Representations of Finite Groups

Lecture notes

October 6, 2015

Contents

Preface

These notes are based on an 8-lecture course taught at Radboud University Nijmegen by Prof. Ieke Moerdijk in the spring of 2015. The course treated the basics of the theory of representations of finite groups, in particular the correspondence between representations and characters, the representation theory of the symmetric group, and induced representations. Since the topic involves a lot of linear algebra, the notes include an appendix that summarizes some important concepts from linear algebra. Furthermore, a section on category theory is included, to put the theory of representations in a broader perspective.

A first draft of these lecture notes was written by Jan Schoone. The notes have been polished and expanded by Frank Roumen.

1 Representations

1.1 Definition and examples

Representation theory is about studying a group by looking at the ways in which it acts on vector spaces. Since vector spaces and linear maps, or matrices, are well understood, this often makes analysing the group easier. Many concepts from linear algebra have analogues in the theory of representations. For a review of some useful linear algebra, see Appendix A.

A group representation is similar to an action of the group on a set, but we replace the set by a vector space and require the action to be linear. Recall that there are two equivalent definitions of a group action: it can be defined as a map $G \times X \to X$ with certain properties, or as a group homomorphism $G \to S_X$. Here S_X is the symmetric group on X, which consists of all bijections from X to itself. Similarly there are three versions of the definition of a group representation. In the following, let G be a group and V a vector space, over either the real or the complex number field.

Version 1. A representation is an action of G on V by linear maps. This means that it is a map

$$G \times V \to V, \quad (g, v) \mapsto g \cdot v$$

with the following properties:

- $g \cdot (-) : V \to V$ is linear, which means that $g \cdot (v + w) = g \cdot v + g \cdot w$ and $g \cdot (\lambda v) = \lambda g \cdot v$ for all $v, w \in V$ and λ in $\mathbb{R}$ or $\mathbb{C}$.
- $e \cdot v = v$ and $(gh) \cdot v = g \cdot (h \cdot v)$.

Version 2. A representation is a group homomorphism $\rho : G \to \mathrm{GL}(V)$. This is the same as Version 1, because such a group homomorphism and an action determine each other via $\rho(g)(v) = g \cdot v$. The homomorphism property translates into the requirements for an action under this correspondence.

Version 3. A representation is a group homomorphism $\rho : G \to \mathrm{GL}_n(\mathbb{R})$ or $G \to \mathrm{GL}_n(\mathbb{C})$. If V is finite-dimensional with dimension n, then such a representation can be obtained from Version 2. Pick a basis $\{e_1, \dots, e_n\}$ for V, and express $\rho(g)$ in terms of this basis as an $n \times n$-matrix. This version is not exactly equivalent to the other two versions, because here we had to choose a basis. If we choose another basis $\{e'_1, \dots, e'_n\}$, then we obtain different matrices $\rho'(g)$. However, there is a connection between the matrices $\rho(g)$ and $\rho'(g)$: there exists one single invertible matrix A such that for all $g \in G$ we have $\rho(g) = A\rho'(g)A^{-1}$.

Convention. In this course we will mainly be interested in representations of *finite* groups on *finite-dimensional* vector spaces. Therefore, we will assume that all groups we encounter are finite and all vector spaces are finite-dimensional, unless otherwise stated. Also, the theory of representations is often easier over the complex numbers than over the real numbers, so we assume that all vector spaces are over $\mathbb{C}$ unless otherwise stated.

Examples 1.1.

1. Recall that the cyclic group C_n consists of the elements $e, r, r^2, \ldots, r^{n-1}$, with the relation $r^n = e$. There is a representation of this group on the vector space $\mathbb{C}^2$, given by

$$\rho : C_n \to \mathrm{GL}_2(\mathbb{C}), \quad \rho(r^k) = \begin{bmatrix} \cos\left(\frac{2\pi k}{n}\right) & -\sin\left(\frac{2\pi k}{n}\right) \\ \sin\left(\frac{2\pi k}{n}\right) & \cos\left(\frac{2\pi k}{n}\right) \end{bmatrix}$$

   Since this assignment satisfies $\rho(e) = I$ and $\rho(gh) = \rho(g)\rho(h)$, it is indeed a representation in Version 3. It is called the *standard representation* of C_n.

