} \mathbb{C}$. □

At this point, it is natural to ask for an explicit formula for the dimensions and characters of the irreducible representations of the symmetric group. Before we come to this, we first make a small detour to the representation theory of general linear groups.

2. Schur-Weyl duality

There exists a beautiful connection between the representations of symmetric groups and general linear groups, based on the commuting action of these groups on tensor powers. We formulate this over an arbitrary field k with $char(k) = 0$ since all irreducible representations of the symmetric group are defined over the rationals: Let $\mathscr{A}$ and $\mathscr{B}$ be k-algebras and U a finite-dimensional k-vector space with commuting actions

$$\begin{aligned} \lambda : \ \mathscr{A} &\longrightarrow \mathrm{End}_{\mathscr{B}}(U) \subseteq \mathrm{End}_k(U), \\ \rho : \ \mathscr{B} &\longrightarrow \mathrm{End}_{\mathscr{A}}(U) \subseteq \mathrm{End}_k(U). \end{aligned}$$

Note that since the actions commute, each factors over the centralizer of the other as indicated above. If one is the full centralizer of the other, then in the semisimple case we get a strong link between the respective isotypic decompositions: Let us call a module *absolutely irreducible* if its base change to the algebraic closure $\overline{k}$ is irreducible, then we have

PROPOSITION 2.1. *Suppose that* $U \simeq \bigoplus_{i=1}^n V_i \otimes \mathrm{Hom}_{\mathscr{A}}(V_i, U)$ *with absolutely irreducible* $V_i \in \mathrm{Mod}(\mathscr{A})$. *If*

$$\rho : \quad \mathscr{B} \twoheadrightarrow \mathrm{End}_{\mathscr{A}}(U)$$

is surjective, then the

$$W_i \ = \ \mathrm{Hom}_{\mathscr{A}}(V_i, U) \ \in \ \mathrm{Mod}(\mathscr{B})$$

are absolutely irreducible modules and they are pairwise non-isomorphic.

Proof. Our assumptions imply that $m : \mathscr{B} \twoheadrightarrow \mathrm{End}_{\mathscr{A}}(U) \simeq \bigoplus_{i=1}^n \mathrm{End}_k(W_i)$, so the maps $m_i : \mathscr{B} \twoheadrightarrow \mathrm{End}_k(W_i)$ are surjective. Therefore each module $W_i \in \mathrm{Mod}(\mathscr{B})$ is absolutely irreducible, and these modules must be pairwise non-isomorphic since otherwise m would factor over some diagonal in $\bigoplus_{i=1}^n \mathrm{End}_k(W_i)$. □

EXAMPLE 2.2. (1) Let G be a finite group. For the group algebra $\mathscr{A} = \mathscr{B} = \mathbb{C}[G]$ acting via the two regular representations on $U = \mathbb{C}[G]$, the above decomposition becomes

$$U \quad \simeq \bigoplus_{V \in \mathrm{Irr}_{\mathbb{C}}(G)} V \boxtimes V^* \ \in \ \mathrm{Rep}_{\mathbb{C}}(G \times G),$$

The dual on the right hand side is needed to make the splitting canonical: Think of $V \boxtimes V^* = \mathrm{Hom}_k(V, V)$, then the right resp. left regular representation gives the action on morphisms by pre- resp. postcomposition.

(2) Let V be a finite-dimensional vector space. On $U = V^{\otimes d}$ for $d \in \mathbb{N}$, we have two commuting actions of $\mathfrak{S}_d$ and $Gl(V)$, so we may apply the above formalism to the algebras

$$\mathscr{A} \ = \ k[\mathfrak{S}_d] \quad \text{and} \quad \mathscr{B} \ = \ \mathrm{Span}_k\big(g \otimes \cdots \otimes g \mid g \in Gl(V)\big) \ \subseteq \ \mathrm{End}_k(U).$$

Let us check that the centralizer condition in the above proposition holds:

THEOREM 2.3. *For any $d \in \mathbb{N}$, the image of $Gl(V)$ spans all of $\mathrm{End}_{\mathfrak{S}_d}(V^{\otimes d})$.*

Proof. Via the natural identification $\mathrm{End}_k(V^{\otimes d}) \simeq E^{\otimes d}$ for $E = \mathrm{End}_k(V)$ we have

$$\mathrm{End}_{\mathfrak{S}_d}(V^{\otimes d}) \;\simeq\; (E^{\otimes d})^{\mathfrak{S}} \;\simeq\; \mathrm{Sym}^d(E).$$

But an exercise in multilinear algebra shows that the symmetric powers $\mathrm{Sym}^d(W)$ of any vector space W are spanned by tensors $w \otimes \cdots \otimes w$ with $w \in W$. In the case at hand, it follows that

$$\mathrm{End}_{\mathfrak{S}_d}(V^{\otimes d}) \;=\; \mathrm{Span}_k\big(f \otimes \cdots \otimes f \mid f \in \mathrm{End}_k(V)\big) \;\subseteq\; \mathrm{End}_k(V^{\otimes d}).$$

Since

$$Gl(V) \;\subset\; \mathrm{End}_k(V)$$

is dense in the Zariski topology or alternatively in the classical one for $k \subseteq \mathbb{C}$, the span $\mathrm{Span}_k(g \otimes \cdots \otimes g \mid g \in Gl(V)) \subseteq \mathrm{End}_{\mathfrak{S}_d}(V^{\otimes d})$ is a dense subspace, hence the whole space since the target is finite-dimensional. □

COROLLARY 2.4 (Schur-Weyl duality). *For any vector space V with $n = \dim_k V$ and $d \in \mathbb{N}$ one has a natural decomposition*

$$V^{\otimes d} \;\simeq\; \bigoplus_{\deg(\lambda)=d} V_\lambda \boxtimes S_\lambda(V) \;\in\; \mathrm{Rep}_k(\mathfrak{S}_d \times Gl(V))$$

where the

$$S_\lambda(V) \;=\; \mathrm{Hom}_{\mathfrak{S}_n}(V_\lambda, V^{\otimes d}) \;\in\; \mathrm{Rep}_k(Gl(V))$$

are absolutely irreducible or zero, the non-zero ones being pairwise non-isomorphic.

Proof. Since the irreducible representations of the symmetric group are precisely the $V_\lambda \in \mathrm{Rep}_k(\mathfrak{S}_d)$ with $\deg(\lambda) = d$, the previous proposition and lemma apply. □

REMARK 2.5. The above construction is natural in the vector space V and we call $S_\lambda : \mathrm{Vect}(k) \longrightarrow \mathrm{Vect}(k)$, $V \mapsto S_\lambda(V)$ a *Schur functor.* These Schur functors generalize the symmetric and alternating powers: The direct summands in the corollary are precisely the isotypic pieces of $V^{\otimes d} \in \mathrm{Rep}_k(\mathfrak{S}_d)$, so for the trivial and sign representations we get

$$S_{(d)}(V) \;\simeq\; \mathrm{Sym}^d(V) \quad \text{and} \quad S_{(1,\dots,1)}(V) \;\simeq\; \mathrm{Alt}^d(V).$$

Notice that the latter vanishes for $d > n$. More generally we have:

LEMMA 2.6. *Let λ be a partition of d. Then for any Young tableau of shape λ the image of the multiplication by the Young symmetrizer c_λ on $V^{\otimes d}$ is isomorphic to $S_\lambda(V)$,*

$$S_\lambda(V) \;\simeq\; c_\lambda \cdot V^{\otimes d} \;\subseteq\; V^{\otimes d} \quad \textit{in} \quad \mathrm{Rep}_k(Gl(V)).$$

Proof. Let $\mathscr{A} = k[\mathfrak{S}_d]$. Recalling that $V_\lambda = \mathscr{A} \cdot c_\lambda$, the desired isomorphism is given by

$$\begin{array}{rcl} S_\lambda(V) \;=\; \mathrm{Hom}_{\mathfrak{S}_d}(V_\lambda, V^{\otimes d}) & \stackrel{\sim}{\longrightarrow} & c_\lambda \cdot V^{\otimes d} \;\subseteq\; V^{\otimes d} \\ f & \mapsto & f(c_\lambda) \end{array}$$

Indeed, we know by the previous section that the symmetrizer satisfies $c_\lambda^2 = c \cdot c_\lambda$ for some $c \in \mathbb{Q}^\times$ and so

$$f(c_\lambda) \;=\; \tfrac{1}{c} \cdot c_\lambda \cdot f(c_\lambda) \;\in\; c_\lambda \cdot V^{\otimes d} \quad \text{for any} \quad f \;\in\; \mathrm{Hom}_{\mathscr{A}}(\mathscr{A} c_\lambda, V^{\otimes d});$$

the same argument furthermore shows that $\mathrm{Ann}_{\mathscr{A}}(c_\lambda \cdot V^{\otimes d}) = \mathrm{Ann}_{\mathscr{A}}(c_\lambda)$ and hence the claim follows. □

This in particular allows to generalize the previous vanishing statement for the alternating powers to a precise vanishing criterion for Schur functors:

COROLLARY 2.7. *We have* $S_\lambda(V) = \{0\}$ *iff* $\ell(\lambda) > \dim_k(V)$.

Proof. For simplicity of notation, we fix the Young tableau of shape λ whose boxes are filled column by column in a monotonically increasing way, such as in the following example:

1	4	6	8
2	5	7	
3			

The reason for taking this tableau is that in terms of the transpose $\mu = (\mu_1, \dots, \mu_m)$ of λ, the image of the column antisymmetrizer

$$b_\lambda \cdot V^{\otimes d} \;=\; \mathrm{Alt}^{\mu_1}(V) \otimes \cdots \otimes \mathrm{Alt}^{\mu_m}(V) \;\subseteq\; V^{\otimes \mu_1} \otimes \cdots \otimes V^{\otimes \mu_m} \;=\; V^{\otimes d}$$

is embedded in the standard way. Clearly

$$\mu_1 \;=\; \ell(\lambda) \;>\; \dim_k(V) \iff b_\lambda \cdot V^{\otimes d} \;=\; \{0\} \implies c_\lambda \cdot V^{\otimes d} \;=\; \{0\},$$

so it only remains to check that for the implication on the right the converse holds as well. Fixing an arbitrary basis $v_1, \dots, v_n \in V$ we consider the basis of $V^{\otimes d}$ given by

$$\mathscr{B} \;=\; \big\{ v_{i_1} \otimes \cdots \otimes v_{i_d} \mid i_1, \dots, i_d \in \{1, 2, \dots, n\} \big\}.$$

One computes by our choice of the tableau that for the vector

$$v \;=\; (v_1 \otimes \cdots \otimes v_{\mu_1}) \otimes (v_1 \otimes \cdots \otimes v_{\mu_2}) \otimes \cdots \otimes (v_1 \otimes \cdots \otimes v_{\mu_m}),$$

one has

$$c_\lambda \cdot v \;\neq\; 0$$

because the coefficient of the basis vector v in $c_\lambda \cdot v$ is strictly positive. □

REMARK 2.8. From a conceptional point of view, theorem 2.3 can be upgraded to the statement that for $U = V^{\otimes d}$ the images

$$\mathscr{A} \;=\; \mathrm{im}\Big(k[\mathfrak{S}_d] \;\longrightarrow\; \mathrm{End}_k(U) \Big) \quad \text{and} \quad \mathscr{B} \;=\; \mathrm{Span}_k\big(g \otimes \cdots \otimes g \mid g \in Gl(V) \big)$$

are the full centralizers of each other in the sense that

$$\mathscr{B} \;=\; \mathrm{End}_{\mathscr{A}}(U) \quad \text{and} \quad \mathscr{A} \;=\; \mathrm{End}_{\mathscr{B}}(U).$$

Indeed, the second of these two equalities is a formal consequence of the first one via the following more symmetric reformulation of proposition 2.1. Here U can be any vector space with $\dim_k U < \infty$. For simplicity we assume the base field to be algebraically closed:

THEOREM 2.9 (Double centralizer). *Suppose that k is algebraically closed, and let $\mathscr{A} \subseteq \mathrm{End}_k(U)$ be a semisimple subalgebra. Then the centralizer $\mathscr{B} = \mathrm{End}_{\mathscr{A}}(U)$ is a semisimple algebra with $\mathscr{A} = \mathrm{End}_{\mathscr{B}}(U)$ and*

$$U \quad \simeq \bigoplus_{V_\lambda \in \mathrm{Irr}_k(\mathscr{A})} V_\lambda \otimes W_\lambda \;\in\; \mathrm{Mod}(\mathscr{A} \otimes \mathscr{B}),$$

where the $W_\lambda = \mathrm{Hom}_{\mathscr{A}}(V_\lambda, U) \in \mathrm{Mod}(\mathscr{B})$ *are irreducible, pairwise non-isomorphic, and form a complete representative set for the isomorphism classes in* $\mathrm{Irr}_k(\mathscr{B})$*. So we have a bijection*

$$\mathrm{Irr}_k(\mathscr{A}) \;\simeq\; \mathrm{Irr}_k(\mathscr{B}), \quad V_\lambda \;\mapsto\; W_\lambda.$$

Proof. From the theory of semisimple algebras and their modules in chapter I we know

$$\mathscr{A} \quad\simeq \bigoplus_{V_\lambda \in \mathrm{Irr}_k(\mathscr{A})} \mathrm{End}_k(V_\lambda)$$

and

$$U \quad\simeq \bigoplus_{V_\lambda \in \mathrm{Irr}_k(\mathscr{A})} V_\lambda \otimes W_\lambda \quad \text{for} \quad W_\lambda \;=\; \mathrm{Hom}_{\mathscr{A}}(V_\lambda, U).$$

For the centralizer of the algebra it follows that $\mathscr{B} = \mathrm{End}_{\mathscr{A}}(U) \simeq \bigoplus_\lambda \mathrm{End}_k(W_\lambda)$, which implies the result. □

3. The representation ring

For a finite group G we define its *character group* $R(G)$ to be the free abelian group generated by the irreducible complex representations of G. Since up to isomorphism any representation is determined by its character, we can equivalently define

$$R(G) \;=\; \big\langle\, \chi_V \mid V \in \mathrm{Irr}_{\mathbb{C}}(G) \,\big\rangle \;\subseteq\; \mathscr{C}(G)$$

to be the subgroup of class functions generated by characters. Its elements are the linear combinations

$$\sum_{V \in \mathrm{Irr}_k(G)} m_V \cdot \chi_V \;\in\; \mathscr{C}(G) \quad \text{with} \quad m_V \in \mathbb{Z}$$

and are called *virtual characters* as opposed to the characters with $m_V \geq 0$.

As a shorthand we also write $[V] \in R(G)$ for the character of $V \in \mathrm{Rep}_{\mathbb{C}}(G)$. For example, the induction from an arbitrary subgroup $H \leq G$ gives rise to a group homomorphism

$$\mathrm{Ind}_H^G : \quad R(H) \;\longrightarrow\; R(G), \quad [V] \;\mapsto\; [\mathrm{Ind}_H^G(V)].$$

In the sequel we combine the character groups of all symmetric groups in the graded abelian group

$$R \;=\; \bigoplus_{d \in \mathbb{N}} R(d) \quad \text{with graded pieces} \quad R(d) \;=\; R(\mathfrak{S}_d).$$

Lemma 3.1. *This* R *is a commutative graded ring with multiplication* $\circ$ *defined on homogenous pieces by*

$$\circ: \quad R(d) \times R(e) \;\longrightarrow\; R(d+e), \quad U \circ V \;=\; \mathrm{Ind}_{\mathfrak{S}_d \times \mathfrak{S}_e}^{\mathfrak{S}_{d+e}}(U \boxtimes V).$$

Proof. Associativity follows from the transitivity of induction in steps, indeed we have

$$\begin{aligned} \mathrm{Ind}_{\mathfrak{S}_{d+e} \times \mathfrak{S}_f}^{\mathfrak{S}_{d+e+f}} \circ \mathrm{Ind}_{\mathfrak{S}_d \times \mathfrak{S}_e \times \mathfrak{S}_f}^{\mathfrak{S}_{d+e} \times \mathfrak{S}_f} \;&\simeq\; \mathrm{Ind}_{\mathfrak{S}_d \times \mathfrak{S}_e \times \mathfrak{S}_f}^{\mathfrak{S}_{d+e+f}} \\ &\simeq\; \mathrm{Ind}_{\mathfrak{S}_d \times \mathfrak{S}_{e+f}}^{\mathfrak{S}_{d+e+f}} \circ \mathrm{Ind}_{\mathfrak{S}_d \times \mathfrak{S}_e \times \mathfrak{S}_f}^{\mathfrak{S}_d \times \mathfrak{S}_{e+f}} \end{aligned}$$

for any $d, e, f \in \mathbb{N}$. Commutativity says that

$$\mathrm{Ind}_{\mathfrak{S}_d \times \mathfrak{S}_e}^{\mathfrak{S}_{d+e}}(U \boxtimes V) \;\simeq\; \mathrm{Ind}_{\mathfrak{S}_e \times \mathfrak{S}_d}^{\mathfrak{S}_{e+d}}(V \boxtimes U)$$

since the two subgroups from which we induce here are conjugate to each other. □

It turns out that the above ring has a very concrete description in terms of symmetric polynomials. Let

$$S(x_1, \dots, x_n) \;=\; \Big(\mathbb{Z}[x_1, \dots, x_n]\Big)^{\mathfrak{S}_n}$$

denote the ring of symmetric polynomials in n variables. Note that this is a graded ring, the grading being given by the total degree of monomials.

Proposition 3.2. *Let V be a complex vector space of dimension n. Then for any partition λ of degree $d = \deg(\lambda)$ there exists a unique homogenous symmetric polynomial*

$$s_\lambda \;\in\; S(x_1, \dots, x_n)$$

of degree d such that

$$tr\Big(\; S_\lambda(V) \xrightarrow{\;S_\lambda(g)\;} S_\lambda(V) \;\Big) \;=\; s_\lambda(t_1, \dots, t_n)$$

for every $g \in Gl(V)$ with eigenvalues $t_1, \dots, t_n \in \mathbb{C}$.

Proof. It suffices to show the existence of a polynomial s_λ with this property for all g in an open dense subset of $Gl(V)$, so we only need to consider diagonalizable matrices. Conjugate matrices have the same trace, so fixing a basis $v_1, \dots, v_n \in V$ we may assume

$$g \;=\; diag(t_1, \dots, t_n)$$

is a diagonal matrix. Furthermore, we know by lemma 2.6 that $S_\lambda(V) = c_\lambda \cdot V^{\otimes d}$ is spanned by the vectors

$$c_\lambda \cdot v_{i_1} \otimes \cdots \otimes v_{i_d} \quad \text{with} \quad i_1, \dots, i_d \;\in\; \{1, 2, \dots, n\}.$$

Each of these vectors is an eigenvector for the action of the element $g \in Gl(V)$, and the corresponding eigenvalue

$$t_{i_1} \cdots t_{i_d}$$

is a polynomial in $t_1, \dots, t_n$. Of course the above eigenvectors will in general not be linearly independent, but any maximal linearly independent subset of them will form a basis of $S_\lambda(V)$. Hence there exists a polynomial $s_\lambda \in \mathbb{Z}[x_1, \dots, x_n]$ such that $tr(S_\lambda(g)) = s_\lambda(t_1, \dots, t_n)$, and this polynomial is necessarily symmetric since any permutation of the eigenvalues can be obtained via conjugation with the corresponding permutation matrix. □

Remark 3.3. An exercise in algebra shows that the ring of symmetric functions is a polynomial ring

$$S(x_1, \dots, x_n) \;=\; \mathbb{Z}[e_1, \dots, e_n] \;=\; \mathbb{Z}[h_1, \dots, h_n]$$

freely generated by either the *elementary* or the *complete symmetric polynomials* given by

$$e_\nu(x_1, \dots, x_n) \;=\; \sum_{i_1 < \cdots < i_\nu} x_{i_1} \cdots x_{i_\nu},$$
$$h_\nu(x_1, \dots, x_n) \;=\; \sum_{i_1 \le \cdots \le i_\nu} x_{i_1} \cdots x_{i_\nu}.$$

For example, for $n = 2$ one has $e_1 = h_1 = x_1 + x_2$, $e_2 = x_1 x_2$, $h_2 = x_1^2 + x_1 x_2 + x_2^2$ so that

$$x_1^3 + x_2^3 \;=\; (x_1 + x_2)^3 - 3x_1 x_2^2 - 3x_1^2 x_2 \;=\; e_1^3 - 3e_1 e_2$$

etc. In general, the homogenous part

$$S(x_1, \dots, x_n)(d) \;\subset\; S(x_1, \dots, x_n)$$

of given degree $d \in \mathbb{N}$ is the free abelian group generated by either the symmetric polynomials

$$e_\lambda(x_1, \dots, x_n) \; = \; \prod_{i=1}^{n} e_{\lambda_i}(x_1, \dots, x_n)$$

or the symmetric polynomials

$$h_\lambda(x_1, \dots, x_n) \; = \; \prod_{i=1}^{n} h_{\lambda_i}(x_1, \dots, x_n)$$

where $\lambda = (\lambda_1, \dots, \lambda_\ell)$ runs over all partitions of degree d with $\lambda_1 \leq n$. Its rank is given by

$$\operatorname{rank}\bigl(S(x_1, \dots, x_n)(d)\bigr) \; = \; \#\bigl\{\text{partitions of degree } d\bigr\} \quad \text{if} \quad d \; \leq \; n.$$

Coming back to the result of proposition 3.2, the polynomials $s_\lambda(x_1, \dots, x_n)$ are called *Schur polynomials*. Like characters, they can be considered as footprints of the irreducible representations of symmetric groups:

Theorem 3.4. *Sending each irreducible representation to its Schur polynomial we get an epimorphism*

$$ch_n : \quad R \twoheadrightarrow S(x_1, \dots, x_n), \quad [V_\lambda] \mapsto s_\lambda(x_1, \dots, x_n)$$

of graded rings. Furthermore, this is an isomorphism in all degrees $d \leq n$.

Proof. Since the irreducible representations V_λ freely generate R as an abelian group, it is clear that there exists a unique group homomorphism sending each $[V_\lambda]$ to the Schur polynomial $s_\lambda(x_1, \dots, x_n)$. Since the latter is homogenous of degree $\deg(\lambda)$, it is also clear that this homomorphism preserves the grading. To see that it is a *ring* homomorphism, recall that the multiplication on the ring R is defined by

$$V_\lambda \circ V_\mu \; = \; \operatorname{Ind}_{\mathfrak{S}_d \times \mathfrak{S}_e}^{\mathfrak{S}_{d+e}}(V_\lambda \boxtimes V_\mu) \quad \text{for} \quad d \; = \; \deg(\lambda), \; e \; = \; \deg(\mu).$$

Using Frobenius reciprocity we therefore get for any complex vector space V the following chain of $Gl(V)$-equivariant isomorphisms:

$$\begin{aligned} \operatorname{Hom}_{\mathfrak{S}_{d+e}}(V_\lambda \circ V_\mu, V^{\otimes(d+e)}) \; &\simeq \; \operatorname{Hom}_{\mathfrak{S}_d \times \mathfrak{S}_e}(V_\lambda \boxtimes V_\mu, V^{\otimes d} \boxtimes V^{\otimes e}) \\ &\simeq \; \operatorname{Hom}_{\mathfrak{S}_d}(V_\lambda, V^{\otimes d}) \otimes \operatorname{Hom}_{\mathfrak{S}_e}(V_\mu, V^{\otimes e}) \\ &\simeq \; S_\lambda(V) \otimes S_\mu(V) \end{aligned}$$

Unravelling the definitions, we obtain after taking traces of $g \in Gl(V)$ on both sides that

$$ch_n(V_\lambda \circ V_\mu) \; = \; ch_n(V_\lambda) \cdot ch_n(V_\mu) \quad \text{for} \quad n \; = \; \dim_{\mathbb{C}} V$$

which shows that ch_n is a ring homomorphism. It is then automatically surjective since the image contains the elementary symmetric functions

$$\begin{aligned} e_\nu(x_1, \dots, x_n) \; &= \; tr\Bigl(diag(x_1, \dots, x_n) \; \Big| \; \operatorname{Alt}^\nu(V)\Bigr) \\ &= \; tr\Bigl(diag(x_1, \dots, x_n) \; \Big| \; S_{(1,\dots,1)}(V)\Bigr) \; = \; ch_n\bigl(sgn\bigr), \end{aligned}$$

which by the previous remark generate the ring of all symmetric functions. Finally, in degrees $d \leq n$ the epimomorphism ch_n is also injective since in these degrees the source and the target are free abelian groups of the same rank, this rank being the number of all partitions of degree d. □

REMARK 3.5. For $d \in \mathbb{N}$, let $\boldsymbol{e}_d = [sgn], \boldsymbol{h}_d = [\mathbf{1}] \in R(d)$ denote the classes of the sign respectively the trivial representation of the symmetric group $\mathfrak{S}_d$. The above argument shows

$$ch_n(\boldsymbol{e}_d) \;=\; e_d(x_1, \dots, x_n) \quad \text{and} \quad ch_n(\boldsymbol{h}_d) \;=\; h_d(x_1, \dots, x_n)$$

for all $n \in \mathbb{N}$. For $d > n$ one has $e_d(x_1, \dots, x_n) = 0$ and then $h_d(x_1, \dots, x_n)$ is a polynomial in lower degree complete symmetric polynomials, which explains the restriction on the degrees. We can get rid of this by passing to polynomials in infinitely many variables:

COROLLARY 3.6. *The ring R is a polynomial ring in either of the above infinite sets of variables:*

$$R \;=\; \mathbb{Z}[\boldsymbol{h}_1, \boldsymbol{h}_2, \dots] \;=\; \mathbb{Z}[\boldsymbol{e}_1, \boldsymbol{e}_2, \dots].$$

Proof. To see that the representation ring is generated by the elements $\boldsymbol{e}_i$, it suffices to show

$$[V] \;\in\; \mathbb{Z}[\boldsymbol{e}_1, \boldsymbol{e}_2, \dots] \quad \text{for every} \quad V \;\in\; \mathrm{Rep}_{\mathbb{C}}(\mathfrak{S}_d) \quad \text{and} \quad d \;\in\; \mathbb{N}.$$

For this we fix $n > d$. Since every symmetric polynomial is a polynomial in the elementary symmetric ones,

$$ch_n([V]) \;=\; f(e_1, \dots, e_n) \quad \text{for some} \quad f \;\in\; \mathbb{Z}[y_1, \dots, y_n].$$

But then

$$[V] \;=\; f(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$$

by the injectivity of ch_n in degrees $d < n$ as required. It remains to see that the generators $\boldsymbol{e}_i \in R$ are algebraically independent. If not, we could find a relation between finitely many of them, say

$$f(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m) \;=\; 0 \quad \text{for some} \quad f \;\in\; \mathbb{Z}[y_1, \dots, y_m] \setminus \{0\}.$$

Applying ch_n we get

$$f(e_1(x_1, \dots, x_n), \dots, e_m(x_1, \dots, x_n)) \;=\; ch_n(f(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)) \;=\; 0$$

in the ring of symmetric functions in n variables. For $n \geq m$ the algebraic independence of the elementary symmetric polynomials $e_1, \dots, e_m$ in this ring implies that $f = 0$ and we are done. □

The above can be reformulated conveniently as follows. For any $n \in \mathbb{N}$ we have restriction homomorphisms

$$\begin{aligned} r_n: \quad S(x_1, \dots, x_{n+1}) &\twoheadrightarrow S(x_1, \dots, x_n), \\ f(x_1, \dots, x_{n+1}) &\mapsto f(x_1, \dots, x_n, 0). \end{aligned}$$

We define the ring of symmetric polynomials in infinitely many variables as the inverse limit

$$S \;=\; S(x_1, x_2, \dots) \;=\; \varprojlim_n S(x_1, \dots, x_n),$$

i.e.

$$S \;=\; \left\{ (f_n)_{n \in \mathbb{N}} \mid f_n \in S(x_1, \dots, x_n) \text{ and } r_n(f_{n+1}) = f_n \text{ for all } n \in \mathbb{N} \right\}$$

is the subring consisting of all elements of the product ring $\prod_{n \in \mathbb{N}} S(x_1, \dots, x_n)$ that are compatible with the restriction maps.