   It is also possible to express the same representation using Version 1. In this case we get the following action of C_n on $\mathbb{C}^2$:

$$r^k \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\left(\frac{2\pi k}{n}\right) & -\sin\left(\frac{2\pi k}{n}\right) \\ \sin\left(\frac{2\pi k}{n}\right) & \cos\left(\frac{2\pi k}{n}\right) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

   That is, r^k acts on a vector by rotating it over an angle $\frac{2\pi k}{n}$. This coincides with the geometric interpretation of the group C_n as the rotations over these angles for $k = 0, 1, \ldots, n-1$.

2. In a similar way we can define a standard representation of the dihedral group D_n. This group is generated by a rotation r and a reflection s, under the relations $r^n = e$, $s^2 = e$, and $rs = sr^{-1}$. Geometrically, r represents rotation over the angle $\frac{2\pi}{n}$ and s represents the reflection in the x-axis. The standard representation is defined by mapping r and s to the matrices corresponding to these geometrical operations. Precisely, this representation is defined on generators by

$$\begin{aligned} \rho : D_n &\to \mathrm{GL}_2(\mathbb{C}) \\ r &\mapsto \begin{bmatrix} \cos\left(\frac{2\pi k}{n}\right) & -\sin\left(\frac{2\pi k}{n}\right) \\ \sin\left(\frac{2\pi k}{n}\right) & \cos\left(\frac{2\pi k}{n}\right) \end{bmatrix} \\ s &\mapsto \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \end{aligned}$$

   This gives rise to a well-defined representation, because the matrices respect the relations between generators: we have $\rho(r)^n = I$, $\rho(s)^2 = I$, and $\rho(r)\rho(s) = \rho(s)\rho(s)^{-1}$, as can be easily verified by matrix multiplications.

3. Many groups can be regarded as groups of symmetries of some geometrical object embedded in a vector space V. Then the group elements are certain linear transformations on V. For example, the tetrahedron group consists of all linear transformations on $\mathbb{R}^3$ that map a tetrahedron placed symmetrically around the origin onto itself. In this case the group G of symmetries is a subgroup of $\mathrm{GL}(V)$. The inclusion $G \to \mathrm{GL}(V)$ is then automatically a group homomorphism and hence a representation of G.

4. There is a representation of the symmetric group S_n on the n-dimensional space $\mathbb{C}^n$. Write the standard basis of $\mathbb{C}^n$ as $e_1, \ldots, e_n$, and define the representation ρ on basis vectors via $\rho(\sigma)(e_i) = e_{\sigma(i)}$. That is, a permutation in S_n acts on a vector by permuting the basis vectors. This is called the standard representation of the symmetric group S_n.

5. For any group G there is a *trivial representation* on the vector space $\mathbb{C}$. This is the homomorphism $\rho(g) = I$ that sends every group element to the identity matrix. In terms of actions, it is given by $g \cdot x = x$ for all $g \in G$ and $x \in \mathbb{C}$. Actually there is a trivial representation of G on any vector space V, defined by $g \cdot v = v$, but the trivial representation on $\mathbb{C}$ is encountered most often.

6. For any group G we can also define a special representation, called the *regular representation* of G. The underlying vector space is $\mathbb{C}[G]$, which is the vector space generated by the basis $\{e_g \mid g \in G\}$. An arbitrary element of $\mathbb{C}[G]$ looks like $\sum_{g \in G} \lambda_g e_g$ for certain scalars $\lambda_g \in \mathbb{C}$. The group G acts on the basis vectors by $g \cdot e_h = e_{gh}$, and this action on basis vectors extends uniquely to a representation of G on $\mathbb{C}[G]$.