Example 3.7. Since $r_n(e_\nu(x_1, \dots, x_{n+1})) = e_\nu(x_1, \dots, x_n)$ for any $\nu \in \mathbb{N}$, we may define the elementary symmetric polynomials in infinitely many variables by the formula

$$e_\nu \;=\; \big(e_\nu(x_1, \dots, x_n)\big)_{n \in \mathbb{N}} \;\in\; S$$

and obtain that

$$S \;\simeq\; \varprojlim_n \mathbb{Z}[e_1, \dots, e_n] \;\simeq\; \mathbb{Z}[e_1, e_2, \dots]$$

is a polynomial ring in this infinite set of variables. For $h_\nu = \big(h_\nu(x_1, \dots, x_n)\big)_{n \in \mathbb{N}}$ we similarly have

$$S \;\simeq\; \mathbb{Z}[h_1, h_2, \dots].$$

We can summarize the above discussion by the following

Corollary 3.8. *We have an isomorphism of graded rings* $ch : R \xrightarrow{\sim} S$.

Proof. By the previous results it only remains to show that $r_n \circ ch_{n+1} = ch_n$ for all $n \in \mathbb{N}$. Both maps send $\boldsymbol{e}_\nu \mapsto e_\nu(x_1, \dots, x_n)$ which by convention is taken to be zero if $\nu > n$, hence the claim follows. □

4. Schur polynomials and character values

In the previous section we have attached to each $V \in \mathrm{Rep}_{\mathbb{C}}(\mathfrak{S}_d)$ a homogenous polynomial

$$ch([V]) \;\in\; S \;=\; S(x_1, x_2, \dots)$$

of degree d, and we have seen that this polynomial determines the representation uniquely up to isomorphism. In order to compute the values of the character χ_V from this polynomial, we need to understand what the dashed arrow at the bottom of the diagram

$$\begin{array}{ccccc} R(d) & \hookrightarrow & \mathscr{C}(\mathfrak{S}_d) & \xrightarrow{\chi \mapsto \chi(g)} & \mathbb{C} \\ \big\downarrow{\scriptstyle ch} & & & & \big\| \\ S(d) & & \overset{???}{\cdots\cdots\cdots\cdots\rightarrow} & & \mathbb{C} \end{array}$$

for $g \in \mathfrak{S}_d$ looks like in terms of homogenous symmetric polynomials. For this we first give another description of the upper row: Recall that the classes $\mathrm{Cl}(g)$ are parametrized by cycle types, partitions of degree d, and if we denote by $C_\lambda \in \mathrm{Cl}(\mathfrak{S}_d)$ the class of cycle type $\lambda = (\lambda_1, \lambda_2, \dots)$, then the centralizer of each of its elements has order

$$z_\lambda \;=\; \frac{d!}{|C_\lambda|} \;=\; \prod_{a \geq 1} m_a(\lambda)! \cdot a^{m_a(\lambda)}$$

where $m_a(\lambda) = \#\{i \mid \lambda_i = a\}$.

Lemma 4.1. *With notations as above, define a class function* $\boldsymbol{p}_\lambda \in \mathscr{C}(\mathfrak{S}_d)$ *by the formula*

$$\boldsymbol{p}_\lambda(x) \;=\; \begin{cases} z_\lambda & \textit{if } x \in C_\lambda, \\ 0 & \textit{otherwise.} \end{cases}$$

Then $\chi(g) = \langle \chi, \boldsymbol{p}_\lambda \rangle$ *for any* $g \in C_\lambda$ *and all* $\chi \in \mathscr{C}(\mathfrak{S}_d)$. *In particular,* $\boldsymbol{p}_\lambda \in R(d)$.

Proof. The first equality is immediate from the definition of the scalar product on class functions since $\langle \chi, \boldsymbol{p}_\lambda \rangle = \frac{1}{d!} \sum_{x \in \mathfrak{S}_d} \chi(x) \overline{\boldsymbol{p}_\lambda(x)} = \frac{1}{d!} \cdot |C_\lambda| \cdot \chi(g) \cdot \boldsymbol{p}_\lambda(g)$. For the second claim, note that

$$\boldsymbol{p}_\lambda \;=\; \sum_\chi \langle \boldsymbol{p}_\lambda, \chi \rangle \cdot \chi$$

where χ runs over the characters of all irreducible representations of $\mathfrak{S}_d$. Since for symmetric groups all character values are integral, it follows that $\boldsymbol{p}_\lambda$ is an integral linear combination of characters. □

Let $\boldsymbol{s}_\lambda \in R(d)$ be the character of the irreducible representation $V_\lambda \in \mathrm{Rep}_\mathbb{C}(\mathfrak{S}_d)$, and consider the polynomials

$$\begin{aligned} s_\lambda \;&=\; ch(\boldsymbol{s}_\lambda) \;\in\; S(d), \\ p_\lambda \;&=\; ch(\boldsymbol{p}_\lambda) \;\in\; S(d). \end{aligned}$$

The first of these are the Schur polynomials, the second will be identified explicitly below. Notice that both form a basis of the vector space $S(d) \otimes_\mathbb{Z} \mathbb{C}$ of degree d symmetric functions since

$$\langle \boldsymbol{s}_\lambda, \boldsymbol{s}_\mu \rangle \;=\; \begin{cases} 1 \\ 0 \end{cases} \qquad \text{and} \quad \langle \boldsymbol{p}_\lambda, \boldsymbol{p}_\mu \rangle \;=\; \begin{cases} z_\lambda & \text{if } \lambda = \mu, \\ 0 & \text{otherwise.} \end{cases}$$

So the problem of finding character values reduces to a base change between these two sets of polynomials:

Corollary 4.2. *Write each symmetric polynomial p_λ as a linear combination of Schur polynomials*

$$p_\lambda \;=\; \sum_\mu c_{\lambda\mu} \cdot s_\mu,$$

then each $c_{\lambda\mu}$ is the value of the character $\boldsymbol{s}_\mu = \chi_{V_\mu}$ on the conjugacy class C_λ.

Proof. Since the $\boldsymbol{s}_\mu$ are orthonormal, this follows from the previous lemma. □

The combinatorial complexity of the problem lies in the Schur polynomials, the polynomials p_λ are very easy to describe explicitly although the proof takes a bit of effort:

Theorem 4.3. *For any partition $\lambda = (\lambda_1, \dots, \lambda_\ell)$ the polynomial $p_\lambda(x)$ is a product*

$$p_\lambda(x) \;=\; p_{\lambda_1}(x) \cdots p_{\lambda_\ell}(x)$$

where the factors are power sums:

$$p_r(x) \;=\; x_1^r + x_2^r + \cdots \;\in\; S(r) \quad \textit{for} \quad r \;\in\; \mathbb{N}.$$

Proof. We will first reduce this to the case of singleton partitions, for which we will proceed by showing that both sides can be computed via the same recursion relation. More precisely we claim

(1) $p_\lambda = p_{\lambda_1} \cdots p_{\lambda_\ell}$,
(2) $n \cdot h_n = \sum_{r=1}^n p_r \cdot h_{n-r}$ for any $n \in \mathbb{N}$,
(3) $n \cdot h_n = \sum_{r=1}^n q_r \cdot h_{n-r}$ for any $n \in \mathbb{N}$,

where in the last part (3) we temporarily denote by $q_r(x) = x_1^r + x_2^r + \cdots \in S(r)$ the power sums, which by comparison with (2) will a posteriori be equal to the symmetric polynomials $p_r(x) = ch(\boldsymbol{p}_r)$.

(1) We prove the corresponding identity in the representation ring. By definition the class function

$$\boldsymbol{p}_{\lambda_1} \circ \cdots \circ \boldsymbol{p}_{\lambda_\ell} \;=\; \mathrm{Ind}_{\mathfrak{S}_{\lambda_1} \times \cdots \times \mathfrak{S}_{\lambda_\ell}}^{\mathfrak{S}_d} (\boldsymbol{p}_{\lambda_1} \boxtimes \cdots \boxtimes \boldsymbol{p}_{\lambda_\ell})$$

is induced from

$$(\boldsymbol{p}_{\lambda_1} \boxtimes \cdots \boxtimes \boldsymbol{p}_{\lambda_\ell})(g_1, \dots, g_\ell) \;=\; \begin{cases} \lambda_1 \cdots \lambda_\ell & \text{if each } g_i \in \mathfrak{S}_{\lambda_i} \text{ is a } \lambda_i\text{-cycle}, \\ 0 & \text{otherwise}, \end{cases}$$

on the subgroup

$$H \;=\; \mathfrak{S}_\lambda = \mathfrak{S}_{\lambda_1} \times \cdots \times \mathfrak{S}_{\lambda_\ell} \;\leq\; G \;=\; \mathfrak{S}_d.$$

Now a general formula for induced characters says

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

for any character χ and any $g \in G$, as one may check as an exercise in Frobenius reciprocity. In our setup it follows that $(\boldsymbol{p}_{\lambda_1} \circ \cdots \circ \boldsymbol{p}_{\lambda_\ell})(g)$ vanishes unless $g \in C_\lambda$, in which case

$$(\boldsymbol{p}_{\lambda_1} \circ \cdots \circ \boldsymbol{p}_{\lambda_\ell})(g) \;=\; \frac{1}{|\mathfrak{S}_\lambda|} \cdot \#\bigl\{x \in \mathfrak{S}_d \mid xgx^{-1} \in \mathfrak{S}_\lambda\bigr\} \cdot \prod_i \lambda_i.$$

To compute this character value, note that the set in the middle is a union of right cosets $\mathfrak{S}_\lambda \cdot y$ where $y \in \mathfrak{S}_d$ permutes the cycles in g of the same length. The total number of such permutations is $\prod_{a \geq 1} m_a(\lambda)!$ in the notations from the beginning of this section, so

$$(\boldsymbol{p}_{\lambda_1} \circ \cdots \circ \boldsymbol{p}_{\lambda_\ell})(g) \;=\; \prod_{a \geq 1} m_a(\lambda)! \cdot a^{m_a(\lambda)} \;=\; z_\lambda \;=\; \boldsymbol{p}_\lambda(g) \quad \text{for} \quad g \;\in\; C_\lambda$$

and hence the claimed identity (1) follows.

(2) Again we prove the corresponding identity $\sum_{r=1}^n \boldsymbol{p}_r \circ \boldsymbol{h}_{n-r} = n \boldsymbol{h}_n$ of class functions. Here

$$\boldsymbol{p}_r \circ \boldsymbol{h}_{n-r} \;=\; \mathrm{Ind}_H^G(\boldsymbol{p}_r \boxtimes \mathbf{1})$$

for the subgroup

$$H \;=\; \mathfrak{S}_{r,n-r} \;=\; \mathfrak{S}_r \times \mathfrak{S}_{n-r} \;\leq\; G \;=\; \mathfrak{S}_n,$$

and for $h \in \mathfrak{S}_r$ we have

$$\boldsymbol{p}_r(h) \;=\; \begin{cases} r & \text{if } h \in C_r \text{ is an } r\text{-cycle}, \\ 0 & \text{otherwise}. \end{cases}$$

So the formula for the induced character gives

$$(\boldsymbol{p}_r \circ \boldsymbol{h}_{n-r})(g) \;=\; \frac{1}{|\mathfrak{S}_{r,n-r}|} \cdot \#\bigl\{x \in \mathfrak{S}_n \mid xgx^{-1} \in C_r \times \mathfrak{S}_{n-r}\bigr\} \cdot r.$$

To compute this character value, note that the set in the middle is a union of right cosets $(\mathfrak{S}_r \times \mathfrak{S}_{n-r}) \cdot y$ where $y \in \mathfrak{S}_d$ maps one of the r-cycles of g to $\mathfrak{S}_r \times \{1\}$. The number of such cosets is therefore equal to $m_r(\lambda)$, and we obtain the character values

$$(\boldsymbol{p}_r \circ \boldsymbol{h}_{n-r})(g) \;=\; r \cdot m_r(\lambda)$$

whose sum is equal to the desired value $\deg(\lambda) = n = n \cdot \boldsymbol{p}_n(g)$ in (2).

(3) This is a formal identity of power series. We start from the formal product expansion

$$\sum_{n\geq 0} h_n(x)\cdot t^n \;=\; \prod_i (1-x_i t)^{-1}$$

in infinitely many variables x_i and a dummy variable t. Taking $\frac{d}{dt}\log$ we obtain that

$$\frac{\sum_{n\geq 1} n\cdot h_n(x)\cdot t^{n-1}}{\sum_{n\geq 0} h_n(x)\cdot t^n} \;=\; \sum_i \tfrac{d}{dt}\log(1-x_i t)^{-1} \;=\; \sum_{r\geq 1}\sum_i x_i^r t^{r-1}$$

and hence

$$\sum_{n\geq 1} n\cdot h_n(x)\cdot t^{n-1} \;=\; \sum_{m\geq 0} h_m(x)\cdot t^m\cdot \sum_{r\geq 1} q_r(x)\cdot t^{r-1}$$

which gives the desired formula by equating coefficients of t^{n-1} on both sides. □

Example 4.4. In the special case $d=3$ the Schur-Weyl decomposition takes the form

$$V^{\otimes 3} \;\simeq\; \mathbf{1}\boxtimes \mathrm{Sym}^3(V) \;\oplus\; W\boxtimes S_{2,1}(V) \;\oplus\; sgn\boxtimes \mathrm{Alt}^3(V)$$

where W is the two-dimensional standard representation of $\mathfrak{S}_3$. We get the Schur polynomials

$$\begin{aligned} s_3 \;&=\; h_3 \;=\; \sum_i x_i^3 + \sum_{i\neq j} x_i^2 x_j + \sum_{i<j<k} x_i x_j x_k,\\ s_{1,1,1} \;&=\; e_3 \;=\; \sum_{i<j<k} x_i x_j x_k,\\ s_{2,1} \;&=\; \frac{e_1^3 - s_3 - s_{1,1,1}}{2} \;=\; \sum_{i\neq j} x_i^2 x_j + 2\sum_{i<j<k} x_i x_j x_k. \end{aligned}$$

One then computes

$$\begin{aligned} p_{1,1,1} \;&=\; \sum_i x_i^3 + 3\sum_{i\neq j} x_i^2 x_j + 6\sum_{i<j<k} x_i x_j x_k &&= s_3 + s_{1,1,1} + 2 s_{2,1}\\ p_3 \;&=\; \sum_i x_i^3 &&= s_3 + s_{1,1,1} - s_{2,1}\\ p_{2,1} \;&=\; \sum_i x_i^3 + \sum_{i\neq j} x_i^2 x_j &&= s_3 - s_{1,1,1}. \end{aligned}$$

The transpose of the matrix of coefficients on the right hand side gives the character table for $\mathfrak{S}_3$ as predicted:

	1	(123)	(12)
1	1	1	1
sgn	1	1	−1
W	2	−1	0

In the next section we will develop a general formula for Schur functions which will allow to give a closed formula for the character values of any symmetric group.

5. The character formula

Let $\lambda = (\lambda_1, \dots, \lambda_r)$ be a partition of degree d. In order to compute the Schur polynomial

$$s_\lambda(x_1, \dots, x_n) \;=\; ch_n(\boldsymbol{s}_\lambda) \;\in\; S(x_1, \dots, x_n) \quad \text{for} \quad n \;\geq\; r$$

directly from the definition, we need to consider the action of diagonal matrices on the Schur functor

$$diag(x_1, \dots, x_n) \;\circlearrowleft\; S_\lambda(\mathbb{C}^n) = c_\lambda \cdot (\mathbb{C}^n)^{\otimes d}$$

and then compute the traces in a suitable basis. This is indeed possible but needs some combinatorial effort; in what follows we will take a different route, so we just sketch the idea:

Remark 5.1. By a *semistandard tableau* of shape λ on n letters, we mean a Young tableau with entries in $[n]$ whose rows are weakly increasing and whose columns are strictly increasing. For instance, we have the following semistandard tableaux of shape $(2,1)$ on three letters:

$$\begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 3 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 3 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 2 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 3 \\ \cline{1-1}\end{array}$$

In general, if V is a complex vector space with an ordered basis $e_1, \dots, e_n$, then for any semistandard tableau T we consider the vector

$$v_T \;=\; (e_{T_{11}} \wedge \cdots \wedge e_{T_{1\mu_1}}) \otimes \cdots \otimes (e_{T_{s,1}} \wedge \cdots \wedge e_{T_{s\mu_s}}) \;\in\; b_\lambda \cdot V^{\otimes d} \;=\; \bigotimes_{i=1}^{s} \mathrm{Alt}^{\mu_i}(V)$$

where $\mu = (\mu_1, \dots, \mu_s) = \lambda^t$ denotes the transpose partition. We then form the image

$$e_T \;=\; a_\lambda \cdot v_T \;\in\; S_\lambda(V) \;=\; c_\lambda \cdot V^{\otimes d}$$

under the row symmetrizer. Note that both the row and column symmetrizers are formed with respect to a fixed Young tableau of shape λ and content $\{1, 2, \dots, d\}$, which should not be confused with the tableau T that describes which basis vectors are selected for the elementary tensors. One may show:

Fact 5.2. *When T ranges over all semistandard tableaux of shape λ on n letters, the vectors e_T defined above form a basis of $S_\lambda(V)$.*

Note that diagonal matrices in $Gl(V)$ with respect to the basis $e_1, \dots, e_n$ act on the e_T by

$$diag(x_1, \dots, x_n) \cdot e_T \;=\; x_1^{m_T(1)} \cdots x_n^{m_T(n)} \cdot e_T$$

where the exponents are the multiplicities $m_T(\nu) = \#\{(i,j) \mid T_{ij} = \nu\}$. For the Schur polynomials this gives

$$s_\lambda(x_1, \dots, x_n) \;=\; \sum_T x_1^{m_T(1)} \cdots x_n^{m_T(n)}$$

where the sum runs over all semistandard tableaux of shape λ on n letters. In what follows we discuss a different approach to Schur polynomials which will result in a determinantal formula more suitable for our application.

Like in the construction of the vectors e_T above, the idea is to write $S_\lambda(V)$ as the image of the composite map given by the column antisymmetrizer and the row symmetrizer:

$$\mathrm{Alt}^{\mu_1}(V) \otimes \cdots \otimes \mathrm{Alt}^{\mu_s}(V) \;\hookrightarrow\; V^{\otimes d} \;\twoheadrightarrow\; \mathrm{Sym}^{\lambda_1}(V) \otimes \cdots \otimes \mathrm{Sym}^{\lambda_r}(V).$$

The Schur functor we are interested in enters both the source and the target of the above map. In the representation ring of the symmetric groups this is reflected by the following result, where for arbitrary partitions λ and μ we write $\lambda \trianglerighteq \mu$ to indicate that

$$\lambda_1 + \cdots + \lambda_i \;\geq\; \mu_1 + \cdots + \mu_i$$

for all i. Note that this *dominance order* is a partial ordering on partitions; as an exercise you may check that $\lambda \trianglerighteq \mu$ iff $\mu^t \trianglerighteq \lambda^t$.

PROPOSITION 5.3. *Let λ and μ be partitions of the same degree. Then*

$$\langle \boldsymbol{h}_\lambda, \boldsymbol{e}_\mu \rangle \;=\; \langle \boldsymbol{e}_\lambda, \boldsymbol{h}_\mu \rangle \;=\; \begin{cases} 0 & \textit{unless } \mu^t \trianglerighteq \lambda, \\ 1 & \textit{if } \mu^t = \lambda. \end{cases}$$

Proof. Recall that by definition of the product in the representation ring we have

$$\begin{aligned} \boldsymbol{h}_\lambda \;&=\; [\mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_d}(\mathbf{1})] \quad \text{for} \quad \mathfrak{S}_\lambda \;=\; \mathfrak{S}_{\lambda_1} \times \cdots \times \mathfrak{S}_{\lambda_r}, \\ \boldsymbol{e}_\lambda \;&=\; [\mathrm{Ind}_{\mathfrak{S}_\mu}^{\mathfrak{S}_d}(sgn)] \quad \text{for} \quad \mathfrak{S}_\lambda \;=\; \mathfrak{S}_{\mu_1} \times \cdots \times \mathfrak{S}_{\mu_s}, \end{aligned}$$

Frobenius reciprocity gives

$$\begin{aligned} \langle \boldsymbol{h}_\lambda, \boldsymbol{e}_\mu \rangle \;&=\; \dim \mathrm{Hom}_{\mathfrak{S}_d}(\mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_d}(\mathbf{1}), \mathrm{Ind}_{\mathfrak{S}_\mu}^{\mathfrak{S}_d}(sgn)) \\ &=\; \dim \mathrm{Hom}_{\mathfrak{S}_\mu}(\mathrm{Res}_{\mathfrak{S}_\mu}^{\mathfrak{S}_d} \mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_d}(\mathbf{1}), sgn). \end{aligned}$$

In the exercises we have seen Mackey's theorem about the restriction of induced representations, which says that if we fix a representative set $S \subset \mathfrak{S}_d$ for the double cosets in $\mathfrak{S}_\mu \backslash \mathfrak{S}_d / \mathfrak{S}_\lambda$ so that

$$\mathfrak{S}_d \;=\; \bigsqcup_{s \in S} \mathfrak{S}_\mu \cdot s \cdot \mathfrak{S}_\lambda,$$

then we have

$$\mathrm{Res}_{\mathfrak{S}_\mu}^{\mathfrak{S}_d} \mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_d}(\mathbf{1}) \;\simeq\; \bigoplus_{s \in S} \mathrm{Ind}_{\mathfrak{S}_\mu \cap s \mathfrak{S}_\lambda s^{-1}}^{\mathfrak{S}_\mu}(\mathbf{1}).$$

For each term on the right hand side, another application of Frobenius reciprocity shows that

$$\mathrm{Hom}_{\mathfrak{S}_\mu}(\mathrm{Ind}_{\mathfrak{S}_\mu \cap s \mathfrak{S}_\lambda s^{-1}}^{\mathfrak{S}_\mu}(\mathbf{1}), sgn) \;=\; \mathrm{Hom}_{\mathfrak{S}_\mu \cap s \mathfrak{S}_\lambda s^{-1}}(\mathbf{1}, sgn).$$

The right hand side does not change if the roles of $\mathbf{1}$ and sgn are interchanged, so it follows that $\langle \boldsymbol{h}_\lambda, \boldsymbol{e}_\mu \rangle = \langle \boldsymbol{e}_\lambda, \boldsymbol{h}_\mu \rangle$. Furthermore, this right hand side is either zero or one-dimensional, and the latter occurs iff

$$\mathfrak{S}_\mu \cap s \mathfrak{S}_\lambda s^{-1} \;\subseteq\; \ker(sgn) \;=\; \mathfrak{A}_d.$$

In the latter case the subgroup on the left contains no transpositions, so if we fill λ and μ row by row from left to right and top to bottom, then s maps the entries of each row of λ into pairwise distinct rows of μ. By definition of the dominance order this can happen only if

$$\mu^t \;\trianglerighteq\; \lambda.$$

If $\mu^t = \lambda$, it remains to show that there exists a unique coset representative s with the above property. This follows as in proposition 1.5: If $s \in S$ maps the entries of each row of λ into distinct rows of μ, then

$$s \cdot p \cdot \lambda \;=\; q \cdot \mu^t \quad \text{for some} \quad p \;\in\; P_\lambda \quad \text{and} \quad q \;\in\; Q_{\mu^t} \;=\; P_\mu.$$

Then $q^{-1} s p \;=\; s_0 \;\in\; \mathfrak{S}_d$ is the unique permutation with $s_0 \cdot \lambda \;=\; \mu^t$ and this determines the double coset of s in $\mathfrak{S}_\mu \backslash \mathfrak{S}_d / \mathfrak{S}_\lambda$ uniquely. $\square$

COROLLARY 5.4. *Let $\lambda = \mu^t$ be conjugate partitions of d. Then V_λ is the unique irreducible representation that enters both in*

$$\mathrm{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_d}(\mathbf{1}) \quad \textit{and} \quad \mathrm{Ind}_{\mathfrak{S}_\mu}^{\mathfrak{S}_d}(sgn).$$

Furthermore, its multiplicity in each of these induced representations is one.

Proof. By the proposition there is a unique irreducible representation V which enters both these induced representations. To see $V \simeq V_\lambda$ we must by Frobenius reciprocity only check

$$\mathbf{1} \hookrightarrow \mathrm{Res}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_d}(V_\lambda) \quad \text{and} \quad sgn \hookrightarrow \mathrm{Res}_{\mathfrak{S}_\mu}^{\mathfrak{S}_d}(V_\lambda),$$

which one may read off from $V_\lambda = \mathscr{A} \cdot a_\lambda b_\lambda = \mathscr{A} \cdot b_\lambda a_\lambda$ for $\mathscr{A} = \mathbb{C}[\mathfrak{S}_d]$. □

REMARK 5.5. Forgetting about Young symmetrizers for a moment, we could as well have *defined* V_λ by the property in the above corollary: Their existence only uses the previous proposition, and the proof of the next theorem will give an independent argument that these V_λ are pairwise non-isomorphic and hence form a complete set of irreducibles. While the combinatorial approach using Young diagrams has its own interest (see e.g. remark 5.1), the approach via representations of $Gl_n(\mathbb{C})$ generalizes to any algebraic group or compact Lie group. The following is a special case of the general *Weyl character formula* for such groups:

THEOREM 5.6. *Let $\lambda = (\lambda_1, \dots, \lambda_r)$ be a partition of length $r \leq n$. Put $\lambda_i = 0$ for $i > r$, then we have*

$$s_\lambda(x_1, \dots, x_n) \;=\; \frac{1}{\Delta} \cdot \det\bigl(x_j^{\lambda_i + n - i}\bigr)_{1 \leq i,j \leq n}$$

for the discriminant

$$\Delta \;=\; \det\bigl(x_j^{n-i}\bigr)_{1 \leq i,j \leq n} \;=\; \prod_{i<j}(x_i - x_j).$$

Proof. In the next section we will show the classical *determinantal formulas* of Jacobi-Trudi and Giambelli

$$\frac{1}{\Delta} \cdot \det\bigl(x_j^{\lambda_i+n-i}\bigr)_{1 \leq i,j \leq n} \;=\; \det\bigl(h_{\lambda_i - i + j}\bigr)_{1 \leq i,j \leq r} \;=\; \det\bigl(e_{\mu_i - i + j}\bigr)_{1 \leq i,j \leq s}.$$

where $\mu = \lambda^t = (\mu_1, \dots, \mu_s)$ denotes the conjugate partition and by convention we put

$$e_\nu(x_1, \dots, x_n) \;=\; h_\nu(x_1, \dots, x_n) \;=\; \begin{cases} 1 & \text{for } \nu = 0, \\ 0 & \text{for } \nu < 0. \end{cases}$$

For the rest of the proof we forget the determinant on the left hand side, we will use only the equality between the determinant in the middle and the one on the right to see that $s_\lambda(x_1, \dots, x_n)$ must be equal to both of them. By direct inspection one verifies that

- the determinant in the middle is a sum of terms $\pm h_{\lambda'}$ with $\lambda' \trianglerighteq \lambda$, and h_λ occurs precisely once,
- the determinant on the right is a sum of terms $\pm e_{\mu'}$ with $\mu' \trianglerighteq \mu$, and e_μ occurs precisely once.

Now by the previous proposition we can have $\langle \boldsymbol{h}_{\lambda'}, \boldsymbol{e}_{\mu'} \rangle \neq 0$ only if $(\mu')^t \trianglerighteq \lambda'$, which for $\mu' \trianglerighteq \mu$ and $\lambda' \trianglerighteq \lambda$ implies

$$\lambda \;=\; \mu^t \;\trianglerighteq\; (\mu')^t \;\trianglerighteq\; \lambda' \;\trianglerighteq\; \lambda$$

and this is possible only if equality holds everywhere. So we obtain for the virtual representation

$$\boldsymbol{d}_\lambda \;=\; \det\bigl(\boldsymbol{h}_{\lambda_i-i+j}\bigr)_{1\le i,j\le r} \;=\; \det\bigl(\boldsymbol{e}_{\mu_i-i+j}\bigr)_{1\le i,j\le s}.$$

that

$$\langle \boldsymbol{d}_\lambda, \boldsymbol{d}_\lambda\rangle \;=\; \langle \boldsymbol{h}_\lambda, \boldsymbol{e}_\mu\rangle + \sum_{\substack{\lambda'\trianglerighteq\lambda,\mu'\trianglerighteq\mu\\ (\lambda',\mu')\neq(\lambda,\mu)}} \pm\langle \boldsymbol{h}_{\lambda'}, \boldsymbol{e}_{\mu'}\rangle \;=\; \langle \boldsymbol{h}_\lambda, \boldsymbol{e}_\mu\rangle \;=\; 1,$$

which by positive definiteness of the scalar product implies that either $+\boldsymbol{d}_\lambda$ or $-\boldsymbol{d}_\lambda$ must be not just a virtual but a true irreducible representation. Since we know that $\langle \boldsymbol{d}_\lambda, \boldsymbol{h}_\lambda\rangle = \langle \boldsymbol{d}_\lambda, \boldsymbol{e}_\mu\rangle = 1$, we then get $\boldsymbol{d}_\lambda = [V_\lambda]$ from the previous corollary and the claim follows. □

Let us see how this allows to compute characters of the symmetric group. We fix a number $n \in \mathbb{N}$ and extend partitions λ of length $\ell \le n$ by zeroes at the end, writing

$$\lambda = (\lambda_1, \dots, \lambda_n) \quad \text{with} \quad \lambda_{\ell+1} = \cdots = \lambda_n = 0.$$

For any such partition let

$$\lambda^\star \;=\; (\lambda_1 + n - 1, \lambda_2 + n - 2, \dots, \lambda_n)$$

be the partition which is obtained from the Young diagram of λ by adding i boxes to row $n-i$ for $i = 1, \dots, n-1$. Note that these are precisely the exponents that occur in the matrix entries of the Weyl character formula from the previous theorem and these form a strictly decreasing sequence, for example

$$\lambda = (3,1,1) \qquad \rightsquigarrow \qquad \lambda^\star = (6,3,2).$$

For polynomials $P(x_1, \dots, x_n)$ we put

$$\bigl[P\bigr]_\lambda \;=\; \text{coefficient of the monomial } x_1^{\lambda_1}\cdots x_n^{\lambda_n} \text{ in } P(x_1,\dots,x_n).$$

The above theorem then shows

Corollary 5.7. *Any symmetric function $P \in S(x_1, \dots, x_n)$ has the expansion in Schur polynomials*

$$P(x_1,\dots,x_n) \;=\; \sum_\mu \bigl[\Delta\cdot P\bigr]_{\mu^\star}\cdot s_\mu(x_1,\dots,x_n).$$

Proof. Note that we have only defined $\mu^\star$ if $\ell(\mu) \le n$, but for all other μ the Schur polynomials $s_\mu(x_1, \dots, x_n)$ vanish. So the right hand side is well-defined with the summation running over the finitely many partitions μ of length $\ell(\lambda) \le n$ and degree $\deg(\mu) \le \deg(P)$. This being said, since the right hand side of the claimed formula is clearly linear in P it suffices to prove it in the special case of Schur polynomials

$$P \;=\; s_\lambda(x_1,\dots,x_n).$$

Then the theorem says

$$\Delta\cdot s_\lambda \;=\; \det\bigl(x_j^{\lambda_i+n-i}\bigr)_{1\le i,j\le n} \;=\; x_1^{\lambda_1+n-1}x_2^{\lambda_2+n-2}\cdots x_n^{\lambda_n} \;+\; \cdots$$

where "$\cdots$" does not involve any monomials with strictly decreasing exponents. In other words

$$\bigl[\Delta\cdot s_\lambda\bigr]_{\mu^\star} \;=\; \begin{cases} 1 & \text{if } \mu=\lambda,\\ 0 & \text{otherwise,}\end{cases}$$

and hence the claim follows. □

We now get a closed formla for the characters of all irreducible representations of the symmetric groups as follows:

Corollary 5.8 (Frobenius character formula). *If λ and μ are partitions of d, then*

$$\chi_{V_\mu}(g) \;=\; \big[\Delta \cdot p_\lambda\big]_{\mu^*} \quad \textit{for all } g \in \mathfrak{S}_d \textit{ of cycle type } \lambda.$$

Proof. We have seen in the previous section that for any g of cycle type λ one has $p_\lambda(x_1, \dots, x_n) = \sum_\mu \chi_{V_\mu}(g) \cdot s_\mu(x_1, \dots, x_n)$, so the claim follows by comparison with the above corollary since Schur polynomials are linearly independent. Notice that the number n of variables only needs to satisfy $n \geq \ell(\mu)$, for any such n the formula gives the same value and works for cycle types of any length $\ell(\lambda)$. □

Let us see how this works in a concrete example: Let χ be the character of the standard representation V_μ of $\mathfrak{S}_3$. Here $\mu = (2,1)$ and we can take $n = 2$, which gives the values

$$\begin{aligned}
\chi((1)(2)(3)) &= \big[(x_1 - x_2)\cdot(x_1 + x_2)^3\big]_{(3,1)} \;=\; 2, \\
\chi((12)(3)) &= \big[(x_1 - x_2)\cdot(x_1^2 + x_2^2)\cdot(x_1 + x_2)\big]_{(3,1)} \;=\; 0, \\
\chi((123)) &= \big[(x_1 - x_2)\cdot(x_1^3 + x_2^3)\big]_{(3,1)} \;=\; -1.
\end{aligned}$$

by taking the coefficient of $x_1^3 x_2$ in the respective polynomials. This is much easier than our previous computations of character tables!

The Frobenius character formula in particular allows to compute the dimensions of the irreducible representations $V_\mu \in \mathrm{Rep}_{\mathbb{C}}(\mathfrak{S}_d)$ as follows. Recall that we have put

$$\mu^\star \;=\; (l_1, l_2, \dots, l_n) \quad \text{where} \quad l_i \;=\; \mu_i + n - i \quad \text{for} \quad i \;=\; 1, 2, \dots, n,$$

extending our partitions by zeroes at the end if needed. Although the l_i depend on the chosen n, the outcome of the following formula does not:

Corollary 5.9 (Frobenius dimension formula). *Let $n \geq \ell(\mu)$. Then with l_i as above we have*

$$\dim_{\mathbb{C}}(V_\mu) \;=\; \frac{d!}{l_1! \cdots l_n!} \prod_{1 \leq i < j \leq n} (l_i - l_j).$$

Proof. We have $\dim_{\mathbb{C}}(V_\mu) = \chi_{V_\mu}(1)$, so we need to take $\lambda = (1, 1, \dots, 1)$ in the character formula:

$$\dim_{\mathbb{C}}(V_\mu) \;=\; \text{coefficient of} \quad x_1^{l_1} \cdots x_n^{l_n} \quad \text{in} \quad p_{1,1,\dots,1} \cdot \Delta.$$

Here

$$p_{1,1,\dots,1} \;=\; (x_1 + \cdots + x_n)^d \;=\; \sum_{e_1 + \cdots + e_n = d} \frac{d!}{e_1! \cdots e_n!} \cdot x_1^{e_1} \cdots x_n^{e_n}$$

and

$$\Delta \;=\; \det \begin{pmatrix} 1 & x_n & \cdots & x_n^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_1 & \cdots & x_1^{n-1} \end{pmatrix} \;=\; \sum_{\sigma \in \mathfrak{S}_n} sgn(\sigma) \cdot x_1^{\sigma(n)-1} \cdots x_n^{\sigma(1)-1}$$

so that

$$\begin{aligned}
\dim_{\mathbb{C}}(V_\mu) &= \sum_{\sigma \in \mathfrak{S}_d} sgn(\sigma) \cdot \frac{d!}{\prod_i (l_i - \sigma(n-i+1) + 1)!} \\
&= \frac{d!}{l_1! \cdots l_n!} \cdot \sum_{\sigma \in \mathfrak{S}_d} sgn(\sigma) \cdot \prod_i l_i(l_i - 1) \cdots (l_i - \sigma(n-i+1) + 2)
\end{aligned}$$

where the sums run over all σ such that $e_i = l_i - \sigma(n-i+1)+1 \geq 0$ for all i. The second sum can be rewritten as a determinant, which by column reduction takes the form

$$\det\begin{pmatrix} 1 & l_n & l_n(l_n-1) & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 1 & l_1 & l_1(l_1-1) & \cdots \end{pmatrix} = \begin{pmatrix} 1 & l_n & l_n^2 & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 1 & l_1 & l_1^2 & \cdots \end{pmatrix} = \prod_{i<j}(l_i - l_j).$$

Hence the claim follows. □

The dimension formula has a nice interpretation in terms Young diagrams. Let us define for any box $b \in \mu$ in the Young diagram of shape μ the *hook length* $h(b)$ to be the number of boxes which are

- to the right of the box in the same row, or
- below the box in the same column,

with the box itself counted precisely once. For example, the following are the hook lengths for $\mu = (4,3,3,2)$:

7	6	4	1
5	4	2	
4	3	1	
2	1		

We then get the following

Corollary 5.10 (Hook length formula). *For any partition μ of degree d one has*

$$\dim_{\mathbb{C}}(V_\mu) \; = \; \frac{d!}{\prod_{b\in\mu} h(b)}.$$

Proof. We label the rows and columns of the Young diagram μ from left to right and top to bottom as usual. Then the key observation is that for any $(i,j) \in \mu$, the numbers

$$\begin{aligned} h(i,b) \quad &\text{for} \quad b \; \geq \; j \\ h(i,j) - h(a,j) \quad &\text{for} \quad a \; > \; i \end{aligned}$$

are pairwise distinct and that each of the numbers $1, 2, \ldots, h(i,j)$ occurs among them precisely once: As an exercise you may check by looking at the jumps in the column lengths that the differences $h(i,j) - h(a,j)$ precisely fill out the gaps between the numbers $h(i,b)$. For $j = 1$ we get that the product of all hook lengths in row i is equal to

$$\frac{h(i,1)!}{\prod_{a>i}\big(h(i,1) - h(a,1)\big)}$$

where

$$h(i,1) \; = \; \mu_i + n - i \; = \; l_i \quad \text{for} \quad 1 \; \leq \; i \; \leq \; n \; = \; \ell(\mu)$$

in the previous notations. Taking the product over all rows i we obtain the claim from the Frobenius character formula. □

Remark 5.11. In the exercises we will see by an independent combinatorial argument that $\dim_{\mathbb{C}}(V_\mu)$ is the number of *standard tableaux of shape* μ, i.e. Young tableaux of shape μ whose rows and columns are increasing from left to right and top to bottom. Viewing the corollary as a formula for the number of such tableaux, we get a heuristic to memorize the hook length formula: A Young tableau is standard iff the label of each box $b \in \mu$ is smallest among those of its hook. The probability for this is $1/h(b)$, and under the boldly unrealistic assumption that these events are independent we get that the number of standard tableaux should be $\prod_{b\in\mu} 1/h(b)$ times the total number $d!$ of Young tableaux.

6. Determinantal formulas

It remains to prove the two determinantal formulas that have been used in the proof of the Weyl character formula in theorem 5.6. Both come from the formal identities

$$H(x,t) \;:=\; \sum_{j\geq 0} h_j(x)\cdot t^j \;=\; \prod_{i=1}^n (1-x_i t)^{-1}$$

$$E(x,t) \;:=\; \sum_{j\geq 0} e_j(x)\cdot t^j \;=\; \prod_{i=1}^n (1+x_i t)$$

for the generating functions of the complete symmetric and elementary symmetric polynomials in $x=(x_1,\dots,x_n)$, for some fixed $n\in\mathbb{N}$. Notice that $e_j(x)=0$ for $j>n$. Put

$$\det(x^\alpha) \;=\; \det\bigl(x_j^{\alpha_i}\bigr)_{1\leq i,j\leq n} \quad \text{for} \quad \alpha \;=\; (\alpha_1,\dots,\alpha_n) \;\in\; \mathbb{N}_0^n.$$

For instance, we have

$$\Delta(x) \;=\; \det(x^\delta) \quad \text{for} \quad \delta \;=\; (n-1,n-2,\dots,1,0).$$

and the first determinantal formula that we used in the proof of theorem 5.6 can be written in the following form, where as usual we extend partitions by zeroes at the end if required:

Theorem 6.1 (Jacobi-Trudi). *For any partition λ of length $\ell(\lambda)\leq n$ we have the formula*

$$\frac{\det(x^{\lambda+\delta})}{\det(x^\delta)} \;=\; \det\bigl(h_{\lambda_i-i+j}\bigr)_{1\leq i,j\leq n}.$$

Proof. In order to separate the various coordinate functions, for $1\leq k\leq n$ we put

$$E^{(k)}(x,t) \;=\; \prod_{i\neq k}(1+x_i t) \;=\; \sum_{j\geq 0} e_j^{(k)}(x)\cdot t^j$$

where

$$e_j^{(k)}(x) \;=\; e_j(x_1,\dots,x_{k-1},0,x_{k+1},\cdots,x_n)$$

is an elementary symmetric polynomial in one variable less. From the generating series we have

$$H(x,t)\cdot E^{(k)}(x,-t) \;=\; (1-x_k t)^{-1}$$

and hence

$$\sum_{j=1}^n h_{a-n+j}(x)\cdot(-1)^{n-j}\, e_{n-j}^{(k)}(x) \;=\; x_k^a \quad \text{for all} \quad a \;\in\; \mathbb{N}_0$$

by comparing the coefficients of t^a on both sides. Note that $e_\nu^{(k)}(x)=0$ for $\nu\geq n$ as these are elementary symmetric polynomials in $n-1$ variables. Taking the above relation with $a = \alpha_i$ ranging over the coordinates of a given $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ we get

$$\mathbb{H}_\alpha \cdot \mathbb{E} \;=\; \det(x^\alpha) \quad \text{and hence} \quad \det(\mathbb{H}_\alpha) \cdot \det(\mathbb{E}) \;=\; \det(x^\alpha)$$

for the matrices

$$\mathbb{H}_\alpha \;=\; \Big(h_{\alpha_i - n + j}(x)\Big)_{1 \le i,j \le n} \quad \text{and} \quad \mathbb{E} \;=\; \Big((-1)^{n-j} e_{n-j}^{(k)}(x)\Big)_{1 \le j,k \le n}.$$

Notice that the second of these two matrices does not depend on the chosen value of α. To compute its determinant, we plug in $\alpha = \delta = (n-1, n-2, \dots, 1, 0)$ and note that $\mathbb{H}_\delta$ is a triangular matrix with entries 1 on the diagonal. Hence the above equation says

$$\det(\mathbb{E}) \;=\; \det(\mathbb{H}_\delta) \cdot \det(\mathbb{E}) \;=\; \det(x^\delta) \;=\; \Delta(x)$$

is the discriminant. This being said, the theorem follows by taking $\alpha = \lambda + \delta$ and by noting that $\mathbb{H}_{\lambda+\delta}$ is a block triangular matrix, the top left block being the matrix that figures in the right hand side of the Jacobi-Trudi formula, and the bottom right block being triangular with entries 1 on the diagonal. □

To finish the proof of theorem 5.6, it remains to show that the determinant in the previous theorem is equal to the corresponding determinant with entries in the elementary rather than complete symmetric polynomials. Before we can do this we need some preparations: For a matrix $A = (a_{ij})_{1 \le i,j \le N}$ and $I, J \in \{1, 2, \dots, N\}^r$ with $r \le N$, let

$$A_{I,J} \;=\; \big(a_{ij}\big)_{(i,j) \in I \times J}$$

be the $I \times J$ *minor* of A. We then have the following generalization of Cramer's formula for the inverse of a matrix:

LEMMA 6.2. *Let $A, B \in \mathrm{Mat}_{N \times N}(\mathbb{C})$ be two matrices such that $A \cdot B = c \cdot id$ for some scalar $c \in \mathbb{C}$. Consider two permutations (I, I') and (J, J') of $(1, 2, \dots, N)$ with $|I| = |J|$. Then*

$$c^{N-|I|} \cdot \det(A_{I,J}) \;=\; \epsilon \cdot \det(A) \cdot \det(B_{J',I'})$$

where $\epsilon \in \{\pm 1\}$ denotes the product of the signs of the two permutations.

Proof. Let $P, Q \in Gl_N(\mathbb{C})$ denote the permutation matrices such that left multiplication by these matrices permutes the rows in any given matrix according to (I, I') resp. (J, J'). Since the transpose of a permutation matrix is equal to the inverse, the right multiplication by P^{-1} resp. Q^{-1} then gives the corresponding permutations of the columns. So we can shift the minors in the top left resp. bottom right position via

$$PAQ^{-1} \;=\; \begin{pmatrix} A_{IJ} & * \\ * & * \end{pmatrix} \quad \text{and} \quad QBP^{-1} \;=\; \begin{pmatrix} * & * \\ * & B_{J'I'} \end{pmatrix}.$$

By assumption

$$\begin{pmatrix} A_{IJ} & * \\ * & * \end{pmatrix} \cdot \begin{pmatrix} * & * \\ * & B_{J'I'} \end{pmatrix} \;=\; PABP^{-1} \;=\; c \cdot PP^{-1} \;=\; c \cdot id.$$

Writing out each block of this matrix equation separately, one sees that then also the identity

$$\begin{pmatrix} A_{IJ} & * \\ * & * \end{pmatrix} \cdot \begin{pmatrix} id & * \\ 0 & B_{J'I'} \end{pmatrix} \;=\; \begin{pmatrix} A_{IJ} & 0 \\ * & c \cdot id \end{pmatrix}$$

holds, where on the left hand side of the equation the entries labelled $*$ have the same meaning as before. Taking determinants we get

$$\det(PAQ^{-1}) \cdot \det(B_{J'I'}) \;=\; c^{N-|I|} \cdot \det(A_{IJ})$$

and the claim follows since $\det(PQ^{-1}) = \det(P)\det(Q) = \epsilon$. □

THEOREM 6.3 (Giambelli). *If $\lambda = (\lambda_1, \dots, \lambda_r)$ and $\mu = (\mu_1, \dots, \mu_s) = \lambda^t$ are conjugate partitions, then*

$$\det\Big(h_{\lambda_i - i + j}(x)\Big)_{1 \le i,j \le r} \;=\; \det\Big(e_{\mu_i - i + j}(x)\Big)_{1 \le i,j \le s}$$

Proof. The idea is again to obtain a matrix relation from the generating series for the elementary and complete symmetric polynomials. Indeed, the identity of power series

$$H(x,t) \cdot E(x,-t) \;=\; 1$$

shows for $N \in \mathbb{N}$ that the matrices

$$A \;=\; \big(h_{i-j}(x)\big)_{1 \le i,j \le N} \quad \text{and} \quad B \;=\; \big((-1)^{i-j} e_{i-j}(x)\big)_{1 \le i,j \le N}$$

are inverse to each other. Furthermore, these matrices are triangular with entries 1 on the diagonal, hence have determinant one. We take $N = r + s$ and apply the previous lemma to

$$I \;=\; \big(\lambda_i + r + 1 - i \mid i = 1, \dots, r\big), \qquad J \;=\; \big(r + 1 - j \mid j = 1, \dots, r\big),$$
$$I' \;=\; \big(r + i - \mu_i \mid i = 1, \dots, s\big), \qquad J' \;=\; \big(r + j \mid j = 1, \dots, s\big).$$

Clearly (J, J') is a permutation of $(1, 2, \dots, N)$. To see the same for (I, I'), notice that

$i \mapsto \lambda_i + r + 1 - i$ is strictly decreasing and maps $\{1, \dots, r\}$ into $\{1, \dots, N\}$,
$i \mapsto r + i - \mu_i$ is strictly increasing and maps $\{1, \dots, s\}$ into $\{1, \dots, N\}$.

and that these two sequences are disjoint: If for some $1 \le i \le r$ and $1 \le k \le s$ we had

$$\lambda_i + r + 1 - i \;=\; r + k - \mu_k,$$

then

$$\lambda_i + \mu_k \;=\; k + i - 1,$$

so $\lambda_i \ge k$ or $\mu_k \ge i$. But then in fact both these inequalities would hold, because each one implies the other since λ and μ are transpose to each other. Thus we obtain $\lambda_i + \mu_k \ge i + k > k + i - 1$, a contradiction. □

REMARK 6.4. We have used both Jacobi-Trudi and Giambelli to prove the Weyl character formula

$$s_\lambda(x) \;=\; \frac{\det(x^{\lambda+\delta})}{\det(x^\delta)}$$

by comparison of two different induced representations. A posteriori we can rewrite the Jacobi-Trudi-Giambelli formulas as

$$s_\lambda(x) \;=\; \det\Big(h_{\lambda_i - i + j}(x)\Big)_{1 \le i,j \le r} \quad \text{and} \quad s_{\lambda^t}(x) \;=\; \det\Big(e_{\lambda_i - i + j}(x)\Big)_{1 \le i,j \le s}.$$

In this more symmetric formulation the Jacobi-Trudi and Giambelli formulae are mirror images of each other in the following sense: The representation ring admits a natural involution

$$\iota: \quad R \longrightarrow R, \quad [V] \;\mapsto\; [V \otimes sgn]$$

which is a ring homomorphism with $\iota(\boldsymbol{s}_\lambda) = \boldsymbol{s}_{\lambda^t}$ and $\iota(\boldsymbol{e}_\lambda) = \boldsymbol{h}_\lambda$ (exercise); under ι the Jacobi-Trudi formula is mapped to the Giambelli formula and vice versa.

The Jacobi-Trudi and Giambelli formulae are usually presented together with yet another determinantal formula which describes the multiplication of symmetric functions in the basis of Schur polynomials:

THEOREM 6.5 (Pieri's formula). *For any partition λ and $m \in \mathbb{N}$, one has the formulas*

$$s_\lambda(x) \cdot e_m(x) \;=\; \sum_\mu s_\mu \qquad \textit{and} \qquad s_\lambda(x) \cdot h_m(x) \;=\; \sum_\nu s_\nu$$

where μ and ν run over all partitions that are obtained from λ by adding m boxes in pairwise distinct rows respectively columns.

Proof. Applying the involution ι from above, it suffices to prove the first of the two formulas. Now

$$\begin{aligned}
\det(x^{\lambda+\delta}) \cdot e_m(x) &= \sum_{\sigma \in \mathfrak{S}_n} sgn(\sigma) \cdot x^{\sigma(\lambda+\delta)} \sum_{1 \le i_1 < \cdots < i_m \le n} x_{i_1} \cdots x_{i_m} \\
&= \sum_{\sigma \in \mathfrak{S}_n} sgn(\sigma) \cdot x^{\sigma(\lambda+\delta)} \sum_{1 \le i_1 < \cdots < i_m \le n} x_{\sigma(i_1)} \cdots x_{\sigma(i_m)} \\
&= \sum_{\epsilon \in \{0,1\}^n, \deg(\epsilon) = m} \det(x^{\lambda+\delta+\epsilon})
\end{aligned}$$

and in the last sum, the only non-zero terms are those where $\lambda + \delta + \epsilon$ is strictly decreasing. These are precisely the ones corresponding to the μ in the statement of the theorem, hence the claim follows after multiplication by Δ on both sides. $\square$

Recall that in the previous section we have characterized $V_\mu \in \mathrm{Rep}_{\mathbb{C}}(\mathfrak{S}_d)$ as the unique irreducible representation of the symmetric group that enters both induced representations

$$\mathrm{Ind}_{\mathfrak{S}_\mu}^{\mathfrak{S}_d}(\mathbf{1}) \quad \text{and} \quad \mathrm{Ind}_{\mathfrak{S}_{\mu^t}}^{\mathfrak{S}_d}(sgn).$$

As an application of Pieri's formula we may completely decompose these induced representations as follows. For partitions λ, μ of degree d the *Kostka numbers* are defined by

$$K_{\lambda\mu} \;=\; \#\big\{ \text{semistandard tableaux of shape } \lambda \text{ and content } \mu \big\},$$

where a tableau is said to have *content* μ if it contains each number i precisely μ_i times. For instance

$K_{(3,2)(2,2,1)} \;=\; 2$

1	1	2
2	3	

1	1	3
2	2	

In general,

- $K_{\mu\mu} = 1$,
- $K_{\lambda\mu} = 0$ if $\lambda < \mu$ in the lexicographic order,
- $K_{\lambda,(1,1,\dots,1)}$ is the number of standard tableaux of shape λ, etc.

The above induced representations then decompose as follows, which in particular recovers our previous characterization of $V_\mu \in \mathrm{Rep}_{\mathbb{C}}(\mathfrak{S}_d)$:

COROLLARY 6.6 (Young's rule). *With notations as above,*

$$\mathrm{Ind}_{\mathfrak{S}_\mu}^{\mathfrak{S}_d}(\mathbf{1}) \;\simeq\; \bigoplus_\lambda K_{\lambda\mu} \cdot V_\lambda \quad \text{and} \quad \mathrm{Ind}_{\mathfrak{S}_{\mu^t}}^{\mathfrak{S}_d}(\mathbf{1}) \;\simeq\; \bigoplus_\lambda K_{\lambda^t\mu^t} \cdot V_\lambda.$$

Proof. The second formula follows from the first one by twisting with the sign character since $V_\lambda \otimes sgn \simeq V_{\lambda^t}$. For the first formula we apply the character map to get

$$ch_n\big[\mathrm{Ind}_{\mathfrak{S}_\mu}^{\mathfrak{S}_d}(\mathbf{1})\big] \;=\; e_{\mu_1} e_{\mu_2} \cdots e_{\mu_\ell} \;=\; s_{\mu_1} e_{\mu_2} \cdots e_{\mu_\ell},$$

so that Pieri's formula implies the claim by induction on the length $\ell = \ell(\mu)$. $\square$

CHAPTER IV

Representations of compact Lie groups

The last two chapters were mostly about representations of finite groups, but we have already encountered some representations of the algebraic group $Gl_n(\mathbb{C})$ in relation with Schur-Weyl duality. The ideas from the Weyl character formula will now be developed in the more general framework of *compact Lie groups*, or equivalently of *reductive algebraic groups over* $\mathbb{C}$.