7. Let X be a set carrying an action of the group G. This induces a representation on the vector space $\mathbb{C}[X]$ with basis vectors e_x for $x \in X$. The representation acts on these basis vectors as $g \cdot e_x = e_{g \cdot x}$. Two of the above examples are special cases of this construction:

   - Any group G acts on itself by multiplication on the left. This action gives rise to the regular representation.
   - The group S_n acts on the set $\{1, \ldots, n\}$ by permuting the numbers. The resulting representation is the standard representation of S_n on $\mathbb{C}[\{1, \ldots, n\}] \cong \mathbb{C}^n$.

We will often denote a representation $\rho : G \to \mathrm{GL}(V)$ using the notation (V, ρ), leaving the group implicit. If no confusion is possible, we sometimes write just V instead of (V, ρ). The idea behind this notation is that a representation in Version 1 can be regarded as a vector space with a group action. In algebra it is common to omit the extra structure from the notation and focus on the underlying set. For example, a group $(G, \cdot)$ is often just written as G. Similarly, the group action in a representation is an additional structure on the underlying vector space, so we abbreviate (V, ρ) to V.

When working with maps between algebraic structures, the most interesting maps are the functions that preserve all structure. For groups, these are the homomorphisms, and for vector spaces these are linear maps. Since we regard representations as vector spaces with the additional structure of a group action, the natural maps between representations are linear maps that respect the group action. We will only consider maps between representations of the same group.

Definition 1.2. Let (V, ρ) and (W, σ) be representations of one group G. An *intertwiner* from V to W is a linear map $\phi : V \to W$ for which $\phi(g \cdot v) =$

$g \cdot \phi(v)$. If ϕ is a bijection, then it is called an *equivalence* or *isomorphism of representations*. In this case the representations (V, ρ) and (W, σ) are said to be *equivalent*.

The above definition is in line with Version 1 of the definition of a representation. We can rewrite it to obtain an equivalent definition in Version 2. The condition for an intertwiner says that $\phi(\rho(g)(v)) = \sigma(g)(\phi(v))$, so an intertwiner can also be defined as a linear map $\phi : V \to W$ such that $\phi \circ \rho(g) = \sigma(g) \circ \phi$ for all $g \in G$. Replacing the linear maps with matrices gives a definition in Version 3: an intertwiner is a matrix A such that for all g we have $A\rho(g) = \sigma(g)A$.

The representations of a fixed group G, together with the intertwiners between them, form a mathematical structure called a category. This category is denoted $\mathbf{Rep}_G$. In Chapter 5 we will say more about categories.

Example 1.3. The cyclic group C_n has a single generator r. The standard representation on $\mathbb{C}^2$ is defined on this generator by

$$\rho(r) = \begin{bmatrix} \cos\left(\frac{2\pi}{n}\right) & -\sin\left(\frac{2\pi}{n}\right) \\ \sin\left(\frac{2\pi}{n}\right) & \cos\left(\frac{2\pi}{n}\right) \end{bmatrix}$$

Define another representation of C_n by

$$\sigma(r) = \begin{bmatrix} e^{2\pi i/n} & 0 \\ 0 & e^{-2\pi i/n} \end{bmatrix}$$

We claim that these two representations are equivalent. To see this, let A be the matrix $\left[\begin{smallmatrix} 1 & i \\ 1 & -i \end{smallmatrix}\right]$. To check that this matrix is an intertwiner from ρ to σ, we have to show that $A\rho(g) = \sigma(g)A$ for all $g \in C_n$. Since C_n is generated by r, it suffices to check this for $g = r$. A quick calculation gives

$$\begin{aligned} A\rho(g) &= \begin{bmatrix} \cos\left(\frac{2\pi}{n}\right) + i\sin\left(\frac{2\pi}{n}\right) & -\sin\left(\frac{2\pi}{n}\right) + i\cos\left(\frac{2\pi}{n}\right) \\ \cos\left(\frac{2\pi}{n}\right) - i\sin\left(\frac{2\pi}{n}\right) & -\sin\left(\frac{2\pi}{n}\right) - i\cos\left(\frac{2\pi}{n}\right) \end{bmatrix} \\ &= \begin{bmatrix} e^{2\pi i/n} & ie^{2\pi i/n} \\ e^{-2\pi i/n} & -ie^{2\pi i/n} \end{bmatrix} = \sigma(g)A. \end{aligned}$$

Furthermore the matrix A is invertible, therefore it is an equivalence from ρ to σ.