1. Lie groups and algebraic groups

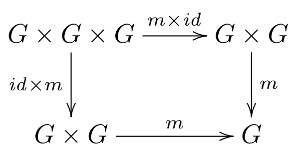
By definition, a *Lie group* is a group whose underlying set carries the additional structure of a smooth manifold such that both structures are compatible, i.e. the multiplication and inversion maps are smooth. It is convenient to reformulate this more abstractly as follows:

Definition 1.1. Let $\mathscr{C}$ be a category with finite products, i.e. there is a final object $\{1\} \in \mathscr{C}$ and any two objects $A, B \in \mathscr{C}$ admit a product $A \times B \in \mathscr{C}$. By a *group object in the category* $\mathscr{C}$ we mean a quadruple consisting of an object $G \in \mathscr{C}$ and morphisms

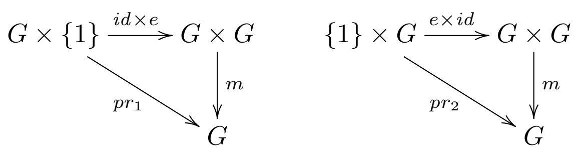
$$\begin{aligned} m : G \times G \to G & \quad \text{(the multiplication)}, \\ i : G \to G & \quad \text{(the inversion map)}, \\ e : \{1\} \to G & \quad \text{(the neutral element)}, \end{aligned}$$

such that the following diagrams commute, where $diag : G \to G \times G$ denotes the diagonal embedding and pr_1, pr_2 are the projections onto the factors:

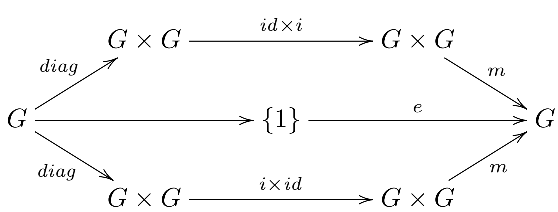
- Associativity:



- Two-sided unit:



- Two-sided inverse:



A *homomorphism* between group objects $G, H \in \mathscr{C}$ is a morphism $f : G \to H$ in $\mathscr{C}$ which is compatible with the unit and multiplication in the sense that the following diagram commutes:

$$\begin{array}{ccc} G \times G & \xrightarrow{m} & G \\ {\scriptstyle f\times f}\downarrow & & \downarrow{\scriptstyle f} \\ H \times H & \xrightarrow{m} & H \end{array}$$

Let us take a look at a few examples:

- If $\mathscr{C}$ is the category of sets, a group object in $\mathscr{C}$ is just a group.
- If $\mathscr{C}$ is the category of finite sets, a group object in $\mathscr{C}$ is a finite group.
- If $\mathscr{C}$ is the category of topological spaces, with continuous maps as morphisms, then a group object in $\mathscr{C}$ is called a *topological group*.

The main topic of this chapter will be the following example:

Example 1.2. If $\mathscr{C}$ is the category of smooth manifolds, with smooth maps as morphisms, then a group object in $\mathscr{C}$ is called a *Lie group*. For instance, $(\mathbb{R}, +)$ and $Gl_n(\mathbb{R})$ are Lie groups. In this chapter we will focus on *compact* Lie groups such as

(1) the *unit circle*

$$U(1) \;=\; \big\{ z \in \mathbb{C} \mid |z| = 1 \big\}$$

and the *compact tori*, i.e. products $U(1) \times \cdots \times U(1)$,

(2) the *unitary groups*

$$U(n) \;=\; \big\{ M \in Gl_n(\mathbb{C}) \mid \overline{M}^{\,t} \cdot M = \mathbf{1} \big\}$$

and the *special unitary groups* $SU(n) = U(n) \cap Sl_n(\mathbb{C})$,

(3) the *orthogonal groups*

$$O(n) \;=\; \big\{ M \in Gl_n(\mathbb{R}) \mid M^t \cdot M = \mathbf{1} \big\}$$

and the *special orthogonal groups* $SO(n) = O(n) \cap Sl_n(\mathbb{C})$,

(4) the *symplectic groups*

$$Sp(n) \;=\; \big\{ M \in Gl_n(\mathbb{H}) \mid \overline{M}^{\,t} \cdot M = \mathbf{1} \big\},$$

where $Gl_n(\mathbb{H})$ denotes the group of invertible $n \times n$ matrices over the Hamiltonian quaternions and the complex conjugation on quaternions $\mathbb{H}$ is defined by

$$\overline{a + bi + cj + dk} \;=\; a - bi - cj - dk \quad \text{for} \quad a, b, c, d \;\in\; \mathbb{R}.$$

We leave it as an exercise to verify that all of these are compact Lie groups.

Example 1.3. If $\mathscr{C}$ is the category of complex manifolds, with holomorphic maps as morphisms, then a group object in $\mathscr{C}$ is called a *complex Lie group*. Examples include

(1) the *multiplicative group*

$$\mathbb{G}_m(\mathbb{C}) \;=\; \mathbb{C}^\times$$

and more generally *complex tori*, i.e. products $\mathbb{G}_m(\mathbb{C}) \times \cdots \times \mathbb{G}_m(\mathbb{C})$,

(2) the *general linear groups* $Gl_n(\mathbb{C})$ and the *special linear groups* $Sl_n(\mathbb{C})$,

(3) the *orthogonal groups*

$$O_n(\mathbb{C}) \;=\; \{M \in Gl_n(\mathbb{C}) \mid M^t \cdot M = \mathbf{1}\}$$

and the *special orthogonal groups* $SO_n(\mathbb{C}) = O_n(\mathbb{C}) \cap Sl_n(\mathbb{C})$,

(4) the *symplectic groups*

$$Sp_{2n}(\mathbb{C}) \;=\; \{M \in Gl_{2n}(\mathbb{C}) \mid M^t \cdot \mathbb{I} \cdot M = \mathbb{I}\}$$

of matrices preserving the standard symplectic form

$$\langle \cdot, \cdot \rangle : \quad \mathbb{C}^{2n} \times \mathbb{C}^{2n} \longrightarrow \mathbb{C}, \quad (u, v) \mapsto u^t \cdot \mathbb{I} \cdot v$$

defined by the $2n \times 2n$ block matrix

$$\mathbb{I} \;=\; \begin{pmatrix} 0 & \mathbf{1} \\ -\mathbf{1} & 0 \end{pmatrix}.$$

Note that each of the above examples is a subgroup of some general linear group, inside of which it is defined by finitely many polynomial equations in the matrix coefficients. This leads to a third incarnation of the objects to be considered, for which we need a bit of language from algebraic geometry:

Definition 1.4. The category $\mathscr{C}$ of *affine algebraic varieties* over a field k is the opposite of the category of reduced k-algebras of finite type. We think of the objects of $\mathscr{C}$ as zero loci of polynomials: For $V \in \mathscr{C}$, write the corresponding algebra as a quotient

$$\mathscr{O}(V) \;\simeq\; k[x_1, \dots, x_n]/I \quad \text{for some ideal} \quad I \;=\; (f_1, \dots, f_m).$$

If k is algebraically closed, then Hilbert's Nullstellensatz says that this ideal can be recovered from the zero locus

$$|V| \;=\; \{a \in k^n \mid f(a) = 0 \,\forall f \in I\} \;\subset\; k^n$$

via

$$I \;=\; \{f \in k[x_1, \dots, x_n] \mid f(a) = 0 \,\forall a \in |V|\} \;\trianglelefteq\; k[x_1, \dots, x_n].$$

If k is not algebraically closed, then in general the above zero locus is not enough to determine the variety, it may even be empty. Nevertheless, intuitively we think of $\mathscr{O}(V)$ as the algebra of polynomial functions on the affine variety $V \in \mathscr{C}$ and call it the *coordinate ring* of the variety.

Example 1.5. By definition, an *affine algebraic group over k* is a group object in the category of affine algebraic varieties from the previous definition. All the complex Lie groups in example 1.3 have underlying algebraic groups, and all of them can be defined over an arbitrary field k. The corresponding coordinate rings have the following shape:

(1) $\mathscr{O}(\mathbb{G}_m) = k[t, t^{-1}]$,

(2) $\mathscr{O}(Gl_n) = k[x_{ij}, \det(x_{ij})^{-1}]$,

$\mathscr{O}(Sl_n) = k[x_{ij}]/(\det(x_{ij}) - 1)$,

(3) $\mathscr{O}(O_n) = k[x_{ij}]/(\delta_{il} - \sum_j x_{ji} x_{jl})$,

$\mathscr{O}(SO_n) = k[x_{ij}]/(\det(x_{ij}) - 1,\ \delta_{il} - \sum_j x_{ji} x_{jl})$,

(4) $\mathscr{O}(Sp_{2n}) = k[x_{ij}]/(\mathbb{I}_{im} - \sum_{j,l} x_{ji} \cdot \mathbb{I}_{jl} \cdot x_{lm})$.

If you unravel the definition of a group object in the category of affine algebraic varieties, you will find the notion of a *Hopf algebra* (which we will not discuss here).

The relation between the various categories of group objects in this section can be summarized by the following

LEMMA 1.6. *One has natural analytification and complexification functors for algebraic groups as indicated by the solid arrows in the following diagram:*

$$\begin{array}{ccc}
\Big\{ \text{affine algebraic groups} \,/\, \mathbb{R} \Big\} & \xrightarrow{(-)_{\mathbb{C}}} & \Big\{ \text{affine algebraic groups} \,/\, \mathbb{C} \Big\} \\
{\scriptstyle (-)^{an}} \big\downarrow & & \big\downarrow {\scriptstyle (-)^{an}} \\
\Big\{ \text{real Lie groups} \Big\} & \overset{(-)_{\mathbb{C}}}{- - - - - - \dashrightarrow} & \Big\{ \text{complex Lie groups} \Big\}
\end{array}$$

Sketch of proof. The complexification functor $G \mapsto G_{\mathbb{C}}$ on algebraic groups is the base change functor from real to complex algebraic varieties, which on coordinate rings is defined by

$$\mathscr{O}(G_{\mathbb{C}}) \ := \ \mathscr{O}(G) \otimes_{\mathbb{R}} \mathbb{C}.$$

This base change functor is compatible with finite products in the respective two categories of varieties, hence it sends group objects to group objects.

The analytification functor $G \mapsto G^{an}$ sends a real affine variety G to the set of its real points

$$G^{an} \ := \ G(\mathbb{R}).$$

Intrinsically the latter is defined as the set of algebra homomorphisms $\mathscr{O}(G) \to \mathbb{R}$, but writing the coordinate ring as $\mathscr{O}(G) \simeq \mathbb{R}[x_1, \dots, x_n]/I$ we can identify it with a subset

$$G(\mathbb{R}) \ \simeq \ \{a \in \mathbb{R}^n \mid f(a) = 0 \, \forall \, f \in I\} \ \subset \ \mathbb{R}^n$$

of some affine space defined by finitely many real polynomial equations. One can check that this subset is a submanifold and the induced manifold structure does not depend on the chosen presentation of the coordinate ring. If G is an algebraic group, then by transport of structure the manifold G^{an} is a Lie group. The analytification functor for complex algebraic groups is defined similarly. □

EXAMPLE 1.7. Each of the real Lie groups in example 1.2 is the analytification of a real algebraic group in a natural way, e.g. $U(1)$ has the underlying real algebraic group given by

$$\mathscr{O}(U(1)) \ \simeq \ \mathbb{R}[x, y]/(x^2 + y^2 - 1).$$

After complexification we get an isomorphism of algebraic groups $Gl_{1,\mathbb{C}} \simeq U(1)_{\mathbb{C}}$ since

$$\mathbb{C}[t, t^{-1}] \xrightarrow{\sim} \mathbb{C}[x, y]/(x^2 + y^2 - 1) \quad \text{via} \quad t^{\pm 1} \mapsto x \pm iy.$$

However, this isomorphism cannot be defined over the reals: The Lie group $U(1)$ is compact while the group $Gl_1(\mathbb{R}) = \mathbb{R}^{\times}$ is not. Similarly you may check as an exercise that

$$U(n)_{\mathbb{C}} \simeq Gl_n(\mathbb{C}), \quad SU(n)_{\mathbb{C}} \simeq Sl_n(\mathbb{C}) \quad \text{and} \quad Sp(n)_{\mathbb{C}} \simeq Sp_{2n}(\mathbb{C}),$$

but none of these are induced by an isomorphism of the underlying real Lie groups.

REMARK 1.8. (a) Via the inclusion $\mathbb{R} \subset \mathbb{C}$ any complex algebraic group is also a real algebraic group and the analytification functor is left adjoint to the forgetful functor

$$\Big\{ \text{affine algebraic groups} \,/\, \mathbb{C} \Big\} \ \longrightarrow \ \Big\{ \text{affine algebraic groups} \,/\, \mathbb{R} \Big\},$$

i.e. for any affine algebraic groups G over $\mathbb{R}$ and H over $\mathbb{C}$ there exists a natural isomorphism

$$\mathrm{Hom}_{\mathbb{R}}(G, H) \xrightarrow{\sim} \mathrm{Hom}_{\mathbb{C}}(G_{\mathbb{C}}, H).$$

By considering universal coverings of Lie groups, one may show that such a left adjoint also exists for the forgetful functor from complex to real Lie groups, yielding the dashed arrow making the diagram in lemma 1.6 commute.

(b) Neither the complexification functor $(-)_{\mathbb{C}}$ nor the analytification $(-)^{an}$ is an equivalence of categories. However, a result of Chevalley says that on the full subcategory of real algebraic groups with *compact* analytification, these functors induce equivalences

$$\begin{array}{ccc} \Big\{ \text{affine alg gps } G \text{ over } \mathbb{R} \mid G^{an} \text{ compact} \Big\} & \xrightarrow[\cong]{(-)_{\mathbb{C}}} & \Big\{ \text{reductive alg gps} / \mathbb{C} \Big\} \\ {\scriptstyle (-)^{an}} \Big\downarrow {\scriptstyle \cong} & & \\ \Big\{ \text{compact real Lie groups} \Big\} & & \end{array}$$

where an affine algebraic group over $\mathbb{C}$ is called *reductive* if all its representations are semisimple. Here we use the following convention that will remain valid throughout the rest of the chapter:

DEFINITION 1.9. Let $\mathscr{C}$ be any of the categories from above, assumed to be clear from the context in what follows. By a *representation* of a group object G in $\mathscr{C}$ we mean a pair (V, ρ) where V is a finite dimensional complex vector space and where

$$\rho: \quad G \longrightarrow Gl(V)$$

is a homomorphism of group objects in the chosen category $\mathscr{C}$. So for Lie groups ρ is assumed to be smooth, for algebraic groups it is given by polynomials, etc.

With this convention, Chevalley's result implies that for any compact real Lie group $K = G^{an}$ with underlying real algebraic group G, the above functors induce an equivalence of categories

$$\mathrm{Rep}_{\mathbb{C}}(K) \xrightarrow{\sim} \mathrm{Rep}_{\mathbb{C}}(G_{\mathbb{C}})$$

between the categories of representations of the compact real Lie group and the associated complex algebraic group. For the rest of this chapter we therefore focus on representations of compact real Lie groups.

2. Haar measure, semisimplicity and characters

The key ingredient in the proof of semisimplicity for the representations of finite groups was the idea of averaging over the group. If one wants to extend this to the case of infinite Lie groups, one needs to replace the summation by an integral, so let us briefly recall some basic notions about integration on manifolds.

REMARK 2.1. A smooth manifold M is said to be *orientable* if it satisfies the following equivalent conditions:

1. There exists an atlas for M such that the transition functions between all coordinate charts have positive Jacobian determinant.
2. Let $n = \dim(M)$, then the top exterior power $\mathrm{Alt}^n(\mathscr{T}_M^*)$ of the cotangent bundle has a nowhere vanishing smooth section.

In this situation, any nowhere vanishing smooth section ω_M of $\mathrm{Alt}^n(\mathscr{T}_M^*)$ will be called a *volume form* and defines an *orientation* on the manifold as follows: For any open subsets $U \subseteq M$, $V \subseteq \mathbb{R}^n$ we say that a coordinate chart $\varphi : U \xrightarrow{\sim} V$ is *oriented* if the standard volume form on $\mathbb{R}^n$ pulls back to the chosen volume form on the chart,

$$\varphi^*(dx_1 \wedge \cdots dx_n) \;=\; c \cdot \omega_M \quad \text{for some} \quad c \;>\; 0.$$

The collection of all oriented charts then forms an atlas for which all transition functions have positive Jacobian determinant. The choice of an orientation allows to define the integral over any compactly supported top differential form ω on M as follows: If $\mathrm{Supp}(\omega) \subset U$ for some oriented coordinate chart (U, φ) as above, we define

$$\int_M \omega \;=\; \int_U (\varphi^{-1})^*(\omega).$$

By the change of variables formula this is invariant under coordinate changes with positive Jacobian determinant, hence it does not depend on the chosen oriented chart (U, φ). For a general compactly supported top differential form ω on M we take a partition of unity $1 = \sum_{i \in I} \varepsilon_i$ subordinate to an open cover by oriented charts (U_i, φ_i) and put

$$\int_M \omega \;=\; \sum_{i \in I} \int_{U_i} (\varphi_i^{-1})^*(\varepsilon_i \cdot \omega),$$

which again only depends on the chosen orientation. □

Now let $M = G$ be a real Lie group. The left and right multiplication by $h \in G$ are diffeomorphisms

$$\begin{aligned} \lambda_h : &\quad G \longrightarrow G, \quad g \mapsto hg \\ \rho_h : &\quad G \longrightarrow G, \quad g \mapsto gh \end{aligned}$$

of the underlying smooth manifold. We say that a differential form ω on the group is *left invariant* if $\lambda_h^*(\omega) = \omega$ for all $h \in G$, and *right invariant* if $\rho_h^*(\omega) = \omega$ for all $h \in G$. This notion allows to single out a distinguished volume form on any compact Lie group:

LEMMA 2.2. *Up to multiplication by a scalar $c \in \mathbb{R}^\times$ there exists a unique left invariant and a unique right invariant volume form on any Lie group G.*

Proof. The action of G on itself by left translations is transitive, therefore any left invariant differential form ω of degree $n \in \mathbb{N}$ is determined uniquely by its value at $1 \in G$,

$$\omega(1) \;\in\; (\mathrm{Alt}^n(\mathscr{T}_G^*))_1 \;=\; \mathrm{Alt}^n(T_1^*(G))$$

For $n = \dim(G)$ the vector space on the right hand side is one-dimensional, hence up to a multiplication by a non-zero scalar there is at most one left invariant volume form. To see that such a volume form does exist, fix any vector $v \in \mathrm{Alt}^n(T_1^*(G))$ and define

$$\omega(g) \;=\; (d\lambda_g)^*(v) \;\in\; \mathrm{Alt}^n(T_g^*(G)) \quad \text{for} \quad g \;\in\; G,$$

using the differential

$$(d\lambda_g)^* : T_g^*(G) \to T_1^*(G).$$

Since the multiplication in a Lie group is a smooth morphism, the above defines a smooth section ω of $\mathrm{Alt}^n(\mathscr{T}_G^*)$ and by construction this is a left invariant volume form. The statement for right invariant volume forms follows similarly. □

Thus any Lie group is orientable. For the rest of this chapter we fix one of the two possible orientations preserved by all *left* translations; this last condition is automatic for connected Lie groups, but for disconnected ones it is a way of orienting the various connected components consistently. Note that working instead with *right* translations would in general give a different orientation convention, as you may check for the orthogonal group $G = O(2)$ as an exercise!

Corollary 2.3. *If G is compact, then there is a unique left invariant volume form $\omega = dg$ such that in the orientation chosen above,*

$$\int_G dg \;=\; +1.$$

Proof. Since G is compact, it can be covered by finitely many compact subsets of oriented charts and hence $0 \;<\; |\int_G \omega| \;<\; \infty$ for any volume form ω. So the claim follows from the previous lemma. □

Definition 2.4. The *Haar measure* on a compact Lie group G is defined to be the functional

$$\mathscr{C}(G,\mathbb{C}) \;\longrightarrow\; \mathbb{C}, \quad F \;\mapsto\; \int_G F(g)\,dg$$

on continuous complex valued functions, using the above volume form dg.

Corollary 2.5. *The Haar measure is invariant under left and right translations and the inverse map, i.e.*

$$\int_G F(g)dg \;=\; \int_G F(g^{-1})dg \;=\; \int_G F(fgh)dg \quad \textit{for all} \quad f,h \;\in\; G.$$

Proof. Invariance under left translations is clear by definition. Since left and right translations in the group commute, for any $h \in G$ the form $\rho_h^*(dg)$ is also left invariant and so

$$\rho_h^*(dg) \;=\; c(h)\cdot dg \quad \text{for some} \quad c(h) \;\in\; \mathbb{R}^\times$$

by the previous lemma. Clearly the map $c : G \to \mathbb{R}^\times$ is a continuous group homomorphism, so its image lies inside the maximal compact subgroup $\{\pm 1\} \subset \mathbb{R}^\times$ since continuous images of compact sets are compact. So

$$c(h) \;=\; \begin{cases} +1 \\ -1 \end{cases} \quad \text{if } \rho_h : G \to G \text{ is} \quad \begin{cases} \text{orientation-preserving,} \\ \text{orientation-reversing,} \end{cases}$$

and in both cases

$$\int_G F(gh)dg \;=\; \int_G F(g)dg$$

by the change of variables formula, since in the orientation-reversing case two minus signs cancel out. The argument for the inverse map is similar. □

Let us now prove the analog of Maschke's theorem in the setting of compact Lie groups. Let V be a representation of a Lie group G. By a Hermitian inner product on the underlying complex vector space we mean a positive definite sesquilinear form

$$\langle\cdot,\cdot\rangle : \quad V \times V \;\longrightarrow\; \mathbb{C},$$

and call it G*-invariant* if

$$\langle gu, gv\rangle \;=\; \langle u,v\rangle \quad \text{for all} \quad u,v \;\in\; V \quad \text{and all} \quad g \;\in\; G.$$

If $\overline{V} \in \mathrm{Rep}_{\mathbb{C}}(G)$ denotes the complex conjugate representation, such a product can be viewed as an isomorphism of representations

$$\overline{V} \;\xrightarrow{\sim}\; V^* \quad \text{given by} \quad u \;\mapsto\; \langle u, -\rangle.$$

Averaging over the group we get

PROPOSITION 2.6. *If G is a compact Lie group, any representation $V \in \mathrm{Rep}_{\mathbb{C}}(G)$ admits a G-invariant Hermitian inner product.*

Proof. Take any Hermitian inner product $(\cdot,\cdot) : V \times V \longrightarrow \mathbb{C}$ on the complex vector space. For any $u, v \in V$ the map

$$G \longrightarrow \mathbb{C}, \quad g \mapsto (gu, gv)$$

is continuous, as by definition of a Lie group representation the action $G \times V \to V$ is smooth. So for $u, v \in V$ put

$$\langle u, v\rangle = \int_G (gu, gv)\, dg.$$

This is again a sesquilinear pairing since the integral is linear. Since the integral over strictly positive real function is strictly positive, the pairing is positive definite, and it is G-invariant because of the invariance of the Haar measure. □

COROLLARY 2.7. *If G is a compact Lie group, any $V \in \mathrm{Rep}_{\mathbb{C}}(G)$ is semisimple.*

Proof. Fix a G-invariant Hermitian inner product $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{C}$. Then for any subrepresentation $U \subset V$ the orthocomplement

$$U^{\perp} = \{\, v \in V \mid \langle u, v\rangle = 0 \text{ for all } u \in U \,\} \subseteq V$$

is stable under the action of G and $V = U \oplus U^{\perp}$ by construction. □

Recall that by definition the *character* of a representation $(V, \rho) \in \mathrm{Rep}_{\mathbb{C}}(G)$ is the function

$$\chi_V : \quad G \longrightarrow \mathbb{C}, \quad g \mapsto tr(\rho(g)).$$

This function is invariant under conjugation, and for Lie group representations it is continuous — even smooth, but this will not be important for us. If G is compact, the space $\mathscr{C}(G)$ of continuous functions $f : G \longrightarrow \mathbb{C}$ comes with the Hermitian product

$$\langle f, h\rangle = \int_G \overline{f(g)} h(g) dg \quad \text{for} \quad f, h \in \mathscr{C}(G),$$

and like for finite groups we obtain

PROPOSITION 2.8. *If G is a compact Lie group and $V, W \in \mathrm{Rep}_{\mathbb{C}}(G)$, then*

(1) $\dim_{\mathbb{C}} V^G = \int_G \chi_V(g) dg$.
(2) $\langle \chi_V, \chi_W\rangle = \dim_{\mathbb{C}} \mathrm{Hom}_G(V, W)$.

Proof. For the first part we claim that the projection onto the invariants V^G is given by the averaging operator

$$p : \quad V \longrightarrow V \quad \text{sending} \quad v \mapsto \int_G gv\, dg,$$

where the integral over vector valued functions is defined by choosing a basis of V and integrating each coordinate functions separately. By definition this integral is compatible with linear maps, so for any fixed $h \in G$ we have $h \cdot p(v) = p(h \cdot v)$ and therefore $p \in \mathrm{End}_G(V)$. Since the Haar measure is invariant under translations, we also see that

$$p(V) \subseteq V^G \quad \text{and} \quad p|_{V^G} = id,$$

which proves the claim and therefore

$$\dim_{\mathbb{C}} V^G = tr(p) = \int_G tr(g)\, dg = \int_G \chi_V(g)\, dg$$

as required. The second part of the proposition follows formally from the first part since

$$\dim_{\mathbb{C}} \mathrm{Hom}_G(V,W) \;=\; \dim_{\mathbb{C}} (V^* \otimes W)^G \;=\; \int_G \chi_{V^*\otimes W}(g)\, dg \;=\; \langle \chi_V, \chi_W \rangle$$

by adjunction. □

Corollary 2.9. *Let G be a compact Lie group. Then*

1. *the characters of the irreducible representations of G are orthonormal,*
2. *a representation $V \in \mathrm{Rep}_{\mathbb{C}}(G)$ is irreducible iff $\langle \chi_V, \chi_V \rangle = 1$,*
3. *any representation is determined by its character: For $V, W \in \mathrm{Rep}_{\mathbb{C}}(G)$ we have*

$$V \;\simeq\; W \quad\Longleftrightarrow\quad \chi_V \;=\; \chi_W$$

Proof. The first part follows from the above proposition since $\mathrm{End}_G(V) = \mathbb{C} \cdot id$ for any irreducible $V \in \mathrm{Rep}_{\mathbb{C}}(G)$ by Schur's lemma. The second and third parts then follow by semisimplicity. □

3. The Peter-Weyl theorem

For finite groups we have seen that every irreducible representation enters the regular representation. In the more general setting of compact Lie groups G a naive idea would be to consider instead the space $\mathscr{C}(G,\mathbb{C})$ of continuous complex-valued functions on the group. We consider this space as a representation of $G \times G$ with the action

$$((g_1,g_2)(f))(g) \;=\; f(g_1^{-1} g g_2) \quad \text{for} \quad g_1, g_2, g \in G,\; f \in \mathscr{C}(G,\mathbb{C}),$$

combining both the left and the right regular representation on functions. Like for finite groups we get

Lemma 3.1. *If $\mathrm{Irr}_{\mathbb{C}}(G) \subset \mathrm{Rep}_{\mathbb{C}}(G)$ denotes a representative system for the isomorphism classes of irreducible representations, we have a $G \times G$-equivariant embedding*

$$\iota : \bigoplus_{V \in \mathrm{Irr}_{\mathbb{C}}(G)} V^* \otimes V \;\hookrightarrow\; \mathscr{C}(G,\mathbb{C}).$$

Proof. We define $\iota : V^* \otimes V \to \mathscr{C}(G,\mathbb{C})$ on simple tensors $f \otimes v \in V^* \otimes V$ by putting

$$\iota(f \otimes v)(g) \;=\; f(gv) \quad \text{for} \quad g \;\in\; G.$$

Then ι is an equivariant map with respect to the action of the product $G \times G$ since we have

$$\begin{aligned} ((g_1,g_2)\cdot \iota(f\otimes v))(g) \;&=\; \iota(f\otimes v)(g_1^{-1} g g_2) \\ &=\; f(g_1^{-1} g g_2 \cdot v) \\ &=\; (g_1 \cdot f)(g g_2 \cdot v) \\ &=\; \iota((g_1\cdot f) \otimes (g_2 \cdot v))(g) \\ &=\; \iota((g_1,g_2)\cdot (f\otimes v))(g) \end{aligned}$$

for $g_1, g_2, g \in G$. Since ι is clearly not the zero map, it must then be an embedding because the representation $V^* \otimes V \in \mathrm{Rep}_{\mathbb{C}}(G \times G)$ is irreducible. □

However, in contrast with the situation for finite groups, the above inclusion is not an equality if $\dim(G) > 0$, since the algebraic direct sum on the left hand side has countable dimension while the target space has not. To get a better statement we consider $\mathscr{C}(G,\mathbb{C})$ as a pre-Hilbert space with respect to the Hermitian inner product

$$\langle f, h\rangle \;=\; \int_G \overline{f(g)}h(g)dg$$

given by the Haar measure of the previous section. Fix a G-invariant Hermitian inner product $\langle -,-\rangle_V$ on each V and make the identification

$$\overline{V} \;\xrightarrow{\sim}\; V^* \quad \text{via} \quad u \;\mapsto\; \langle u, -\rangle_V.$$

Then ι is given by

$$\iota: \quad \overline{V}\otimes V \;\hookrightarrow\; \mathscr{C}(G,\mathbb{C}), \quad u\otimes v \;\mapsto\; \rho_{uv}$$

where we put

$$\rho_{uv}(g) \;=\; \langle u, gv\rangle_V \;=\; \langle g^{-1}u, v\rangle_V \quad \text{for} \quad g \;\in\; G.$$

To explain this notation, note that if the vectors $u, v\in V$ are part of an orthonormal basis, then $\rho_{uv}(g)$ is the corresponding coefficient in the matrix for $\rho(g)\in \mathrm{End}_{\mathbb{C}}(V)$ in this basis. Summing up:

Remark 3.2. The image of the embedding $\iota : \bigoplus_{V\in \mathrm{Irr}_{\mathbb{C}}(G)} V^*\otimes V \hookrightarrow \mathscr{C}(G,\mathbb{C})$ is spanned by the *matrix coefficients* of the irreducible representations.