1.2 Constructions

There are several constructions of new representations from old ones. Most of them are derived from constructions on vector spaces.

Direct sum. Let V and W be two representations of the same group G. Their direct sum $V \oplus W$ as vector spaces can be made into a representation by letting G act pointwise:

$$g \cdot (v, w) = (g \cdot v, g \cdot w)$$

The properties of the direct sum of vector spaces from Appendix A.2 also hold for the direct sum of representations.

Tensor product. Using tensor products of vector spaces and linear maps as described in Appendix A.3, we can define the tensor product of two representations. Suppose that (V, ρ) and (W, σ) are representations of G. Then $(V \otimes W, \rho \otimes \sigma)$ is again a representation, where the tensor product of ρ and σ is defined by

$$(\rho \otimes \sigma)(g) = \rho(g) \otimes \sigma(g).$$

Since $\rho(g) : V \to V$ and $\sigma(g) : W \to W$ are both linear maps, their tensor product as linear maps indeed gives a linear map $V \otimes W \to V \otimes W$, and it is easy to check that it yields a representation. It can also be written as an action:

$$g \cdot (v \otimes w) = (g \cdot v) \otimes (g \cdot w)$$

This defines the representation on all generating vectors $v \otimes w$. It extends uniquely to an assignment on all of $V \otimes W$.

The tensor product of representations has the same properties as the tensor product of vector spaces. We have the following analogue of Proposition A.8, replacing linear maps by intertwiners.

Proposition 1.4. *There is a one-to-one correspondence between bilinear intertwiners $\varphi : U \oplus V \to W$ and linear intertwiners $\psi : U \otimes V \to W$. On the generating vectors of the tensor product, this correspondence is given by $\psi(u \otimes v) = \varphi(u, v)$.*

From this it follows that the tensor product of representations is also commutative and associative. It has the trivial representation $\mathbb{C}$ as tensor unit, and distributes over direct sums. The proofs are almost the same as in the case of vector spaces. The only required extra step is checking that the isomorphisms obtained in the proofs are intertwiners.

Dual representation. We will use dual spaces and maps (see Appendix A.1) to define dual representations. If (V, ρ) is a representation, then the dual representation is (V^*, ρ^*), where V^* is the dual space of V and $\rho^*(g) = \rho(g^{-1})^*$. As an action, the dual representation is defined by $(g \cdot \phi)(v) = \phi(g^{-1} \cdot v)$. This is again a representation because:

$$\begin{aligned}\rho^*(gh) &= \rho((gh)^{-1})^* = \rho(h^{-1}g^{-1})^* \\ &= (\rho(h^{-1})\rho(g^{-1}))^* = \rho(g^{-1})^*\rho(h^{-1})^* = \rho^*(g)\rho^*(h)\end{aligned}$$

Note that we have to use g^{-1} instead of g in the definition, since otherwise ρ^* need not preserve multiplication, and hence we need not obtain a representation.

Linear Maps. If V and W are vector spaces, then the space of linear maps $\mathrm{Hom}(V, W)$ carries a representation of G whenever both V and W do. Suppose that (V, ρ) and (W, σ) are representations of G. The resulting representation on $\mathrm{Hom}(V, W)$ is denoted $\tau = \mathrm{Hom}(\rho, \sigma)$, and given by

$$\tau(g)(\varphi) = \sigma(g) \circ \varphi \circ \rho(g^{-1})$$

for $\varphi \in \mathrm{Hom}(V, W)$. In other words, to obtain the action from g on a linear map $\varphi : V \to W$, we first let g^{-1} act on an element of V, then we apply the function φ, and finally we act with g on the resulting vector in W:

$$\begin{tikzcd}
V \arrow[r, "\varphi"] & W \arrow[d, "\sigma(g)"] \\
V \arrow[u, "\rho(g^{-1})"] \arrow[r, dashed, "g\cdot\varphi"'] & W
\end{tikzcd}$$

In terms of actions, this says that

$$(g \cdot \varphi)(v) = g \cdot \varphi(g^{-1} \cdot v).$$

This construction is a generalization of the dual representation. If we take the trivial representation $\mathbb{C}$ for W, then the representation $\mathrm{Hom}(V, \mathbb{C})$ coincides with V^*. This is the reason why we include $\rho(g^{-1})$ in the definition.