The direct summands for the various irreducible representations are orthogonal to each other by the following

Lemma 3.3 (Schur orthogonality). *For $V = (V,\rho)$ and $W = (W,\sigma)$ in $\mathrm{Irr}_{\mathbb{C}}(G)$ we have*

$$\big\langle \rho_{uv}, \sigma_{wx}\big\rangle \;=\; \begin{cases} 0 & \text{if } V\not\simeq W,\\ \frac{1}{\dim V}\cdot \langle w,u\rangle_V\langle v,x\rangle_V & \text{if } V = W.\end{cases}$$

Proof. Put $M = \langle v, -\rangle_V\cdot x\in \mathrm{Hom}_{\mathbb{C}}(V,W)$. Averaging over the matrix-valued function

$$G \;\ni\; g \;\mapsto\; \sigma(g)\circ M\circ \rho(g^{-1}) \;\in\; \mathrm{Hom}_{\mathbb{C}}(V,W)$$

we get that

$$\mathbb{M} \;=\; \int_G \sigma(g)\circ M\circ\rho(g^{-1})\,dg \;\in\; \mathrm{Hom}_G(V,W) \;=\; \begin{cases} 0 & \text{if } V\not\simeq W,\\ \mathbb{C}\cdot id_V & \text{if } V = W.\end{cases}$$

By construction

$$\begin{aligned}\big\langle w, \mathbb{M}(u)\big\rangle_W \;&=\; \int_G \langle w, gMg^{-1}u\rangle_W\,dg\\ &=\; \int_G \langle v, g^{-1}u\rangle_V\cdot\langle w, gx\rangle_W\,dg\\ &=\; \int_G \overline{\langle g^{-1}u, v\rangle_V}\cdot\langle w, gx\rangle_W\,dg \;=\; \big\langle \rho_{uv}, \sigma_{xw}\big\rangle.\end{aligned}$$

So for $V\not\simeq W$ we are done since then the left hand side vanishes, while for $V = W$ it only remains to show

$$\big\langle w, \mathbb{M}(u)\big\rangle_V \;=\; \frac{1}{\dim V}\cdot\langle w,u\rangle_V\langle v,x\rangle.$$

Now $\mathbb{M} = c\cdot id$ for some scalar $c\in\mathbb{C}$ that can be computed from the trace $tr(\mathbb{M})$ via

$$c\cdot\dim V \;=\; tr(\mathbb{M}),$$

so the claim reduces to

$$tr(\mathbb{M}) \;=\; \int_G tr(gMg^{-1})\,dg \;=\; tr(M) \;=\; tr\Big(y \mapsto \langle v, y\rangle_V \cdot x\Big) \;=\; \langle v, x\rangle_V$$

where the last equality follows by computing the trace in some orthonormal basis containing the vector x. □

We have already noted that unlike in the case of finite groups, ι is not an isomorphism if $\dim G > 0$. Let us look at a simple example:

EXAMPLE 3.4. Let $G = U(1)$ be the unit circle. Since this is an abelian group, any irreducible representation of it must be one-dimensional, and you may check as an exercise that every continuous homomorphism $\rho : G \to \mathbb{C}^\times$ has the form $z \mapsto z^n$ for a unique $n \in \mathbb{Z}$, so

$$\mathrm{Irr}_{\mathbb{C}}(G) \;=\; \mathbb{Z}.$$

On the other hand, in terms of the universal covering map $\mathbb{R} \twoheadrightarrow G, x \mapsto e^{2\pi i x}$ we can identify

$$\mathscr{C}(G, \mathbb{C}) \;=\; \big\{\, f \in \mathscr{C}(\mathbb{R}, \mathbb{C}) \mid f(x+1) = f(x)\ \forall x \in \mathbb{R} \,\big\}$$

with the space of all continuous periodic functions with period one, and the image of ι with the subspace

$$\Big\{ \sum_{n=-N}^{N} a_n e^{2\pi i n x} \mid N \in \mathbb{N}, a_n \in \mathbb{C} \Big\}$$

of *trigonometric polynomials*. It is a basic result of Fourier analysis on the circle that this subspace is dense with respect to the $\mathscr{L}^2$ norm and that with a suitable normalization the functions $e^{2\pi i n x}$ form a Hilbert space basis of $\mathscr{L}^2(\mathbb{R}/\mathbb{Z})$.

We can now do the same for any compact Lie group G. Considering $\mathscr{C}(G, \mathbb{C})$ as a pre-Hilbert space under

$$\langle f, h\rangle \;=\; \int_G \overline{f(g)} h(g) dg,$$

we denote by

$$\mathscr{L}^2(G) = \mathscr{L}^2(G, dg)$$

its *Hilbert space completion*. We endow this Hilbert space with the action of $G \times G$ given by the left and right regular representations as before. This is the proper generalization of the group algebra:

THEOREM 3.5 (Peter-Weyl). *We have a $G \times G$-equivariant decomposition as a Hilbert space direct sum*

$$\mathscr{L}^2(G) \;\simeq\; \widehat{\bigoplus}_{V \in \mathrm{Irr}_{\mathbb{C}}(G)} V^* \otimes V.$$

Proof. Recall that the Hilbert space direct sum is by definition the completion of the algebraic orthogonal direct sum. This latter algebraic direct sum is identified with the image of ι under the embedding $\mathscr{C}(G, \mathbb{C}) \subset \mathscr{L}^2(G)$, hence in view of the above it only remains to be shown that this image is dense inside $\mathscr{L}^2(G)$, or equivalently in $\mathscr{C}(G, \mathbb{C})$.

For the proof of this density result we assume that G is a closed subgroup of some $Gl(V)$, or equivalently that it has a *faithful* representation $V \in \mathrm{Rep}_{\mathbb{C}}(G)$. In fact one can prove the same result in general and deduce a posteriori that *every* compact Lie group is a matrix group, but we will not do this here. The advantage of working only with matrix groups is that one can use the

Stone-Weierstrass Theorem. Let X be a compact metrizable topological space and $\mathscr{A} \subset \mathscr{C}(X, \mathbb{C})$ a $\mathbb{C}$-subalgebra such that

- $\mathscr{A}$ is stable under conjugation: $\forall f \in \mathscr{A}$ one has $\overline{f} \in \mathscr{A}$,
- $\mathscr{A}$ separates points: $\forall x \neq y \in X \, \exists f \in \mathscr{A}$ with $f(x) \neq f(y)$,
- $\mathscr{A}$ vanishes nowhere: $\forall x \in X \, \exists f \in \mathscr{A}$ with $f(x) \neq 0$.

Then $\mathscr{A}$ is dense in $\mathscr{C}(X, \mathbb{C})$ with respect to the norm $\|f\|_\infty = \sup_{x \in X} |f(x)|$.

We apply the theorem of Stone-Weierstrass to the compact Lie group $X = G$ and the complex vector space $\mathscr{A}$ which is spanned by the matrix coefficients of all its representations. Note that this vector space is in fact a subalgebra: Suppose that $(V, \rho), (W, \sigma) \in \mathrm{Rep}_\mathbb{C}(G)$ and that the vectors $u, v \in V$ and $w, x \in W$ are part of an orthonormal basis on the respective spaces for some G-invariant Hermitian inner product, then the product

$$\rho_{uv} \cdot \sigma_{wx} \;=\; (\rho \otimes \sigma)_{(u \otimes w)(v \otimes x)}$$

is a matrix coefficient for the tensor product of the two representations, hence again in $\mathscr{A}$. Similarly, since the dual of a representation is given by the complex conjugate matrix, $\mathscr{A}$ is stable under complex conjugation, and $\mathscr{A}$ vanishes nowhere because representations are given by invertible matrices. Finally, the existence of a faithful representation implies that $\mathscr{A}$ separates points, so we are done. □

We have already seen earlier that the characters χ_V of all $V \in \mathrm{Irr}_\mathbb{C}(G)$ form an orthonormal set. We can now show by harmonic analysis that when considered as a subset of

$$\mathscr{L}^2(G)^G \;=\; \{f \in \mathscr{L}^2(G) \mid f(g^{-1}xg) = f(x) \forall g, x \in G\},$$

this orthonormal set is complete:

Corollary 3.6. *For any $f \in \mathscr{L}^2(G)^G$ we have an $\mathscr{L}^2$-convergent expansion as a series*

$$f \;=\; \sum_{V \in \mathrm{Irr}_\mathbb{C}(G)} \langle \chi_V, f \rangle \cdot \chi_V$$

and Plancherel's formula

$$\langle f, f \rangle \;=\; \sum_{V \in \mathrm{Irr}_\mathbb{C}(G)} |\langle \chi_V, f \rangle|^2$$

Proof. For each $V \in \mathrm{Irr}_\mathbb{C}(G)$ pick an orthonormal basis of vectors e_{Vi} with respect to some invariant Hermitian inner product. Then by the Peter-Weyl theorem and Schur orthogonality the functions

$$e_{Vik} \;=\; \frac{1}{\sqrt{\dim V}} \cdot \rho_{e_{Vi}, e_{Vk}} \quad \text{with} \quad 1 \leq i, k \;\leq\; \dim V \quad \text{and} \quad V \in \mathrm{Irr}_\mathbb{C}(G)$$

form an orthonormal basis for the Hilbert space $\mathscr{L}^2(G)$. Thus any $f \in \mathscr{L}^2(G)$ has an $\mathscr{L}^2$-convergent expansion

$$f \;=\; \sum_{V \in \mathrm{Irr}_\mathbb{C}(G)} \sum_{i,k=1}^{\dim V} c_{Vik} \cdot e_{Vik} \quad \text{with unique} \quad c_{Vik} \;\in\; \mathbb{C}.$$

and by uniqueness such a function is invariant under the conjugation by $g \in G$ iff for each $V \in \mathrm{Irr}_\mathbb{C}(G)$,

$$\sum_{i,k=1}^{\dim V} c_{Vik} \cdot e_{Vik} \;\in\; \iota\Big((V^* \otimes V)^G\Big)$$

is invariant. Since under the identifications

$$\begin{aligned} \overline{V} \otimes V &\simeq \mathrm{End}_{\mathbb{C}}(V) \\ (\overline{V} \otimes V)^G &\simeq \mathrm{End}_G(V) \;=\; \mathbb{C} \cdot id_V \end{aligned}$$

the identity id_V is the image of the vector $v = \sum_{i=1} e_{Vi} \otimes e_{Vi} \in \overline{V} \otimes V$ with corresponding function

$$\iota(v)(g) \;=\; \sum_{i=1}^{\dim V} \langle e_{Vi}, g \cdot e_{Vi} \rangle \;=\; \sum_{i=1}^{\dim V} \rho_{e_{Vi}, e_{Vi}} \;=\; tr(\rho_V(g)) \;=\; \chi_V(g),$$

the characters of the irreducible representations span a dense subspace of $\mathscr{L}^2(G)^G$ and we are done by orthonormality. □

4. Maximal tori and weights

We have seen that for a compact Lie group G every irreducible $V \in \mathrm{Irr}_{\mathbb{C}}(G)$ can be found inside $\mathscr{L}^2(G)$, but we really want an explicit and more algebraic parametrization of the irreducibles. Recall that in Schur-Weyl duality for $Gl_n(\mathbb{C})$ we characterized the Schur functors via the action of diagonal matrices. In the setting of compact Lie groups the role of diagonal matrices is taken by compact tori

$$T \;\simeq\; U(1) \times \cdots \times U(1).$$

The representation theory of such tori is extremely simple: Since T is an abelian group, it follows from Schur's lemma that every irreducible representation of T is one-dimensional, and we have

LEMMA 4.1. *Any irreducible representation of the compact torus $T = U(1)^r$ has the form*

$$\rho: \quad T \longrightarrow U(1), \quad z \mapsto z^a \;=\; z_1^{a_1} \cdots z_r^{a_r} \quad \textit{with unique} \quad a_i \in \mathbb{Z}.$$

Proof. It suffices to prove this for $r = 1$, so suppose that $\rho : T = \mathbb{R}/\mathbb{Z} \longrightarrow \mathbb{C}^\times$ is a homomorphism of Lie groups. By continuity the image of this homomorphism is a compact subgroup of $\mathbb{C}^\times$ and hence contained in $U(1)$. Again by continuity there exists $\epsilon > 0$ with

$$\mathrm{Re}\Big(\rho\left(x \bmod \mathbb{Z}\right)\Big) > 0 \quad \text{for all} \quad x \;\in\; I(2\epsilon) \;=\; (-2\epsilon, 2\epsilon) \;\subset\; \mathbb{R}.$$

Then there is a unique lift

$$\varphi: \quad I(2\epsilon) \longrightarrow \mathbb{R} \quad \text{with} \quad \rho|_{I(2\epsilon)} \;=\; e^{2\pi i \varphi} \quad \text{and} \quad \varphi(0) \;=\; 0.$$

Since ρ is a group homomorphism, we have $\varphi(x+y) = \varphi(x) + \varphi(y)$ for all $x, y \in I(\epsilon)$ and may therefore extend φ from the chosen small interval to the whole real line by putting

$$\varphi(x) \;:=\; \lim_{n\to\infty} n \cdot \varphi(\tfrac{x}{n}) \quad \text{for any} \quad x \;\in\; \mathbb{R}.$$

This extension φ is additive, hence $\mathbb{Q}$-linear and then by continuity even $\mathbb{R}$-linear, and

$$\rho\big(x \bmod \mathbb{Z}\big) \;=\; e^{2\pi i \varphi(x)} \quad \text{for all} \quad x \;\in\; \mathbb{R}.$$

Hence $\varphi(x) = ax$ for a unique integer $a \in \mathbb{Z}$ and the claim follows. □

Definition 4.2. For a compact torus T we define the group of its *characters* and *cocharacters* by

$$X^*(T) \;=\; \mathrm{Hom}_{cont}(T, U(1)) \;\simeq\; \mathbb{Z}^r \quad \text{and} \quad X_*(T) \;=\; \mathrm{Hom}_{cont}(U(1), T) \;\simeq\; \mathbb{Z}^r$$

respectively, where as usual all homomorphisms are assumed to be continuous. The number r is called the *rank* of the torus. The above lemma implies that composition gives a perfect pairing

$$\langle\cdot,\cdot\rangle: \quad X^*(T)\times X_*(T) \;\longrightarrow\; X^*(U(1)) \;=\; \mathbb{Z}.$$

We want to describe the representations of a compact Lie group G in terms of their restriction to subtori $T \subseteq G$. Such subtori can be specified by a single element of the group:

Lemma 4.3 (Kronecker). *Let T be a compact torus. Then for general $t \in T$ the subgroup*

$$\bigl\{t^n \in T \mid n \in \mathbb{Z}\bigr\} \;\subset\; T \quad \textit{is dense.}$$

Proof. Write $T = (\mathbb{R}/\mathbb{Z})^r$ for some $r \in \mathbb{N}$. For general $t = (t_1, \dots, t_r) \in \mathbb{R}^n$ the numbers $1, t_1, \dots, t_r$ are linearly independent over $\mathbb{Q}$, since there are only countably many possible linear relations with rational coefficients. We claim that any such t generates a dense subgroup of the given torus. Indeed, if this were not the case, then the closure

$$S \;=\; \overline{\bigl\{t^n \mid n \in \mathbb{Z}\bigr\}} \;\hookrightarrow\; T$$

would be a proper closed subgroup. But the quotient of any compact Lie group by a closed subgroup is a compact Lie group, hence in our case $Q = T/S \neq \{1\}$ would be a non-trivial compact abelian Lie group and as such it admits a non-trivial one-dimensional representation

$$\rho: \quad Q \;\longrightarrow\; U(1).$$

By the above lemma the pull-back of this representation under the map $T \twoheadrightarrow Q$ is of the form

$$\rho(x) \;=\; e^{2\pi i a\cdot x} \quad \text{for some} \quad a \in \mathbb{Z}^r \setminus \{0\}.$$

But then $\rho(t) = 1$ implies $a \cdot t \in \mathbb{Z}$ so that $1, t_1, \dots, t_r$ are not linearly independent over $\mathbb{Q}$. □

Definition 4.4. By a *subtorus* of a compact Lie group G we mean a closed subgroup $T \subseteq G$ isomorphic to a compact torus. A *maximal torus* is a subtorus that does not lie in any larger subtorus. Maximal tori clearly exist in any compact Lie group and we will soon see that any two or them are conjugate via some inner automorphism of the group. Let us take a look at the classical groups:

Example 4.5. (1) Inside $G = U(n)$ we have the subtorus $T \subset G$ of all diagonal matrices

$$diag(z_1, \dots, z_n) \quad \text{with} \quad z_1, \dots, z_n \;\in\; U(1).$$

This is a maximal torus because it is equal to its own centralizer, $T = Z_G(T)$.

(2) Inside $G = SO(2n)$ we have the torus T of all block diagonal matrices of the form

$$diag(B_1, \dots, B_n) \quad \text{with} \quad B_i \;=\; \begin{pmatrix} \cos(\varphi_i) & \sin(\varphi_i) \\ -\sin(\varphi_i) & \cos(\varphi_i) \end{pmatrix} \quad \text{for} \quad \varphi_i \;\in\; \mathbb{R}/2\pi\mathbb{Z}.$$

Similarly, inside $G = SO(2n+1)$ we have the torus of block diagonal matrices of the form

$$diag(B_1, \dots, B_n, 1).$$

In both cases the subtori are maximal since they are their own centralizer.

(3) Inside $G = Sp(n) \subset Gl_n(\mathbb{H})$ we have the subtorus $T \subset G$ consisting of all diagonal matrices

$$diag(z_1, \dots, z_n) \quad \text{with} \quad z_1, \dots, z_n \in U(1) \subset \mathbb{H},$$

which again is a maximal torus because it is equal to its own centralizer.

In general there is no distinguished choice of a maximal torus in a compact Lie group. However, the following result says that any of them will do, which explains why maximal tori are so important in representation theory:

Cartan's theorem. *If G is a compact Lie group, $g \in G$ and $T \subset G$ a maximal torus, then*

$$hgh^{-1} \in T \quad \textit{for some} \quad h \in G.$$

This is a deep theorem, the shortest proof is due to A. Weil and uses a topological fixed point formula as discussed in the appendix. We here only illustrate a special case:

Example 4.6. For $G = U(n)$, Cartan's theorem says that any unitary matrix can be diagonalized via a unitary matrix. This can be checked directly: Unitary matrices $A \in U(n)$ are *normal*, i.e.

$$\overline{A}^t A = A \overline{A}^t$$

and the spectral theorem from linear algebra says that for any normal matrix A there is a basis $v_1, \dots, v_n \in \mathbb{C}^n$ which is orthonormal with respect to the standard Hermitian inner product and satisfies $Av_i = t_i \cdot v_i$ for some $t_i \in \mathbb{R}$. But then we have

$$M^{-1}AM = diag(t_1, \dots, t_n) \quad \text{where} \quad M = (v_1, \dots, v_n) \in U(n).$$

Returning to the case of arbitrary compact Lie groups G we get from Cartan's theorem the

Corollary 4.7. *Any two maximal tori in G are conjugate.*

Proof. Let $S, T \subset G$ be maximal tori. By Kronecker's lemma we may pick $s \in S$ which generates a dense subgroup inside S. Cartan's theorem gives an $h \in G$ such that $hsh^{-1} \in T$, and taking the closure of the subgroup generated by this element we get an inclusion $hSh^{-1} \subseteq T$. This last inclusion must be an equality because with S also hSh^{-1} is a maximal torus. □

Summing up, for the purpose of representation theory all choices of maximal tori are equivalent. In what follows we fix a maximal torus $T \subseteq G$ and put $X_* = X_*(T)$ and $X^* = X^*(T)$. Any representation $V \in \mathrm{Rep}_{\mathbb{C}}(G)$ restricts under this maximal torus to

$$V|_T = \bigoplus_{\chi \in X^*} V_\chi \quad \text{where} \quad V_\chi \simeq \chi^{\oplus \dim V_\chi}$$

and we call the occuring characters $\chi \in X^*$ with $V_\chi \neq \{0\}$ the *weights* of the representation with respect to the chosen maximal torus. These give a fingerprint of the representation: Let

$$R(G) = K^0(\mathrm{Rep}_{\mathbb{C}}(G))$$

be the free abelian group which is generated by the isomorphism classes in $\mathrm{Irr}_{\mathbb{C}}(G)$, or equivalently the group of all class functions which are $\mathbb{Z}$-linear combinations of characters χ_V with $V \in \mathrm{Rep}_{\mathbb{C}}(G)$. We view $R(G)$ as a ring with the product given by the tensor product of representations, or equivalently by the point-wise product of class functions, and call it the *representation ring* of G. Looking at weights we embed it into the group ring of X^* as follows:

PROPOSITION 4.8. *For any connected compact Lie group G one has a natural embedding of rings*

$$R(G) \hookrightarrow R(T) \;\simeq\; \mathbb{Z}[X^*], \quad [V] \;\mapsto\; \sum_{\chi \in X^*} \dim(V_\chi) \cdot [\chi].$$

Proof. The elements of the representation ring $R(G)$ are linear combinations of characters. But characters are invariant under conjugation, hence determined uniquely by their restriction to a maximal torus $T \subset G$ using corollary 4.7. So the restriction map $R(G) \to R(T)$ is an embedding, and the claim follows. □

The above embedding is a counterpart of the one in the previous chapter that sent representations of $Gl_n(\mathbb{C})$ to the polynomials given by the trace of the action of diagonal matrices. These latter were *symmetric* functions, and similarly the image of the above embedding has extra symmetries:

DEFINITION 4.9. Let G be a connected compact Lie group and $T \subset G$ a maximal torus. The conjugation action

$$\varphi: \quad G \;\longrightarrow\; Aut(G), \quad g \;\mapsto\; \big(x \mapsto gxg^{-1}\big)$$

moves the maximal torus around, but its restriction to the normalizer $N_G(T) \leq G$ gives an action

$$\varphi: \quad N_G(T) \;=\; \big\{g \in G \mid \varphi(g)(T) \subseteq T\big\} \;\longrightarrow\; Aut(T)$$

which is trivial precisely on the centralizer

$$Z_G(T) \;=\; \big\{g \in G \mid \varphi(g)|_T \;=\; id_T\big\}.$$

We define the *Weyl group* as the quotient

$$W(G,T) \;=\; N_G(T)/Z_G(T).$$

One may show that for a connected compact Lie group always $Z_G(T) = T$ but this will not be used in what follows. By construction the Weyl group acts faithfully on the maximal torus and hence also on its character group: For $w = [g] \in W(G,T)$ and $\chi \in X^*(T) = \mathrm{Hom}(T, U(1))$, define $w \cdot \chi \in X^*(T)$ by

$$G \xrightarrow{\varphi(g^{-1})} G \xrightarrow{\;\chi\;} U(1), \qquad G \xrightarrow{\;w\cdot\chi\;} U(1)$$

As a result of the above we obtain:

COROLLARY 4.10. *The Weyl group $W = W(G,T)$ is a finite group. It permutes the weights of any representation, and the above embedding factors over the subring of Weyl group invariants:*

$$R(G) \;\hookrightarrow\; \mathbb{Z}[X^*]^W$$

Proof. One may check that the normalizer $N_G(T)$ is a closed subset of G and hence compact. Thus $W(G,T)$ is compact as well when endowed with the quotient topology, so to see that it is finite it will be enough to show it is discrete. But this follows by noting that the conjugation action $N_G(T) \to Aut(X^*) = Gl_n(\mathbb{Z})$ is continuous with respect to the discrete topology on the target and that its image is precisely the quotient $W(G,T) = N_G(T)/Z_G(T)$.

The Weyl group permutes the weights of any $V \in \mathrm{Rep}_{\mathbb{C}}(G)$ because χ_V is a class function on G, so that $\chi_V|_T$ is invariant under conjugation by $N_G(T)$. □

We will see later that the embedding $R(G) \hookrightarrow \mathbb{Z}[X^*]^W$ is an isomorphism, and on the way we will develop a general formula for the images of the irreducible representations under this isomorphism. But let us first look at a concrete example:

EXAMPLE 4.11. Let $G = U(n)$. Since the maximal torus $T \subset G$ of unitary diagonal matrices is normalized by all permutation matrices, we have a natural embedding

$$\iota: \quad \mathfrak{S}_n \hookrightarrow N_G(T), \quad \sigma \mapsto M \quad \text{where} \quad M_{ij} = \begin{cases} 1 & \text{if } j = \sigma(i), \\ 0 & \text{otherwise.} \end{cases}$$

We claim that the composite map

$$\mathfrak{S}_n \hookrightarrow N_G(T) \twoheadrightarrow W(G,T) = N_G(T)/T$$

is an isomorphism: Indeed, the injectivity of this map is clear since $\iota(\mathfrak{S}_n) \cap T = \{1\}$; for surjectivity note that

$$\begin{aligned} g \in N_G(T) &\iff \forall s \in T \ \exists t \in T \text{ with } sg = gt \\ &\iff \begin{cases} \text{any rescaling of the rows of } g \text{ by scalars from } U(1) \\ \text{can be obtained via some rescaling of the columns} \end{cases} \\ &\iff \text{each row and column of } g \text{ contains only one entry} \neq 0 \end{aligned}$$

and the last condition says that g is the product of a permutation matrix and a diagonal matrix. Hence

$$W(G,T) \simeq \mathfrak{S}_d$$

and this group acts on

$$X^* = X^*(T) \simeq \mathbb{Z}^n$$

in the standard way by permuting the factors. We view the group ring $\mathbb{Z}[X^*]$ as a ring of Laurent polynomials in n variables $x_1, \dots, x_n$ by attaching to each character the corresponding Laurent monomial which describes the action of diagonal matrices. Then

$$\mathbb{Z}[X^*]^W \simeq \Big(\mathbb{Z}\big[x_1^{\pm 1}, \dots, x_n^{\pm 1}\big]\Big)^{\mathfrak{S}_n} = \mathbb{Z}\big[e_1, e_2, \dots, e_n, \frac{1}{e_n}\big]$$

becomes the ring of symmetric Laurent polynomials in these variables, which can be obtained from the ring of ordinary symmetric polynomials by inverting the determinant $e_n = x_1 \cdots x_n$. Like the corresponding map in the previous chapter, the character map

$$R(G) \hookrightarrow \Big(\mathbb{Z}\big[x_1^{\pm 1}, \dots, x_n^{\pm 1}\big]\Big)^{\mathfrak{S}_n} \quad \text{is given by} \quad [V] \mapsto \chi_V(diag(x_1, \dots, x_n)).$$

Thus if $V = \mathbb{C}^n$ denotes the standard representation of $G = U(n) \subset Gl_n(\mathbb{C})$, then for the Schur functors we get

$$R(G) \ni \big[S_\alpha(V)\big] \mapsto s_\alpha(x_1, \dots, x_n) \in \mathbb{Z}[e_1, \dots, e_n]$$

for any partition α of length $\ell(\alpha) \leq n$. As a result we obtain

COROLLARY 4.12. *For the irreducible representations of $G = U(n)$ we have a bijection*

$$\Big\{a = (a_1, \dots, a_n) \in \mathbb{Z}^n \simeq X^*(T) \ \Big| \ a_1 \geq \cdots \geq a_n\Big\} \xrightarrow{\sim} \mathrm{Irr}_{\mathbb{C}}(G)$$

given by

$$a \mapsto V_a = \det(V)^{\otimes a_n} \otimes S_\alpha(V) \quad \text{for} \quad \alpha = (a_1 - a_n, a_2 - a_n, \dots, a_{n-1} - a_n).$$

Proof. Since the ring $\mathbb{Z}[e_1, \dots, e_n]$ of symmetric polynomials is spanned by the Schur polynomials, it is clear that the images of the classes $[V_a] \in R(G)$ span the ring of symmetric Laurent polynomials; hence these classes a fortiori span the representation ring $R(G)$. Furthermore, the $V_a \in \mathrm{Rep}(G)$ are irreducible and pairwise non-isomorphic since they are so as representations of $G_{\mathbb{C}} = Gl_n(\mathbb{C})$; if one wants to avoid a reference to the previous chapter one can also check this directly from the linear independence of the Schur polynomials $s_\alpha(x_1, \dots, x_n)$ with α of length $\ell(\alpha) \leq n$ (exercise). $\square$

5. From Lie groups to Lie algebras

In order to generalize the previous example to arbitrary compact Lie groups G we need some basic structure results about such groups. We consider the action of the group on itself by conjugation

$$\varphi : \quad G \longrightarrow Aut(G), \quad g \mapsto \varphi_g = \big(x \mapsto gxg^{-1}\big).$$

For $g \in G$, taking the differential of $\varphi(g) : G \to G$ at the neutral element $x = 1 \in G$ we get an endomorphism

$$Ad_g = (D\varphi_g)(1) : \quad T_1(G) \longrightarrow T_1(G)$$

of the tangent space. This tangent space is called the *Lie algebra* of G and will be denoted

$$\mathfrak{g} = Lie(G) = T_1(G)$$

in what follows. The above endomorphisms for varying $g \in G$ fit together to a smooth homomorphism

$$Ad : \quad G \longrightarrow Aut_{\mathbb{R}}(\mathfrak{g}), \quad g \mapsto Ad_g$$

which is called the *adjoint action* of G. Differentiating once more, we obtain a linear map

$$ad : \quad \mathfrak{g} \longrightarrow \mathrm{End}_{\mathbb{R}}(\mathfrak{g}), \quad X \mapsto ad_X = (DAd)(X)$$

between the tangent spaces. Let us make this more explicit:

Remark 5.1. Choosing a faithful representation we may assume $G \subset Gl_n(\mathbb{C})$ for some $n \in \mathbb{N}$, so that $\mathfrak{g} \subset \mathrm{Mat}_{n\times n}(\mathbb{C})$ as a real vector subspace. In this case one computes

$$Ad_g(Y) = gYg^{-1} \quad \text{for} \quad g \in G \subset Gl_n(\mathbb{C}), \quad Y \in \mathfrak{g} \subset \mathrm{Mat}_{n\times n}(\mathbb{C}),$$

so the adjoint action is given by the conjugation of matrices. To compute ad, take any smooth test curve

$$g : \quad (-\epsilon, \epsilon) \longrightarrow G, \quad t \mapsto g(t) \quad \text{with} \quad g(0) = 1 \quad \text{and} \quad \tfrac{d}{dt} g(t)|_{t=0} = X.$$

Note that

$$\tfrac{d}{dt}(g(t)^{-1})|_{t=0} = -\tfrac{d}{dt} g(t)|_{t=0} = -X$$

by the product rule, since the derivative of the constant curve $t \mapsto g(t)^{-1} \cdot g(t) \equiv 1$ vanishes. So again by the product rule

$$ad_X(Y) = \tfrac{d}{dt}\big(g(t)Yg(t)^{-1}\big)|_{t=0} = XY - YX = [X, Y]$$

is given by the commutator of matrices. In particular, the tangent space $\mathfrak{g} = Lie(G)$ is indeed a Lie algebra in the following sense:

DEFINITION 5.2. A *Lie algebra* over a field k is a vector space $\mathfrak{g}$ over k endowed with a k-bilinear map

$$[\cdot,\cdot]: \quad \mathfrak{g}\times\mathfrak{g} \longrightarrow \mathfrak{g}$$

such that

$$[x,x]=0 \quad \text{and} \quad [x,[y,z]]+[y,[z,x]]+[z,[x,y]]=0 \quad \text{for all} \quad x,y,z\in\mathfrak{g}.$$

The last identity is called the *Jacobi identity.* The finite-dimensional Lie algebras over a given field k form an abelian category, where by a *homomorphism* between two Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ we mean a k-linear map $f:\mathfrak{g}\to\mathfrak{h}$ such that $[f(x),f(y)] = f([x,y])$ for all $x,y\in\mathfrak{g}$.

In the above remark we have verified the Jacobi identity for the Lie algebra of a Lie group via matrix representations, but for a more intrinsic interpretation of the bracket on $\mathfrak{g} = Lie(G)$ one can use that every tangent vector extends uniquely to a left invariant vector field. This gives an isomorphism

$$\big\{\text{left invariant vector fields on } G\big\} \xrightarrow{\sim} \mathfrak{g} \;=\; T_e(G), \quad (X_g)_{g\in G} \;\mapsto\; X_e$$

under which the bracket $[\cdot,\cdot]:\mathfrak{g}\times\mathfrak{g}\longrightarrow\mathfrak{g}$ corresponds to the commutator of vector fields. In what follows, for any given tangent vector $X\in\mathfrak{g}$ we denote by $(X_g)_{g\in G}$ the unique left invariant vector field with $X=X_e$. The above correspondence gives a way to go back from the Lie algebra to the Lie group:

PROPOSITION 5.3. *For any $X\in\mathfrak{g}=Lie(G)$ there is a unique homomorphism of Lie groups*

$$\alpha_X: \quad \mathbb{R} \longrightarrow G \quad \textit{with} \quad \frac{d\alpha_X}{dt}\Big|_{t=t_0} := \; (D\alpha_X(t_0))\big(\tfrac{d}{dt}\big) \;=\; X_{\alpha_X(t_0)} \quad \forall\, t_0 \,\in\, \mathbb{R},$$

where $\frac{d}{dt}\in T_{t_0}\mathbb{R}$ denotes the derivation by the standard coordinate on the real line.

Proof. We know from differential geometry that locally any smooth vector field admits unique integral curves, i.e. for sufficiently small $\epsilon>0$ there is a unique smooth map

$$\alpha_X: \quad (-\epsilon,\epsilon) \longrightarrow G$$

with

$$\alpha_X(0) \;=\; e \quad \text{and} \quad \tfrac{d\alpha_X}{dt}|_{t=t_0} \;=\; X_{\alpha_X}(t_0) \quad \forall\, t_0 \,\in\, (-\epsilon,\epsilon).$$

Then for any $g\in G$ the curve

$$g\alpha_X: \quad (-\epsilon,\epsilon) \longrightarrow \mathbb{R}, \quad t \;\mapsto\; g\cdot\alpha_X(t)$$

has the same differential property as the previous one but is centered at g rather than at e. Taking $g=\alpha_X(\pm\epsilon/2)$, we may by uniqueness of integral curves extend our curve smoothly to

$$\alpha_X: \quad (-3\epsilon/2, 3\epsilon/2) \longrightarrow G$$

by putting

$$\alpha_X(t) \;=\; \begin{cases} \alpha_X(\mp\epsilon/2)\cdot\alpha_X(t\pm\epsilon/2) & \text{for } t\pm\epsilon/2 \,\in\, (-\epsilon/2,\epsilon/2), \\ \alpha_X(t) & \text{else,} \end{cases}$$

indeed by uniqueness the various curves glue smoothly on their overlap. Continuing like this we obtain a smooth extension

$$\alpha_X: \quad \mathbb{R} \longrightarrow G$$

and the uniqueness implies that this is in fact a homomorphism, since for $t_0\in\mathbb{R}$ the curves $t\mapsto\alpha_X(t_0+t)$ and $t\mapsto\alpha_X(t_0)\cdot\alpha_X(t)$ are both integral curves of the same vector field centered at the same point. □

For $G = U(1)$ one computes that $\alpha_X : \mathbb{R} \to \mathbb{R}$ is given by $\alpha_X(t) = exp(2\pi t)$. Motivated by this example, we call

$$exp: \quad \mathfrak{g} \longrightarrow G, \quad exp(X) \;=\; \alpha_X(1)$$

the *exponential map* of the group. Its basic properties are summarized in

LEMMA 5.4. *The exponential map* $exp : \mathfrak{g} \longrightarrow G$ *is smooth, it restricts to a local diffeomorphism in a neighborhood of the origin, and for any homomorphism of Lie groups* $f : G \to H$ *of Lie groups we have a commutative diagram*

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{Lie(f)} & \mathfrak{h} \\ {\scriptstyle exp}\downarrow & & \downarrow{\scriptstyle exp} \\ G & \xrightarrow{f} & H \end{array}$$

Proof. The smooth dependence of integral curves on the initial values implies that exp is a smooth map. By construction its differential at the origin is the identity map $(Dexp)(e) = id : \mathfrak{g} = T_0(\mathfrak{g}) \longrightarrow \mathfrak{g}$, hence exp is a local diffeomorphism by the inverse function theorem. The commutativity of the diagram holds by uniqueness of integral curves. □

REMARK 5.5. For nonabelian Lie groups G the exponential map $exp : \mathfrak{g} \to G$ is not a homomorphism. However, the group structure on the connected component can be recovered in terms of the Lie algebra via the Baker-Campbell-Hausdorff formula

$$exp(x) \cdot exp(y) \;=\; exp\Big(x + y + \tfrac{1}{2}[x, y] + \cdots\Big)$$

where "$\cdots$" is a complicated power series in iterated commutators of x and y convergent for x, y in a small neighborhood of the origin. We will not use this in what follows but point to some other subtleties in the passage from Lie groups to Lie algebras, illustrated by the following examples:

- we have $SU(2) \not\simeq SO(3)$ but $Lie(SU(2)) \simeq Lie(SO(3))$;
- for the rank two compact torus $T = U(1) \times U(1)$, any linear subspace of the form

  $$\{(x, ax) \mid x \in \mathbb{R}\} \;\subset\; \mathbb{R}^2 \;=\; Lie(T) \quad \text{with} \quad a \;\in\; \mathbb{R}$$

  is a Lie subalgebra but comes from a closed Lie subgroup only if $a \in \mathbb{Q}$.

The situation is clarified by

THEOREM 5.6 (Lie's three theorems).

(1) *Let* G *be a Lie group with Lie algebra* $\mathfrak{g}$. *For any Lie subalgebra* $\mathfrak{h} \subseteq \mathfrak{g}$ *there exists a unique connected Lie group* H *together with an injective immersion* $H \hookrightarrow G$ *that induces an isomorphism* $Lie(H) \simeq \mathfrak{h}$.

(2) *If* G, H *are Lie groups with Lie algebras* $\mathfrak{g}, \mathfrak{h}$ *and if* G *is connected, then the natural map*

$$\mathrm{Hom}_{LieGps}(G, H) \;\longrightarrow\; \mathrm{Hom}_{LieAlg}(\mathfrak{g}, \mathfrak{h}), \quad f \;\mapsto\; Lie(f) \;=\; Df(e)$$

is injective. If G *is simply connected, then this map is an isomorphism.*

(3) *On connected simply connected Lie groups, the functor $G \mapsto Lie(G)$ is an equivalence of categories*

$$Lie(-): \quad \{\text{connected simply connected Lie groups}\} \longrightarrow \{\text{finite-dimensional real Lie algebras}\}.$$

Proof. (1) In differential geometry one proves the following higher-dimensional generalization of the existence and uniqueness of integral curves which is known as *Frobenius' theorem*: If M is a smooth manifold and $\mathscr{F} \subset \mathscr{T}_M$ a vector subbundle of the tangent bundle, the following are equivalent:

- $\mathscr{F}$ is *involutive*, i.e. stable under the Lie bracket $[\cdot,\cdot]: \mathscr{T}_M \times \mathscr{T}_M \longrightarrow \mathscr{T}_M$.
- $\mathscr{F}$ is *integrable*, i.e. for any $p_0 \in M$ there exists a neighborhood $U \ni p_0$ and a unique connected closed submanifold $N \subseteq U$ containing p_0 such that $T_pN = \mathscr{F}_p$ for all $p \in N$.

One also calls the submanifolds N the *leaves* of the *foliation* defined by $\mathscr{F}$. Note that any point $p_0 \in M$ lies one a unique maximal leaf of the foliation.

Now let $M = G$ be a Lie group with Lie algebra $\mathfrak{g}$. For any Lie subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ we consider the subbundle $\mathscr{F} \subset \mathscr{T}_G$ whose local sections are the left invariant vector fields with fibers

$$\mathscr{F}_g \;=\; \lambda_{g,*}(\mathfrak{h}) \;\subseteq\; \lambda_{g,*}(\mathfrak{g}) \;=\; T_gG \quad \text{for} \quad g \,\in\, G.$$

Since $[\mathfrak{h},\mathfrak{h}] \subseteq \mathfrak{h}$ and since the Lie bracket of vector fields is compatible with left translations, it is clear that the subbundle $\mathscr{F}$ is involutive. Hence by Frobenius' theorem it is integrable, and a similar glueing argument as for the exponential map shows that any leaf extends to a unique subset $H \subset G$ which is the image of an injective immersion $H \to G$ and whose points form a subgroup of G.

(2) By the lemma the map $exp: \mathfrak{g} \to G$ is a local diffeomorphism, hence its image contains some open neighborhood of $e \in G$. Every connected topological group is generated by any neighborhood of the neutral element, hence it follows that the image of exp generates G. But then by the commutative diagram in the lemma, any $f \in \mathrm{Hom}(G,H)$ is determined uniquely by $(Df)(e) \in \mathrm{Hom}(\mathfrak{g},\mathfrak{h})$.

Now assume that G is connected and simply connected, and let $\varphi: \mathfrak{g} \to \mathfrak{h}$ be any homomorphism of Lie algebras. We endow $\mathfrak{g} \times \mathfrak{h}$ with the Lie algebra structure given by $[g_1 + h_1, g_2 + h_2] := [g_1, g_2] + [h_1, h_2]$ for $g_i \in \mathfrak{g}$ and $h_i \in \mathfrak{h}$. Since φ is a Lie algebra homomorphism, the graph

$$\gamma \;=\; \{(g, \varphi(g)) \mid g \in \mathfrak{g}\} \;\subseteq\; \mathfrak{g} \times \mathfrak{h}$$

is a Lie subalgebra. So by (1) there is a unique connected Lie group Γ with an injective immersion

$$\Gamma \;\hookrightarrow\; G \times H \quad \text{inducing an isomorphism} \quad Lie(\Gamma) \;\simeq\; \gamma.$$

Composing this immersion with the projection on the first factor, we get a Lie group homomorphism

$$p_1: \quad \Gamma \;\hookrightarrow\; G \times H \;\twoheadrightarrow\; G$$

whose differential at the identity is the projection from the graph $\gamma \subseteq \mathfrak{g} \times \mathfrak{h}$ onto the first factor. Thus

$$Lie(p_1): \quad \gamma \;\xrightarrow{\sim}\; \mathfrak{g}$$

is an isomorphism, and one easily checks that then $p_1: \Gamma \twoheadrightarrow G$ is a topological covering map (exercise). Since G was assumed to be simply connected, it does not admit any non-trivial topological cover, hence p_1 must be an isomorphism. Then the composite

$$f \;:=\; p_2 \circ p_1^{-1}: \quad G \;\longrightarrow\; \Gamma \;\hookrightarrow\; G \times H \;\twoheadrightarrow\; H$$

is a homomorphism of Lie groups with $Lie(f) = \varphi$ as desired.

(3) The functor is fully faithful by the above, so it only remains to see that it is essentially surjective, i.e. that every finite-dimensional real Lie algebra $\mathfrak{h}$ is the Lie algebra of a connected simply connected Lie group. For this we cheat a bit like in the proof of the Peter-Weyl theorem: We assume that $\mathfrak{h}$ is a Lie algebra of matrices, i.e. it embeds in $\mathfrak{gl}_n(\mathbb{R}) = Lie(GL_n(\mathbb{R}))$ for some $n \in \mathbb{N}$. In fact a theorem by Ado says that this is always true, but we will not prove this here. By (1) we get a Lie group H with an injective immersion $H \hookrightarrow Gl_n(\mathbb{R})$ which induces on Lie algebras an isomorphism

$$Lie(H) \;\simeq\; \mathfrak{h}.$$

If $\tilde{H} \to H$ denotes the universal cover of the topological space H, then $\tilde{H}$ inherits a natural structure of a Lie group with $Lie(\tilde{H}) = \mathfrak{h}$ and hence the claim follows. □

6. Roots and the adjoint representation

Recall that by definition the adjoint representation $(\mathfrak{g}, Ad) \in \mathrm{Rep}_{\mathbb{R}}(G)$ is a representation on a real vector space. By extension of scalars we obtain a complex representation

$$Ad: \quad G \;\longrightarrow\; Aut_{\mathbb{C}}(\mathfrak{g}_{\mathbb{C}}) \quad \text{on} \quad \mathfrak{g}_{\mathbb{C}} \;=\; \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}.$$

Fixing a maximal torus $T \subset G$ with character group $X^* = X^*(T)$, we consider the weight decomposition

$$\mathfrak{g}_{\mathbb{C}}|_T \;\simeq\; \bigoplus_{\alpha \in X^*} \mathfrak{g}_\alpha.$$

Note that

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \;\subseteq\; \mathfrak{g}_{\alpha+\beta} \quad \text{for all} \quad \alpha, \beta \;\in\; X^*(T).$$

Since the adjoint representation is the complexification of a real representation, it is clear that its non-trivial weights come in pairs of complex conjugate ones. This already implies

LEMMA 6.1. *The only connected compact Lie groups of rank one are $U(1)$, the orthogonal group $SO(3)$ and its double cover $SU(2) \twoheadrightarrow SO(3)$.*

Proof. Let G be connected compact of rank one. We may assume $G \neq U(1)$, so $\dim(G) > 1$. Then in fact

$$n \;=\; \dim(G) \;\geq\; 3$$

by the above remark on non-trivial weights in the adjoint representation. Choose an $Ad(G)$-invariant inner product $\langle \cdot, \cdot \rangle : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ on $\mathfrak{g} = Lie(G)$, and let $T \subset G$ be a maximal torus. Fixing

$$X \;\in\; \mathfrak{t} \;=\; Lie(T) \quad \text{with} \quad \langle X, X \rangle \;=\; 1,$$

we get a smooth map

$$f: \quad G/T \;\longrightarrow\; S^{n-1} \;=\; \big\{ Y \in \mathfrak{g} \mid \langle Y, Y \rangle = 1 \big\}, \quad \overline{g} \;\mapsto\; Ad_g(X).$$

We claim that this map is finite-to-one (in fact injective). Indeed $f(\overline{g}_1) = f(\overline{g}_2)$ implies $Ad_g(X) = X$ for $g = g_1^{-1} g_2$. Then $Ad_g : \mathfrak{t} \to \mathfrak{t}$ is the identity map and therefore $g \in Z_G(T)$, but we know that $Z_G(T)/T$ is finite (in fact trivial).

Since $\dim(G/T) = n - 1 = \dim(S^{n-1})$, it follows that f is surjective, the image of f being closed by compactness. Hence there exists $g \in G$ with $Ad_g(X) = -X$ and then

$$Ad_g = -id \ \in \ Aut(T) = \mathbb{Z}/2\mathbb{Z} = Aut(\pi_1(T)).$$

Now the inclusion $i : T \hookrightarrow G$ gives a diagram

$$\begin{array}{ccc} \pi_1(T) & \xrightarrow{Ad_g = -id} & \pi_1(T) \\ {\scriptstyle i_*}\downarrow & & \downarrow{\scriptstyle i_*} \\ \pi_1(G) & \xrightarrow{Ad_g = id} & \pi_1(G) \end{array}$$

where the lower row is the identity map because $Ad_g : G \to G$ is homotopic to the identity by connectedness of G (note that this fails for $Ad_g : T \to T$ because $N_G(T)$ is not connected). So

$$i_* : \quad \pi_1(T) \longrightarrow \pi_1(G)$$

is not injective. But from algebraic topology the fiber bundle $T \hookrightarrow G \twoheadrightarrow S^{n-1}$ gives rise to an exact sequence

$$\cdots \longrightarrow \pi_2(S^{n-1}) \longrightarrow \pi_1(T) \longrightarrow \pi_1(G) \longrightarrow \pi_1(S^{n-1})$$

of homotopy groups, and the group on the left vanishes for $n > 3$. For a more direct argument, if i_* is not injective, the fiber bundle

$$T \ \hookrightarrow \ G \ \twoheadrightarrow \ S^{n-1}$$

is not trivial. But it is well-known from topology that non-trivial fiber bundles with fiber $T = S^1$ over S^{n-1} exist only for $n \leq 3$. Indeed, write $S^{n-1} = D_+ \cup D_-$ as the union of two hemispheres; since these are contractible, any fiber bundle over them is trivial, so we may pick trivializations $\varphi_\pm : G|_{D_\pm} \simeq T \times D_\pm$. Now the fiber bundle $G \twoheadrightarrow S^{n-1}$ is obtained by glueing the two trivial bundles via the clutching function

$$\varphi_+ \circ \varphi_-^{-1} : \quad S^{n-2} \ = \ D_+ \cap D_- \ \longrightarrow \ T \ = \ S^1$$

and the fiber bundle is trivial iff this clutching function is homotopic to a constant map. For $n > 3$ this is always the case, since then S^{n-2} is simply connected so that we get a lift

$$\begin{array}{ccc} & & \mathbb{R} \\ & \nearrow & \downarrow{\scriptstyle exp} \\ S^{n-2} & \longrightarrow & S^1 \end{array}$$

and can use that the universal cover of the circle is contractible. Hence we must have $n = 3$, and incidentally we have rediscovered the Hopf fibration.

Via the invariant inner product from above, the adjoint representation then gives a homomorphism $\rho : G \to SO(3)$. Now for any connected Lie group the kernel of the adjoint representation is precisely the center $Z(G)$. In our case $Z(G)$ is finite since it lies inside but differs from the torus T of rank one. Accordingly $\ker(\rho)$ is finite and $\rho : G \to SO(3)$ is a topological covering map. But we have seen in the exercises that

$$\pi_1(SO(3)) \ \simeq \ \mathbb{Z}/2\mathbb{Z}$$

and that $SU(2) \to SO(3)$ is the universal cover, hence the claim follows. $\square$

Returning to a general compact connected Lie group with maximal torus $T \subset G$, let us now study the adjoint representation in more detail.

Definition 6.2. The non-zero weights of the adjoint representation are called the *roots* of the compact Lie group G. In what follows we denote the set of roots by

$$\Phi \;=\; \Phi(G,T) \;=\; \{\, 0 \neq \alpha \in X^* \mid \mathfrak{g}_\alpha \neq 0 \,\}.$$

Via the diagram

$$\begin{array}{ccc} \mathfrak{t} & \xrightarrow{\;exp\;} & T \\ {\scriptstyle Lie(\alpha)}\downarrow & & \downarrow{\scriptstyle \alpha} \\ \mathbb{R} & \xrightarrow{\;exp\;} & U(1) \end{array}$$

we embed the characters as a lattice $X^* \subset \mathfrak{t}^*$ in the dual vector space of the Lie algebra $\mathfrak{t} = Lie(T)$ and view the roots as a subset of this real vector space.

Example 6.3. For the unitary group $G = U(n)$ the complexified Lie algebra takes the form

$$\mathfrak{g}_\mathbb{C} \;=\; \left\{ M \in \mathrm{Mat}_{n\times n}(\mathbb{C}) \mid \overline{M}^t = -M \right\} \otimes_\mathbb{R} \mathbb{C} \;\simeq\; \mathfrak{gl}_n(\mathbb{C})$$

via the assignment

$$M \otimes z \;\mapsto\; z \cdot M \quad \text{for} \quad M \;\in\; \mathfrak{g} \;\subset\; \mathrm{Mat}_{n\times n}(\mathbb{C}), \quad z \;\in\; \mathbb{C}.$$

Hence

$$\mathfrak{g}_\mathbb{C} \;\simeq\; \mathfrak{t}_\mathbb{C} \oplus \bigoplus_{j\neq k} \mathfrak{g}_{jk}$$

where we put

$$\mathfrak{t}_\mathbb{C} \;=\; \bigoplus_{k=1}^{n} \mathbb{C} \cdot E_{kk} \quad \text{and} \quad \mathfrak{g}_{jk} \;=\; \mathbb{C} \cdot E_{jk}.$$

for the elementary matrices $E_{jk} \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ with an entry one at position (j,k) and zeroes everywhere else. Thus the adjoint action of the maximal torus of diagonal matrices $T = \{ diag(z_1, \dots, z_n) \mid z_1, \dots, z_n \in U(1)\} \subset G$ is trivial on $\mathfrak{t}_\mathbb{C}$, but on $\mathfrak{g}_{jk}$ it is given by

$$diag(z_1, \dots, z_n) \cdot E_{jk} \;=\; z_j \overline{z}_k \cdot E_{jk}.$$

So

$$\Phi \;=\; \{\pm(e_j - e_k) \mid j < k\} \;\subset\; X^* \;=\; \mathbb{Z}^n$$

where $e_j \in \mathbb{Z}^n$ are the standard basis vectors. All roots enter with multiplicity one, and this is a general pattern:

Theorem 6.4. *Let G be a compact connected Lie group and $T \subset G$ a maximal torus with Lie algebra*

$$\mathfrak{t} \;=\; Lie(T) \;\subset\; \mathfrak{g} \;=\; Lie(G).$$

Then

(1) $\mathfrak{g}_0 = \mathfrak{t}_\mathbb{C}$,

(2) *for all $\alpha \in \Phi$ one has* $\dim \mathfrak{g}_\alpha = 1$, $\mathbb{Q}\alpha \cap \Phi = \{\pm\alpha\}$, *and a Lie algebra isomorphism*

$$\mathfrak{sl}_2(\mathbb{C}) \;\simeq\; \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \;\subseteq\; \mathfrak{g}_\mathbb{C},$$

(3) *there exists an element* $s_\alpha \in W(G,T)$ *of order two acting trivially on the hyperplane*

$$Lie(\ker(\alpha)) \;\subset\; Lie(T) \;=\; \mathfrak{t}.$$

Proof. (1) Clearly $\mathfrak{t}_\mathbb{C} \subseteq \mathfrak{g}_0$. On the other hand, since T acts trivially on $\mathfrak{g}_0$, we have

$$[\mathfrak{t}_\mathbb{C}, \mathfrak{g}_0] = 0.$$

Hence if $\mathfrak{t}_\mathbb{C} \neq \mathfrak{g}_0$, we can find a strictly bigger subspace $\mathfrak{t}_\mathbb{C} \subsetneq \mathfrak{h} \subseteq \mathfrak{g}_0$ with $[\mathfrak{h}, \mathfrak{h}] = 0$, and since on commutative Lie subalgebras the exponential map is a homomorphism, the image

$$H \;=\; exp(\mathfrak{h}) \;\subseteq\; G$$

will then be a connected abelian subgroup. A priori this subgroup might not be closed, but its closure $\overline{H} \subseteq G$ is still a connected abelian subgroup and, being closed, it is automatically a Lie subgroup as we will see in the exercises using the exponential map. As such it is compact and hence a subtorus with $T \subsetneq \overline{H}$, which contradicts our assumption that $T \subset G$ is a maximal torus.

(2) For any $\beta \in \Phi$ denote by $U_\beta = \ker(\beta) \subset T$ the kernel of β. Then for $t \in T$ we have

$$(\star) \qquad Lie(Z_G(t))_\mathbb{C} \;=\; \{X \in \mathfrak{g} \mid Ad_t(X) = X\} \;=\; \mathfrak{t}_\mathbb{C} \oplus \bigoplus_{U_\beta \ni t} \mathfrak{g}_\beta$$

since $Z_G(t) = \{z \in G \mid t^{-1}zt = z\}$. Now fix $\alpha \in \Phi$ and consider the connected component

$$U_\alpha^\circ \;\subseteq\; U_\alpha,$$

which is a subtorus of dimension $r-1$ where $r = \dim(T)$. Taking $t \in U_\alpha^\circ$ to be a topological generator for this subtorus, i.e. an element generating a dense subgroup, we claim that

$$(\star\star) \qquad \dim Z_G(t) \;=\; r+2.$$

Note that in this case on the right hand side of equation $(\star)$ only $\beta = \pm\alpha$ occur and both do with multiplicity one, so the first two properties in part (2) follow and the third one will become clear in the proof. To prove claim $(\star\star)$ we first observe that

$$U_\alpha^\circ \;\subseteq\; T \;\subseteq\; Z_G(t) \quad \text{and} \quad U_\alpha^\circ \;=\; \overline{\{t^n \mid n \in \mathbb{Z}\}} \;\subseteq\; Z(Z_G(t)),$$

which together implies that

$$T/U_\alpha^\circ \;\hookrightarrow\; Z_G(t)/U_\alpha^\circ$$

is again a maximal torus (exercise). But this torus has dimension $r - (r-1) = 1$, hence the connected compact Lie group $Z_G(t)/U_\alpha^\circ$ has rank one. Now the only connected compact Lie groups of rank one are

- the circle $U(1)$,
- the orthogonal group $SO(3)$,
- its double cover $SU(2) \twoheadrightarrow SO(3)$,

as we have seen in the previous lemma. In our case the circle group $U(1)$ cannot occur since by $(\star)$ the inclusion $T \hookrightarrow Z_G(t)$ is strict. Hence it follows that there exists an isomorphism

$$\mathfrak{su}(2) \;\xrightarrow{\sim}\; Lie(Z_G(t)/U_\alpha^\circ),$$

which implies ($\star\star$). After complexification we obtain that $\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus [\mathfrak{g}_\alpha, \mathfrak{g}_\alpha] \subseteq \mathfrak{g}_\mathbb{C}$ is a Lie subalgebra which can be at most three-dimensional, and it follows that the composite

$$\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus [\mathfrak{g}_\alpha, \mathfrak{g}_\alpha] \hookrightarrow \mathit{Lie}(Z_G(t))_\mathbb{C} \twoheadrightarrow \mathit{Lie}(Z_G(t)/U_\alpha^\circ)_\mathbb{C} \simeq \mathfrak{su}(2)_\mathbb{C} \simeq \mathfrak{sl}_2(\mathbb{C})$$

is an isomorphism of Lie algebras.

(3) For the corresponding Weyl groups, a look at the above rank one Lie groups implies

$$\mathbb{Z}/2\mathbb{Z} \simeq W(Z_G(t)/U_\alpha^\circ) \simeq N_{Z_G(t)}(T)/Z_{Z_G(t)}(T) \hookrightarrow N_G(T)/Z_G(T) = W(G)$$

and we can take $s_\alpha \in W(G)$ to be the non-trivial element on the left hand side. $\square$

Definition 6.5. For any $\alpha \in \Phi \subset \mathfrak{t}^*$ the involution $s_\alpha \in W(G)$ in the above theorem acts non-trivially but fixes a hyperplane, hence there exists a unique $\alpha^\vee \in \mathfrak{t}$ such that

$$s_\alpha(\alpha^\vee) = -\alpha^\vee \quad \text{and} \quad \langle \alpha^\vee, \alpha \rangle = 2.$$

Then a short computation shows that the adjoint action of s_α on $\mathfrak{t}$ and $\mathfrak{t}^*$ is given by

$$\begin{aligned} s_\alpha(\beta) &= \beta - \langle \beta, \alpha \rangle \cdot \alpha^\vee \quad \text{for} \quad \beta \in \mathfrak{t} \\ s_\alpha(\gamma) &= \gamma - \langle \alpha^\vee, \gamma \rangle \cdot \alpha \quad \text{for} \quad \gamma \in \mathfrak{t}^* \end{aligned}$$

We call $\alpha^\vee$ the *coroot* attached to α and denote by

$$\Phi^\vee = \Phi^\vee(G, T) = \{\alpha^\vee \mid \alpha \in \Phi\} \subset \mathfrak{t}$$

the set of all such coroots. With respect to the embedding $X_* \subset \mathfrak{t} = \mathrm{Hom}_\mathbb{R}(\mathbb{R}, \mathfrak{t})$ given by the exponential map

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{exp} & U(1) \\ {\scriptstyle Lie(\beta)}\downarrow & & \downarrow{\scriptstyle \beta} \\ \mathfrak{t} & \xrightarrow{exp} & T \end{array}$$

all the coroots sit inside the cocharacter lattice:

Lemma 6.6. *We have* $\Phi^\vee \subset X_* \subset \mathfrak{t}$.

Proof. Let $\alpha \in \Phi$ be a root, and put $\xi = \frac{1}{2}\alpha^\vee \in \mathfrak{t}$. Then $\langle \alpha^\vee, \alpha \rangle = 2$ implies that $\alpha(exp(\xi)) = 1$, hence

$$exp(\xi) \in U_\alpha^\circ = \ker(\alpha)^\circ$$

Since $s_\alpha|_{U_\alpha^\circ} = id_{U_\alpha^\circ}$, it follows that $exp(s_\alpha(\xi)) = s_\alpha(exp(\xi)) = exp(\xi)$. But for commutative Lie groups the exponential map is a homomorphism, hence it follows that

$$\xi - s_\alpha(\xi) \in \ker\Big(\mathfrak{t} \xrightarrow{exp} T \Big) = X_*.$$

Inserting

$$s_\alpha(\xi) = s_\alpha\big(\tfrac{1}{2}\alpha^\vee\big) = -\tfrac{1}{2}\alpha^\vee = -\xi$$

we get $\alpha^\vee = 2\xi \in X_*$ as required. $\square$

EXAMPLE 6.7. (a) For $G = SU(2)$, the adjoint representation is the conjugation action on complex trace-free matrices via $\mathfrak{su}(2)_{\mathbb{C}} \simeq \mathfrak{sl}_2(\mathbb{C})$. Thus $\Phi = \{\pm\alpha\}$ for the root

$$\begin{aligned}\alpha &\mathrel{\widehat{=}} \Big(diag(a, a^{-1}) \mapsto a^2\Big) \;\in\; \mathrm{Hom}(T, U(1)) \\ &\mathrel{\widehat{=}} \Big(diag(X, -X) \mapsto 2X\Big) \;\in\; \mathrm{Hom}(Lie(T), \mathbb{R}),\end{aligned}$$

where $T \subset G$ denotes the maximal torus of diagonal matrices. Here $s_\alpha \in W(G,T)$ is the coset

$$s_\alpha \;=\; \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot T$$

and we have

$$\begin{aligned}\alpha^\vee &\mathrel{\widehat{=}} \Big(a \mapsto diag(a, a^{-1})\Big) \;\in\; \mathrm{Hom}(U(1), T) \\ &\mathrel{\widehat{=}} \Big(X \mapsto diag(X, -X)\Big) \;\in\; \mathrm{Hom}(\mathbb{R}, Lie(T)).\end{aligned}$$

Therefore

$$X^*(T) \;=\; \mathbb{Z} \;\supset\; \Phi(G,T) \;=\; \{\pm 2\} \quad \text{and} \quad X_*(T) \;=\; \mathbb{Z} \;\supset\; \Phi^\vee(G,T) \;=\; \{\pm 1\}.$$

(b) Now consider the quotient $G \twoheadrightarrow \overline{G} = G/\{\pm 1\} \simeq SO(3)$. Putting $\overline{T} = T/\{\pm 1\}$ but still using the previous character and cocharacter lattices as a reference point, we get the same roots and coroots but in $X^*(\overline{T}) = 2\mathbb{Z}$ and $X_*(\overline{T}) = \frac{1}{2}\mathbb{Z}$. So after rescaling

$$X^*(\overline{T}) \;=\; \mathbb{Z} \;\supset\; \Phi(\overline{G},\overline{T}) \;=\; \{\pm 1\} \quad \text{and} \quad X_*(\overline{T}) \;=\; \mathbb{Z} \;\supset\; \Phi^\vee(\overline{G},\overline{T}) \;=\; \{\pm 2\}.$$

Note that the configuration of roots and coroots is different from the previous one, even though both groups share the same Lie algebra.

REMARK 6.8. The above structure can be summarized concisely as follows. We define a *root datum* to be a tuple $(X, \Phi, X^\vee, \Phi^\vee)$ where X and $X^\vee$ are finite rank free $\mathbb{Z}$-modules with a perfect pairing

$$\langle \cdot, \cdot \rangle : \quad X^\vee \times X \;\longrightarrow\; \mathbb{Z}$$

and $\Phi \subset X$, $\Phi^\vee \subset X^\vee$ are finite subsets between which we are given a bijective map

$$(-)^\vee : \quad \Phi \;\xrightarrow{\sim}\; \Phi^\vee, \quad \alpha \;\mapsto\; \alpha^\vee$$

such that the following axioms hold:

- for all $\alpha \in \Phi$ one has $\langle \alpha^\vee, \alpha \rangle = 2$,
- the reflection $s_\alpha : X \to X, \beta \mapsto \beta - \langle \alpha^\vee, \beta \rangle \cdot \alpha$ sends Φ to itself,
- its dual with respect to $\langle \cdot, \cdot \rangle$ sends $\Phi^\vee$ to itself.

The root datum is called *reduced* if furthermore $\mathbb{Q}\alpha \cap \Phi = \{\pm\alpha\}$ for all $\alpha \in \Phi$. In the above we have assigned to each compact connected Lie group a reduced root datum via

$$G \;\mapsto\; \big(X^*(T), \Phi(G,T), X_*(T), \Phi^\vee(G,T)\big)$$

where T is a maximal torus; the last two axioms follow from the fact that the Weyl group action permutes the roots and hence also the coroots. With the obvious notion of isomorphism of root data, the isomorphism class of the above root datum does not depend on the chosen maximal torus T and one can show that this gives a bijection

$$\left\{\begin{array}{c}\text{isomorphism classes of compact}\\ \text{connected Lie groups}\end{array}\right\} \xrightarrow{\sim} \left\{\begin{array}{c}\text{isomorphism classes of}\\ \text{reduced root data}\end{array}\right\}$$

This classifies compact connected Lie groups by discrete combinatorial data. We also note that on the isomorphism classes of reduced root data one has a natural involution

$$(X, \Phi, X^\vee, \Phi^\vee) \;\mapsto\; (X^\vee, \Phi^\vee, X, \Phi)$$

which interchanges the role of roots and coroots. The corresponding involution on compact connected Lie groups is known as *Langlands duality*.

For the study of compact connected Lie groups *modulo their centre* one can use the weaker notion of a *root system*. Here X and $X^\vee$ are replaced by a real vector space V and its dual; the axioms for a root system are the same as for a root datum plus the extra condition that $V = \langle \Phi \rangle_\mathbb{R}$ and $\langle \alpha^\vee, \beta \rangle \in \mathbb{Z}$ for all $\alpha, \beta \in \Phi$.

In talking about root systems one often identifies V with its dual V^* by choosing an inner product

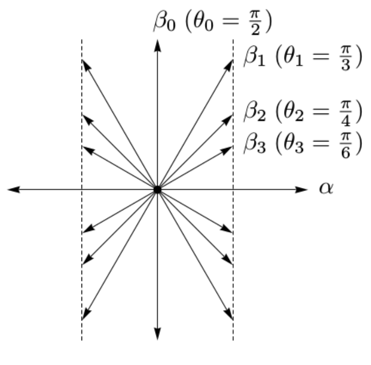
$$(\cdot, \cdot): \quad V \times V \;\longrightarrow\; \mathbb{R}$$

which is invariant under the Weyl group. Then s_α is the reflection in the hyperplane orthogonal to α, so

$$n_{\alpha\beta} \;:=\; 2\frac{(\alpha,\beta)}{(\alpha,\alpha)} \;=\; \langle \alpha^\vee, \beta \rangle \;\in\; \mathbb{Z} \quad \text{for all} \quad \alpha, \beta \;\in\; \Phi\,.$$

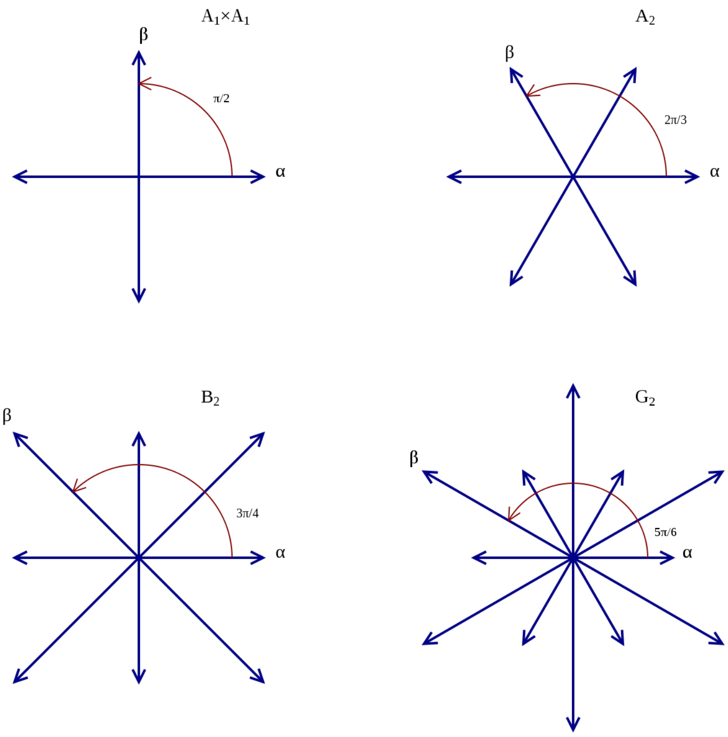
This integrality condition imposes strong restrictions on the possible configurations of roots, which only leaves finitely many cases in each dimension. If $\theta \in (0, \pi)$ denotes the angle between two roots $\alpha \neq \pm\beta \in \Phi$, then $4\cos^2(\theta) = n_{\alpha\beta} \cdot n_{\beta\alpha} \in \mathbb{Z}$ implies $4\cos^2(\theta) \in \{0, 1, 2, 3\}$:

$4\cos^2\theta$	$n_{\beta\alpha}$	$n_{\alpha\beta}$	$\lVert\alpha\rVert \,/\, \lVert\beta\rVert$
3	$+1$	$+3$	$\sqrt{3}$
	-1	-3	$\sqrt{3}$
2	$+1$	$+2$	$\sqrt{2}$
	-1	-2	$\sqrt{2}$
1	$+1$	$+1$	1
	-1	-1	1
0	0	0	any



Since the coroots $\alpha^\vee$ are determined uniquely by $\langle \alpha^\vee, - \rangle = 2(\alpha, -)/(\alpha, \alpha)$, the root system is given by the configuration of roots in the ambient Euclidean vector space.

Example 6.9 (Reduced root systems of rank two). The following pictures from `en.wikipedia.org/wiki/Root_system` show all two-dimensional reduced root systems:



The first three come from the classical groups $G = SU(2) \times SU(2)$, $SU(3)$, $SO(5)$ and the last one comes from $G = G_2$, one of the so-called *exceptional groups* that only exist in small dimensions and do not fit in the classical series of unitary, orthogonal or symplectic Lie groups.

Example 6.10 (Root systems of the classical groups). In the exercises we have obtained the following list of roots and Weyl groups, where $e_1, \dots, e_n \in \mathbb{R}^n$ are the standard basis vectors on which the Weyl groups act by signed permutations and we put $(\pm 1)_0^n = \ker(\det) \subset (\pm 1)^n$ for the even orthogonal groups:

G	$\Phi(G,T)$	$W(G,T)$
$SU(n)$	$\{\pm(e_i - e_j) \mid 1 \leq i < j \leq n\}$	$\mathfrak{S}_n$
$SO(2n)$	$\{\pm(e_i \pm e_j) \mid 1 \leq i < j \leq n\}$	$(\pm 1)_0^n \trianglelefteq \mathfrak{S}_n$
$SO(2n+1)$	$\{\pm e_i \mid 1 \leq i \leq n\} \cup \Phi(SO(2n))$	$(\pm 1)^n \trianglelefteq \mathfrak{S}_n$
$Sp(n)$	$\{\pm(e_i \pm e_j) \neq 0 \mid 1 \leq i \leq j \leq n\}$	$(\pm 1)^n \trianglelefteq \mathfrak{S}_n$

In each case the standard scalar product on $\mathbb{R}^n$ is preserved under the Weyl group action, so $V = \langle \Phi \rangle_{\mathbb{R}} \subseteq \mathbb{R}^n$ may be endowed with the induced scalar product. This in particular identifies the first three root systems from the previous example.

We will not enter further in the rich geometry of root systems $\Phi \subset V$ and only state the most basic facts without proof:

(1) The subgroup $W(\Phi) := \langle s_\alpha \mid \alpha \in \Phi \rangle \subset Gl(V)$ generated by the reflections in the root hyperplanes is a finite group. It is called the *Weyl group* of the root system. For the root system of a compact connected group G with maximal torus $T \subset G$ we recover the Weyl group defined in the previous section, i.e.

$$W(\Phi(G,T)) \;=\; W(G,T).$$

(2) A subset $\Phi^+ \subset \Phi$ is called a set of *positive roots* if
- $\Phi = \Phi^+ \sqcup (-\Phi^+)$,
- if $\alpha, \beta \in \Phi^+$ and $\alpha + \beta$ is a root, then $\alpha + \beta \in \Phi^+$.

Such sets always exist: For any vector $v \in V$ with $(v, \alpha) \neq 0$ for all $\alpha \in \Phi$ one can take
$$\Phi^+ \;=\; \big\{\, \alpha \in \Phi \mid (v, \alpha) > 0 \,\big\}.$$
In fact this sets up a bijection between sets of positive roots and connected components of
$$V \setminus \bigcup_{\alpha \in \Phi} H_\alpha \quad \text{where} \quad H_\alpha \;:=\; (\mathbb{R} \cdot \alpha)^\perp \;=\; \ker(s_\alpha + id).$$
These connected components are called *Weyl chambers*; the group $W(\Phi)$ permutes them simply transitively. Note that by the above any set of positive roots lies in some open half-space $\{w \in V \mid (v, w) > 0\}$.

(3) Fixing a set $\Phi^+ \subset \Phi$ of positive roots, a root $\alpha \in \Phi^+$ is called *simple* if it cannot be written as the sum of two positive roots. The set $\Delta \subset \Phi^+$ of all simple positive roots is a basis of V with the additional property that every $\beta \in \Phi^+$ has the form
$$\beta \;=\; \sum_{\alpha \in \Delta} n_\alpha \cdot \alpha \quad \text{with} \quad n_\alpha \;\in\; \mathbb{N}_0.$$

All this can be shown entirely in the combinatorial framework of root systems, with no Lie groups involved (except for the comparison between Weyl groups).

Corollary 6.11. *For any $\alpha \in \Delta$, we have $s_\alpha(\Phi^+ \setminus \{\alpha\}) = \Phi^+ \setminus \{\alpha\}$.*

Proof. Any $\beta = \sum_{\gamma \in \Delta} n_\gamma \cdot \gamma \in \Phi^+ \setminus \{\alpha\}$ must satisfy $n_\gamma > 0$ for some $\gamma \neq \alpha$, because no non-trivial multiple of a root is again a root. But $s_\alpha(\beta) = \beta - n_{\alpha\beta} \cdot \alpha$, so the expansions of the elements $s_\alpha(\beta)$ and β differ only in the coefficient of the simple root α. Hence in $s_\alpha(\beta)$ some simple root $\gamma \neq \alpha$ enters with multiplicity > 0, and then all of them enter with non-negative multiplicity by (3). □

7. The Weyl character formula

Recall that for any compact connected Lie group G with maximal torus $T \subset G$ we have an embedding
$$R(G) \;\hookrightarrow\; \mathbb{Z}[X]^W \quad \text{where} \quad X \;=\; X^*(T) \quad \text{and} \quad W \;=\; W(G,T).$$
and we want to show this is an isomorphism. The representation ring on the left hand side can be viewed as a subring of the ring of continuous functions $\mathscr{C}(G, \mathbb{C})$ by attaching to a representation its character, and then the above embedding comes from the restriction map
$$\mathscr{C}(G, \mathbb{C}) \;\longrightarrow\; \mathscr{C}(T, \mathbb{C}).$$
Both sides come with a Hermitian inner product defined by the the integral over the Haar measure on the respective group, and we want to relate these two. For this we define
$$\Delta(t) \;=\; \prod_{\alpha \in \Phi^+} (1 - t^{-\alpha}) \;\in\; \mathscr{C}(T, \mathbb{C}),$$
where we use the exponential notation $t^{\pm\alpha} = (\alpha(t))^{\pm 1}$ since we want to view the group $X = \mathrm{Hom}(T, U(1))$ as an additive rather than a multiplicative group. On the subring of conjugation invariant functions we have

Theorem 7.1 (Weyl integration formula). *On class functions, the multiplication map*

$$\psi : \quad \mathscr{C}(G,\mathbb{C})^G \longrightarrow \mathscr{C}(T,\mathbb{C}), \quad f \mapsto \tfrac{1}{\sqrt{|W|}}\Delta \cdot f|_T$$

is an isometry in the sense that for all conjugation-invariant $f, h \in \mathscr{C}(G,T)^G$ we have

$$\langle f,h\rangle_G \;=\; \langle \psi(f),\psi(h)\rangle_T.$$

Proof. Since $\Phi = \Phi^+ \sqcup (-\Phi^+)$ and complex conjugation acts by inversion on the roots, the claim amounts to the formula

$$\int_G f(g)dg \;=\; \frac{1}{|W|}\cdot \int_T f(t)\prod_{\alpha\in\Phi}(1-t^\alpha)dt$$

for all functions $f : G \to \mathbb{C}$ that are invariant under conjugation. Here dg and dt denote the volume forms on the compact Lie groups G and T that give rise to the normalized Haar measure on these groups. For the proof of the above formula we put $\overline{G} = G/T$ and consider the map

$$\pi : \quad T \times \overline{G} \longrightarrow G, \quad (t,\overline{g}) \mapsto gtg^{-1}$$

It follows from Cartan's theorem that this map is a generically finite cover of degree $\deg(\pi) = |W|$, so

$$\int_G f\,dg \;=\; \frac{1}{|W|}\int_{T\times\overline{G}} \pi^*(f)\,\pi^*(dg).$$

Since f is a class function,

$$(\pi^*(f))(t,\overline{g}) \;=\; f(gtg^{-1}) \;=\; f(t) \quad \text{for all} \quad (t,\overline{g}) \in T\times\overline{G}$$

so we need to show

$$\pi^*(dg) \;=\; \prod_{\alpha\in\Phi}(1-t^\alpha)\cdot dt\wedge d\overline{g}$$

where $d\overline{g}$ is the normalized left invariant measure on $\overline{G}$ of total mass one. We check this at any given point $(t_0,\overline{g}_0)$ as follows: First of all, since both sides are invariant under left translations by the group $\{1\}\times G$, we can assume $g_0 = 1$. Now look at the diagram

$$\begin{array}{ccccccc}
T_{t_0}(T)\times T_{\overline{1}}(\overline{G}) & = & T_{(t_0,\overline{1})}(T\times\overline{G}) & \xrightarrow{D\pi} & T_{t_0}(G) & \xrightarrow{\simeq} & T_{t_0}(T)\times T_{\overline{1}}(\overline{G}) \\
\downarrow \simeq & & & & \downarrow \simeq & & \downarrow \simeq \\
\mathfrak{t}\oplus\mathfrak{g}/\mathfrak{t} & & & & \mathfrak{g} & = & \mathfrak{t}\oplus\mathfrak{g}/\mathfrak{t}
\end{array}$$

where the vertical arrows are given by left translation with t_0. Note that the identification on the bottom right is canonical since it is the splitting of the adjoint representation into its invariants and the remaining isotypic pieces. One checks that via this identification

$$dg_{(t_0,\overline{1})} \;=\; \big(dt\wedge d\overline{g}\big)_{(t_0,\overline{1})},$$

indeed dt extends uniquely to a left-translation-invariant differential form $d\tau$ on G and then

$$dg \;=\; d\tau\wedge q^*(d\overline{g}) \quad \text{for the quotient map} \quad q: \quad G \twoheadrightarrow \overline{G}$$

by uniqueness of the Haar measure. So it only remains to compute the determinant of the linear map

$$(D\pi)(t_0,\overline{1}) : \quad \mathfrak{t}\oplus\mathfrak{g}/\mathfrak{t} \longrightarrow \mathfrak{t}\oplus\mathfrak{g}/\mathfrak{t}$$

Now

$$\begin{aligned}\pi\Big(t_0(1+\epsilon X), 1+\epsilon Y\Big) &= \big(1+\epsilon Y\big)\big(t_0 + t_0\epsilon X\big)\big(1+\epsilon Y\big)^{-1} \\ &= \big(1+\epsilon Y\big)\big(t_0 + t_0\epsilon X\big)\big(1 - t_0\epsilon Y + O(\epsilon^2)\big)^{-1} \\ &= t_0 \cdot \Big(1 + \epsilon\big(X - Y + t_0^{-1} Y t_0\big) + O(\epsilon^2)\Big)\end{aligned}$$

so that with respect to the decomposition $\mathfrak{t} \oplus \mathfrak{g}/\mathfrak{t}$ the map in question is given by a block diagonal matrix

$$(D\pi)(t_0, \bar{1}) \;=\; \begin{pmatrix} \mathbf{1} & 0 \\ 0 & Ad_{t_0^{-1}} - \mathbf{1} \end{pmatrix}.$$

Hence

$$\det\big((D\pi)(t_0, \bar{1})\big) \;=\; \prod_{\alpha \in \Phi} (1 - t_0^{\alpha})$$

as the roots are precisely the weights of the adjoint representaiton on $(\mathfrak{g}/\mathfrak{t})_{\mathbb{C}}$. □

For an irreducible representation $V \in \mathrm{Irr}_{\mathbb{C}}(G)$ with character $ch(V) = \chi_V|_T$ and for Δ as above put

$$\Delta \cdot ch(V) \;=\; \sum_{\mu \in X} c_\mu t^\mu \quad \text{with} \quad c_\mu \;\in\; \mathbb{Z}.$$

The coefficients c_μ may be negative, but the Weyl integration formula already gives an upper bound on their total number:

Corollary 7.2. *With notations as above* $\sum_{\mu \in X} c_\mu^2 \;=\; |W|$.

Proof. Since $\langle \chi_V, \chi_V \rangle_G = 1$, we know that $\langle \Delta\chi_V, \Delta\chi_V \rangle_T = |W|$ by the Weyl integration formula. Therefore the claim follows from the orthonormality of the characters $\mu : T \longrightarrow U(1), t \mapsto t^\mu$ of the torus. □

In order to compute the coefficients c_μ, we will exploit the fact that $ch(V)$ is invariant under the Weyl group action, so we need to understand the transformation behaviour of Δ under the Weyl group. For this it will be convenient to consider the half-sum

$$\delta \;:=\; \tfrac{1}{2} \cdot \sum_{\alpha \in \Phi^+} \alpha$$

of all positive roots:

Lemma 7.3. *For* $s \in W(G)$ *one has* $(s \cdot \Delta)(t) \;=\; \det(s) \cdot t^{\delta - s\cdot\delta} \cdot \Delta(t)$.

Proof. Let us introduce the shorthand notation $\Delta_P(t) = \prod_{\alpha \in P}(1 - t^{-\alpha})$ for any subset $P \subseteq \Phi$. Clearly $s \cdot \Delta_P \;=\; \Delta_{s\cdot P}$. Now for the set of all positive roots $P = \Phi^+$ the identity

$$s \cdot P \;=\; (P \cap s \cdot P) \sqcup -(P \setminus s \cdot P)$$

implies

$$\Delta_{s\cdot P}(t) \;=\; \Delta_P(t) \cdot \prod_{\alpha \in P \setminus s\cdot P} \frac{1 - t^\alpha}{1 - t^{-\alpha}} \;=\; \Delta_P(t) \cdot (-1)^{|P \setminus s\cdot P|} \cdot t^N$$

where

$$N \;=\; \sum_{\alpha \in P \setminus s\cdot P} \alpha \;=\; \tfrac{1}{2} \sum_{\alpha \in P \cap sP} (\alpha - \alpha) \;+\; \tfrac{1}{2} \sum_{\alpha \in P \setminus sP} (\alpha - (-\alpha)) \;=\; \delta - s \cdot \delta.$$

Furthermore, writing s as a product of reflections at simple roots α and recalling that every such simple reflection s_α permutes the roots in $P \setminus \{\alpha\}$, one inductively checks $(-1)^{|P \setminus s\cdot P|} = \det(s)$ and hence the claim follows. □

COROLLARY 7.4. *We have* $c_{s\cdot(\mu+\delta)-\delta} = \det(s)\cdot c_\mu$ *for all* $\mu \in X$ *and* $s \in W$.

Proof. Since $ch(V)$ is invariant under the Weyl group, the function $f = \Delta\cdot ch(V)$ transforms in the same way as Δ in the previous lemma. Thus

$$\sum_\nu c_\nu t^{s\nu} \;=\; s\cdot f(t) \;=\; \det(s)\cdot t^{\delta-s\delta}\cdot f(t) \;=\; \sum_\mu \det(s)c_\mu t^{\mu+\delta-s\delta}$$

and the claim follows by a comparison of coefficients, putting $s\nu = \mu+\delta-s\delta$ or equivalently $\nu = s(\mu+\delta)-\delta$. □

We are now prepared for the classification of all irreducible representations. We begin with the following

DEFINITION 7.5. A weight $\lambda \in X$ is called a *highest weight* of $V \in \mathrm{Irr}_{\mathbb{C}}(G)$ if we have

- $V_\lambda \neq 0$ but
- $V_\mu = 0$ for all $\mu = \lambda + \sum(Q)$ with $\varnothing \neq Q \subset \Phi^+$.

Here we use the notation $\sum(Q) := \sum_{\alpha\in Q}\alpha$. Clearly, any $V \in \mathrm{Irr}_{\mathbb{C}}(G)$ has a highest weight in the above sense. It turns out that the highest weight is unique and the irreducible representations are classified by their highest weights. We prove this alongside with an explicit formula for the character of a representation, where we put

$$\det(t^{\mu+\delta}) \;:=\; \sum_{w\in W}\det(s)\cdot t^{s\mu+s\delta-\delta} \quad \text{for} \quad \mu \;\in\; X$$

to underline the relation with our previous character formula for $Gl_n(\mathbb{C})$:

THEOREM 7.6 (Weyl character formula). *Any* $V \in \mathrm{Irr}_{\mathbb{C}}(G)$ *has a unique highest weight* λ. *This highest weight determines the representation up to isomorphism, more precisely*

$$ch(V) \;=\; \frac{\det(t^{\lambda+\delta})}{\det(t^\delta)}.$$

Proof. Put $P = \Phi^+$ for brevity. Let $V \in \mathrm{Irr}_{\mathbb{C}}(G)$, and write $ch(V) = \sum_\mu m_\mu\cdot t^\mu$ with $m_\mu \in \mathbb{N}_0$. By definition

$$\Delta(t) \;=\; \sum_{Q\subseteq P}(-1)^{|Q|}\cdot t^{-\sum(Q)}$$

so

$$\varphi(t) \;:=\; \Delta(t)\cdot ch(V)(t) \;=\; \sum_{\mu,Q}(-1)^{|Q|}\cdot m_\mu\cdot t^{\mu-\sum(Q)}$$

and by comparison of coefficients

$$c_\alpha \;=\; \sum_{\mu,Q}(-1)^{|Q|}\cdot m_{\alpha+\sum(Q)}.$$

Now the fact that $P = \Phi^+$ is a proper cone, i.e. contained in a half-space lying on one side of a hyperplane, implies that we can have $\sum(Q) = 0$ only if $Q = \varnothing$. Hence it follows that

$$c_\lambda \;=\; m_\lambda \quad \text{for any highest weight } \lambda.$$

We next claim that the stabilizer of λ with respect to the shifted Weyl group action defined by

$$s \bullet \mu \;:=\; s(\mu+\delta)-\delta \quad \text{for} \quad \mu \;\in\; X,\; s \;\in\; W(G)$$

is trivial: Indeed

$$\begin{aligned} s \bullet \lambda \;=\; \lambda &\iff s\lambda = \lambda + (\delta - s\delta) \\ &\iff s\lambda = \lambda + \sum (P \setminus sP) \end{aligned}$$

Since $m_{s\lambda} = m_\lambda > 0$ and λ is a highest weight, this can happen only if $P = sP$ and then $s = 1$ because $W(G)$ acts faithfully on systems of positive roots.

So it follows that the Weyl group orbit $O(\lambda) = \{w \bullet \lambda \mid w \in W\}$ has size $|W|$, but then

$$|W| \cdot c_\lambda^2 \;=\; \sum_{\mu \in O(\lambda)} c_\mu^2 \;\leq\; \sum_{\mu \in X} c_\mu^2 \;=\; |W|$$

implies that

$$c_\mu \;=\; \begin{cases} 0 & \text{if } \mu \notin O(\lambda), \\ \det(s) & \text{if } \mu = s \bullet \lambda. \end{cases}$$

Therefore

$$\Delta(t) \cdot ch(V)(t) \;=\; \det(t^{\lambda+\delta}).$$

Applying the above formula in the special case of the trivial representation $V = \mathbf{1}$ we get $\Delta(t) = \det(t^\delta)$, hence the claim follows. □

To complete the classification of irreducible representations, it only remains to clarify what weights can occur as highest weights. For this we introduce the following

Definition 7.7. A weight $\lambda \in X$ is *dominant* if $\langle \alpha^\vee, \lambda \rangle \geq 0$ for all $\alpha \in \Phi^+$.

For example, for $G = U(n)$ we have $X = \mathbb{Z}^n \supset \Phi^+ = \{e_i - e_j \mid e < j\}$ and in this case

$$\lambda \;=\; (\lambda_1, \dots, \lambda_n) \in \mathbb{Z}^n \quad \text{is dominant iff} \quad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n.$$

Note that any such dominant weight occurs as the highest weight of the irreducible representation

$$V_\lambda \;:=\; \det(\mathbb{C}^n)^{\otimes \lambda_n} \otimes S_\alpha(\mathbb{C}^n) \quad \text{where} \quad \alpha \;=\; (\lambda_1 - \lambda_n, \dots, \lambda_{n-1} - \lambda_n, 0).$$

This extends to the general case:

Theorem 7.8. *For any compact connected Lie group G, the highest weights of the irreducible representations are precisely the dominant weights.*

Idea of the proof. To see that the highest weight of $V \in \mathrm{Irr}_\mathbb{C}(G)$ is dominant, one uses for $\alpha \in \Phi$ the representation theory of

$$\mathfrak{sl}_2(\mathbb{C}) \;=\; \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus [\mathfrak{g}_\alpha, \mathfrak{g}_\alpha] \;\subset\; \mathfrak{g}_\mathbb{C}\}$$

to see that for any $\mu \in X$,

$$V_\mu \neq \{0\} \quad \Longrightarrow \quad V_{\mu - n\alpha} \neq \{0\} \quad \text{for all} \quad n \;\in\; \{0, 1, \dots, \langle \alpha^\vee, \mu \rangle\}.$$

If the highest weight λ of V would satisfy $k = \langle \alpha^\vee, \lambda \rangle < 0$, then taking $\mu = s_\alpha(\lambda)$ and $n = 1 - k > 0$ we get a contradiction. For the converse that every dominant weight occurs as a highest weight of some irreducible representation, one computes that the

$$\frac{\det(t^{\lambda+\delta})}{\det(t^\delta)} \;\in\; \mathbb{Z}[X]^W \;\subset\; \mathscr{C}(T, \mathbb{C})^W \;=\; \mathscr{C}(G, \mathbb{C})^G \quad \text{with } \lambda \text{ dominant}$$

form an orthonormal system of class functions on G. By what we have seen above, this system contains the characters of all irreducible representations. If it contained any other member, this member would be orthogonal to all class functions, which is impossible by the Peter-Weyl theorem. □

REMARK 7.9. For the representation ring a closer look at the character formula shows that we get an isomorphism

$$R(G) \xrightarrow{\sim} \mathbb{Z}[X]^W \quad \text{via} \quad [V] \mapsto ch(V).$$

APPENDIX A

Topological proof of Cartan's theorem

In this appendix we discuss a very short and elegant proof of Cartan's theorem given by A. Weil. The basic input for this is the Lefschetz fixed point theorem, for which we first recall some basic intersection theory on smooth manifolds.

1. Some intersection theory

Let M be a smooth manifold. In intersection theory one wants to understand the intersections of a closed submanifold $Y \subset M$ with another submanifold, or more generally the preimages

$$X \cap Y \;:=\; f^{-1}(Y)$$

where $f : X \to M$ is a smooth map from a compact manifold X. This preimage behaves in the expected way if the subvarieties are in general position:

Definition 1.1. We say that f is *transversal* to Y at a point $p \in X$ if the image of the differential

$$Df(p): \quad T_pX \longrightarrow T_{f(p)}M$$

satisfies

$$(Df(p))(T_pX) \;+\; T_{f(p)}Y \;=\; T_{f(p)}M.$$

Note that the sum on the left hand side is not required to be direct. We say that f is *transversal* to Y if it is so at every point, and we then write $f \pitchfork Y$. If f is a closed immersion, we also write $X \pitchfork Y$ and say that X and Y are *transversal*.

Remark 1.2. If $f \pitchfork Y$, then the preimage $f^{-1}(Y) \subseteq X$ is a smooth submanifold of dimension

$$\dim f^{-1}(Y) \;=\; \dim X + \dim Y - \dim M$$

and the map $f^{-1}(Y) \longrightarrow Y$ is a submersion. Note that the notion of transversality depends on the ambient manifold M. For instance, two submanifolds $X, Y \subset M$ with $\dim X + \dim Y < \dim M$ cannot be transversal unless they are disjoint.

In order to deal with non-transverse intersections, we look for small deformations that become transverse:

Definition 1.3. Two maps $f_0, f_1 : X \longrightarrow M$ are called *homotopic* if there exists a smooth map

$$F: \quad X \times [0,1] \longrightarrow M \quad \text{with} \quad F|_{X\times\{\nu\}} \;=\; f_\nu \quad \text{for } \nu = 0, 1.$$

Any such map is called a *homotopy* and we then say that f_1 is a *deformation* of f_0.

Fact 1.4. *In the above setting,*

1. *if f is transversal to $Y \subset M$, then so is any small deformation of it,*
2. *conversely any smooth f is homotopic to one transversal to $Y \subset M$,*
3. *if $f_0, f_1 \pitchfork Y$ are homotopic, then there exists a homotopy F between them which is transversal to Y, i.e. $F|_{X\times(0,1)} \pitchfork Y$.*

Now suppose $\dim X + \dim Y = \dim M$. In the transverse case $f^{-1}(Y)$ will then consist of finitely many points. In general one could try to define the intersection number $X \cdot_f Y$ as

$$\# f_\epsilon^{-1}(Y)$$

where $f_\epsilon \pitchfork Y$ is a small transverse deformation of the map $f : X \to M$. However, the following picture illustrates that this is ill-defined:

[PICTURE]

This problem disappears if we only consider intersection numbers modulo 2. For example, if $M = [0,1] \times (-1,1)/((0,a) \sim (1,-a))$ is the open Möbius strip, the circle

$$f : \quad X = \mathbb{R}/\mathbb{Z} \hookrightarrow M, \quad x \mapsto (x,0) \quad \text{with image} \quad Y = f(X)$$

satisfies

$$\# f_\epsilon^{-1}(Y) \equiv 1 \mod 2$$

for any small transversal deformation $f_\epsilon \pitchfork Y$ as illustrated by the following picture:

[PICTURE]

In order to define intersection numbers with values in $\mathbb{Z}$ rather than $\mathbb{Z}/2\mathbb{Z}$, we need to take orientations into account. So for the rest of this sections we will always assume that all our manifolds are oriented. We can then count the points of the intersection $f_\epsilon^{-1}(Y)$ with the appropriate signs:

Definition 1.5. If $\dim X + \dim Y = \dim M$ and f is transversal to Y at p, consider the sum map

$$(Df(p))(T_pX) \oplus T_{f(p)}Y \longrightarrow T_{f(p)}M.$$

We endow the source with the orientation induced by the given orientations of X and Y and define the *oriented local intersection number* of X and Y at the point p by

$$(X \cdot_f Y)_p = \begin{cases} +1 & \text{if the sum map is orientation-preserving,} \\ -1 & \text{if the sum map is orientation-reversing.} \end{cases}$$

For instance, in the first example the local intersection numbers look as follows:

[PICTURE]

Now the ambiguities in the choice of the deformations f_ϵ cancel out, and this is what happens in general:

Lemma 1.6. *If $f_0, f_1 \pitchfork Y$ are homotopic transversal maps, then*

$$\sum_{p \in f_1^{-1}(Y)} (X \cdot_{f_1} Y)_p = \sum_{p \in f_2^{-1}(Y)} (X \cdot_{f_2} Y)_p.$$

Proof. By the deformation property (3) we find a homotopy F between the f_ν that is transversal, i.e.

$$F|_{X\times(0,1)} \pitchfork Y \quad \text{and} \quad F|_{X\times\{\nu\}} \pitchfork Y \quad \text{for} \quad \nu \ = \ 0, 1.$$

In that case the preimage

$$F^{-1}(Y) \ \subset \ X \times [0,1]$$

is a 1-dimensional smooth manifold with boundary. Any such manifold is a disjoint union of circles and closed intervals as illustrated in the following picture

[PICTURE]

and hence that the orientations at the boundary points cancel out. □

Definition 1.7. If $\dim X + \dim Y = \dim M$, then for a smooth map $f : X \to M$ we define the *intersection number*

$$X \cdot_f Y \ := \ \sum_{p\in f_\epsilon^{-1}(Y)} (X \cdot_{f_\epsilon} Y)_p$$

where $f_\epsilon \pitchfork Y$ is any small deformation of f transversal to Y. If f is a closed immersion, we simply write $X \cdot Y$ for this intersection number.

For example, the two oriented circles $X, Y \subset M$ in the torus $M = U(1) \times U(1)$ shown below

[PICTURE]

have the intersection numbers $X \cdot Y = -Y \cdot X = 1$ and $X \cdot X = Y \cdot Y = 0$.

2. Lefschetz numbers

Now suppose that we have a smooth self-map $f : X \to X$ of an oriented compact manifold X. We define the number of fixed points, counted with multiplicities, as follows:

Definition 2.1. For $M = X \times X$, let $\Delta_X \subset X \times X$ be the diagonal, and consider the graph

$$\Gamma_f \ = \ \{(x, f(x)) \mid x \in X\} \ \subset \ M.$$

Then the integer

$$L(f) \ := \ \Delta_X \cdot \Gamma_f \ \in \ \mathbb{Z}$$

is called the *global Lefschetz number* of f. Note that homotopic maps have the same Lefschetz number. Almost by definition we have

Theorem 2.2 (Lefschetz). *If $L(f) \neq 0$, then f has a fixed point.*

Proof. Otherwise $\Delta_X \cap \Gamma_f = \varnothing$ set-theoretically, but then Δ_X and Γ_f are trivially transverse, and their intersection number is zero. □

REMARK 2.3. Using Poincaré duality and the Künneth decomposition for the cohomology of $X \times X$, one may give a global interpretation of the Lefschetz number as the sum

$$L(f) \;=\; \sum_{\nu} (-1)^{\nu}\, tr\Big(f^* \mid H^{\nu}(X,\mathbb{Q})\Big).$$

This result is sometimes called the *Lefschetz-Hopf trace formula.* For $f = id_X$ it says that

$$\Delta \cdot \Delta \;=\; \chi(X)$$

is the topological Euler characteristic, which can be seen as a reformulation of the Gauss-Bonnet theorem. The trace formula also has an arithmetic analog for the étale cohomology of varieties over finite fields, where it is used in the proof of the Weil conjectures by counting fixed points of the Frobenius endomorphisms.

EXAMPLE 2.4. The trace formula says that if $f : X \to X$ is homotopic to the identity but has no fixed point, then $\chi(X) = 0$. It follows for instance that

(1) the antipodal map $f : S^2 \to S^2, x \mapsto -x$ is not homotopic to the identity, since it has no fixed points but the Euler characteristic is $\chi(S^2) = 2$.

(2) any compact Lie group G has Euler characteristic $\chi(G) = 0$, since the left translation by any non-trivial element has no fixed point, but is homotopic to the identity if the element lies in the connected component.

How can one compute the global Lefschetz number in concrete applications? Going back to the local definition of intersection numbers we get

PROPOSITION 2.5. *With notations as above,*

(1) *We have* $\Gamma_f \pitchfork \Delta_X$ *at* (p,p) *iff* $Df(p) - id \in \mathrm{End}_{\mathbb{R}}(T_pX)$ *is invertible.*

(2) *If this condition is satisfied for all fixed points* $p \in X$*, then* $L(f)$ *can be written as a sum*

$$L(f) \;=\; \sum_{p=f(p)} L_p(f)$$

of the local Lefschetz numbers $L_p(f) := sgn \det(Df(p) - id|T_pX)$.

Proof. For the first claim, note that the tangent spaces to the graph and the diagonal are given by

$$\begin{aligned} T_{(p,p)}\Gamma_f \;&=\; \Gamma_{Df(p)} \;\subset\; T_pX \times T_pX, \\ T_{(p,p)}\Delta_X \;&=\; \Delta_{T_pX} \;\subset\; T_pX \times T_pX. \end{aligned}$$

Putting $A = Df(p)$ we get

$$\begin{aligned} \Gamma_f \pitchfork \Delta_X \text{ at } (p,p) \;&\Longleftrightarrow\; \Gamma_A + \Delta_{T_pX} \;=\; T_pX \times T_pX \\ &\Longleftrightarrow\; \Gamma_A \cap \Delta_{T_pX} \;=\; \{0\} \\ &\Longleftrightarrow\; A \,\in\, \mathrm{End}_{\mathbb{R}}(T_pX) \text{ has no non-zero fixed point} \\ &\Longleftrightarrow\; A \,\in\, \mathrm{End}_{\mathbb{R}}(T_pX) \text{ does not have 1 as an eigenvalue} \\ &\Longleftrightarrow\; A - id \text{ is invertible} \end{aligned}$$

and hence the first claim follows. For the second claim, we must see that

$$(\Gamma_f \cdot \Delta_X)_{(p,p)} \;=\; sgn \det(A - id|T_pX).$$

For this let $v_1, \dots, v_n$ be an oriented basis of the tangent space T_pX. We then get oriented bases

$$\begin{array}{rll} (v_1, v_1), \dots, (v_n, v_n) & \text{for} & T_{(p,p)}\Delta_X, \\ (v_1, Av_1), \dots, (v_n, Av_n) & \text{for} & T_{(p,p)}\Gamma_f, \\ (v_1, 0), \dots, (v_n, 0), (0, v_1), \dots, (0, v_n) & \text{for} & T_{(p,p)}X \times X. \end{array}$$

Denoting by $sgn(\mathscr{B}) \in \{\pm 1\}$ the orientation of any basis $\mathscr{B}$ of $T_{(p,p)}X \times X$ with respect to the product orientation chosen in the previous line, we get

$$\begin{aligned} (\Gamma_f \cdot \Delta_X)_{(p,p)} &= sgn((v_1, v_1), \dots, (v_n, v_n), (v_1, Av_1), \dots, (v_n, Av_n)) \\ &= sgn((v_1, v_1), \dots, (v_n, v_n), (0, (A - id)v_1), \dots, (0, (A - id)v_n)) \\ &= sgn((v_1, 0), \dots, (v_n, 0), (0, (A - id)v_1), \dots, (0, (A - id)v_n)) \\ &= sgn \det(A - id) \end{aligned}$$

where for the second equality we have subtracted the first n basis vectors from the last n ones, noting that this does not change the orientation of the basis, and for the third equality we have similarly subtracted $(A - id)^{-1}$ times the last n basis vectors from the first n basis vectors, which again does not affect orientations. □

EXAMPLE 2.6. Let $\dim X = 2$. If in a suitable basis of the tangent space T_pX the Jacobi matrix is in diagonal form $D_pf = diag(\lambda_1, \lambda_2)$ with $\lambda_1, \lambda_2 > 0$, then

$$L_p(f) = sgn((\lambda_1 - 1)(\lambda_2 - 1)) = \begin{cases} +1 & \text{if } p \text{ is a source or sink,} \\ -1 & \text{if } p \text{ is a saddle point,} \end{cases}$$

as illustrated by the following picture:

[PICTURE]

As a fun application, we may compute the Euler characteristic of all compact oriented surfaces. Any such surface has the form of a doughnut with g holes. If we cover it with a chocolate topping and let this flow down vertically as indicated in the following picture, we obtain a family of maps f homotopic to the identity but with precisely $2g + 2$ fixed points: A source at the top, a sink at the bottom, and two saddle points in each hole.

[PICTURE]

Thus for the total Lefschetz number we get the value $\chi(X) = L(f) = 2 - 2g$.

3. Proof of Cartan's theorem

Now let G be a compact connected Lie group with a maximal torus $T \subset G$, and pick an arbitrary element $g \in G$. Recall that Cartan's theorem claims that we have

$$g \in xTx^{-1} \quad \text{for some} \quad x \in G.$$

The idea of A. Weil is to interpret this as a fixed point relation. For this we consider the coset space $X = G/T$ (which may be shown to be again a smooth manifold) and the smooth map

$$\lambda_g : \quad X \longrightarrow X, \quad xT \mapsto gxT$$

given by left translation. In what follows we write $\overline{x} = xT$ for the cosets.

Lemma 3.1. *We have* $g \in xTx^{-1}$ *iff* $\overline{x} \in X$ *is a fixed point of* λ_g.

Proof. We have $\lambda_g(xT) = xT$ iff $gxT = xT$ iff $gx \in xT$ iff $g \in xTx^{-1}$. □

In these terms, Cartan's theorem translates to the statement that each of the left translations $\lambda_g : X \to X$ has a fixed point. By the Lefschetz fixed point theorem it will thus be enough to show

$$L(\lambda_g) \neq 0 \quad \text{for all} \quad g \in G.$$

Note that by connectedness of G all the left translations λ_g are homotopic, so it suffices in fact to show the above non-vanishing for a single $g \in G$. In particular, if we knew that $\chi(G/T) \neq 0$, then we could take $g = 1$ and would be finished; but we will take a different route which does not require any information about cohomology: We choose $g = t \in T$ to be a topological generator of the maximal torus and obtain

Corollary 3.2. *If* $t \in T$ *is a topological generator, then* $\overline{x} \in X$ *is fixed by* λ_t *iff* $x \in N_G(T)$. *In particular,* $\lambda_t : X \to X$ *has only finitely many fixed points.*

Proof. The first statement is clear from the above lemma, since for a topological generator $t \in xTx^{-1}$ implies by maximality that $T = xTx^{-1}$. For the finiteness statement, note that $x \in N_G(T)$ iff $\overline{x} \in W(G,T)$, and the Weyl group is finite. □

We must now compute the contribution of each of these finitely many fixed points $\overline{x} \in W(G,T)$ to the global Lefschetz number $L(\lambda_t)$. However, In fact all of them yield the same contribution and so it suffices to consider the contribution at the neutral element $e \in G$:

Lemma 3.3. *For all* $x \in N_G(T)$,

$$\det\Big(D\lambda_t(\overline{x}) - id \,\Big|\, T_{\overline{x}}X\Big) \;=\; \det\Big(D\lambda_t(\overline{e}) - id \,\Big|\, T_{\overline{e}}X\Big).$$

Proof. Since $x \in N_G(T)$, the right translation by x gives a well-defined smooth map

$$\rho_x : \quad X \longrightarrow X, \quad \overline{g} \;=\; gT \;\mapsto\; gTx \;=\; gxT \;=\; \overline{gx}$$

of the left coset space $X = G/T$. Since right and left translations commute, we get a commutative diagram

$$\begin{array}{ccc} T_{\overline{e}}X & \xrightarrow{D\rho_x(\overline{e})} & T_{\overline{x}}X \\ {\scriptstyle D\lambda_t(\overline{e})}\downarrow & & \downarrow{\scriptstyle D\lambda_t(\overline{x})} \\ T_{\overline{e}}X & \xrightarrow[D\rho_x(\overline{e})]{} & T_{\overline{x}}X \end{array}$$

which shows that $D\lambda_t(\overline{e})$ and $D\lambda_t(\overline{x})$ are conjugate matrices. □

Since by the above lemma all the local Lefschetz numbers are the same, the proof of Cartan's theorem is finished by the following

Lemma 3.4. *For any topological generator* $t \in T$ *we have* $\det(D\lambda_t(\overline{e}) - id) \neq 0$.

Proof. On the coset space $X = G/T$ the left translation λ_t acts in the same way as conjugation by t, hence on the level of tangent spaces we get a commutative diagram of short exact sequences

$$\begin{array}{ccccccccc}
0 & \longrightarrow & Lie(T) & \longrightarrow & Lie(G) & \longrightarrow & T_{\overline{e}}X & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow{\scriptstyle Ad_t} & & \downarrow{\scriptstyle D\lambda_t(\overline{e})} & & \\
0 & \longrightarrow & Lie(T) & \longrightarrow & Lie(G) & \longrightarrow & T_{\overline{e}}X & \longrightarrow & 0
\end{array}$$

where Ad_t denotes the differential of $t(-)t^{-1} : G \to G$. Now the complexification of the adjoint representation

$$\Big(Ad : \; G \;\longrightarrow\; Aut(Lie(G)) \Big) \;\in\; \mathrm{Rep}_{\mathbb{R}}(G)$$

restricts on T to

$$\big(Lie(G) \otimes_{\mathbb{R}} \mathbb{C}\big)|_T \;\simeq\; Lie(T) \otimes_{\mathbb{R}} \mathbb{C} \;\oplus \bigoplus_{\alpha \in \Phi} V_\alpha$$

for some $\Phi \subset X^*(T) \setminus \{0\}$, in particular the subspace $Lie(T) \otimes_{\mathbb{R}} \mathbb{C}$ consists precisely of the invariants since otherwise the subtorus $T \subset G$ would not be maximal. Thus the representation

$$T_e(X) \otimes_{\mathbb{R}} \mathbb{C} \;\simeq\; \bigoplus_{\alpha \in \Phi} V_\alpha \;\in\; \mathrm{Rep}_{\mathbb{C}}(T)$$

does not contain trivial subrepresentations. Since $t \in T$ is a topological generator, it follows that the action of t on this representation has all its eigenvalues $\neq 1$, which implies that

$$\det\Big(D\lambda_t(\overline{e}) - id \,\Big|\, (T_{\overline{e}}X) \otimes_{\mathbb{R}} \mathbb{C} \Big) \;\neq\; 0.$$

Since an $\mathbb{R}$-linear map is invertible iff its complexification is so, this proves the claim and hence also Cartan's theorem. □