We have the following analogues of Propositions A.10 and A.11, replacing the vector spaces with representations.

Proposition 1.5. *Let (U, ρ), (V, σ) and (W, τ) be representations of G.*

1. *The representations $\mathrm{Hom}(U \otimes V, W)$ and $\mathrm{Hom}(U, \mathrm{Hom}(V, W))$ are equivalent.*

2. *The representations $V^* \otimes W$ and $\mathrm{Hom}(V, W)$ are equivalent.*

The proofs are the same as for vector spaces, but additionally we have to check that the isomorphisms obtained in the proofs are intertwiners.

Invariant subspaces. Let (V, ρ) be a representation of a group G. Sometimes it is useful to restrict ρ to a subspace of V to obtain a smaller representation. This only works if the subspace is closed under the action of the group. A subspace $U \subseteq V$ is called *invariant* if $g \cdot u \in U$ for every $u \in U$. The representation $(U, \rho|_U)$ is then called a *subrepresentation* of (V, ρ).

Change of group. Given a group homomorphism $\varphi : H \to G$, it is possible to transform representations of G into representations of H. If (V, ρ) is a representation of G, then $\rho \circ \varphi : H \to \mathrm{GL}(V)$ is a representation of H. This representation is written as $\varphi^*(V, \rho)$. In Version 1, its action is given by $h \cdot v = \varphi(h) \cdot v$.

Change of vector space. Similarly, a group homomorphism $\psi : \mathrm{GL}(V) \to \mathrm{GL}(W)$ can be used to transform a representation (V, ρ) of G into the representation $(W, \psi \circ \rho)$ of G. This representation is denoted $\psi^*(V, \rho)$. As a special case, if $\alpha : V \to W$ is a linear isomorphism, then $\psi(A) = \alpha \circ A \circ \alpha^{-1}$ defines a homomorphism $\mathrm{GL}(V) \to \mathrm{GL}(W)$.

1.3 Decomposition into irreducible representations

Given a finite group, we can obtain much information about the group by describing all of its representations, up to equivalence. Finding all representations is a hard problem in general. It helps if we can find certain “building blocks” for the representations, and then show that any representation can be decomposed into these building blocks. If the building blocks are easy to describe, this will also make the problem of finding all representations more achievable. We will use irreducible representations as building blocks.

Definition 1.6. A non-zero representation (V, ρ) is called *irreducible* if its only invariant subspaces are 0 and V itself.

Example 1.7. Consider the representation of C_4 on $\mathbb{C}^2$ determined by

$$\rho(r) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

We would like to know whether ρ is irreducible.

Suppose that $U \subseteq \mathbb{C}^2$ is invariant and $0 \neq U \neq \mathbb{C}^2$. Then U must be one-dimensional, say that it is spanned by the single vector $u = (u_1, u_2)$. Since U is invariant, we have $\rho(r)(u) = \lambda u$ for some $\lambda \in \mathbb{C}$. This gives the equation

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \lambda \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

This amounts to saying that λ is an eigenvalue of the matrix $\left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$. The eigenvalues are $\lambda = \pm i$. The eigenvector $(1, -i)$ belonging to the eigenvalue i spans an invariant subspace, so ρ is not irreducible.

Irreducible representations are suitable as building blocks, since they cannot be decomposed into smaller representations. We will first show that every representation of a group can be decomposed as a direct sum of irreducible representations. Later we will develop techniques to describe all irreducible representations. This will give a complete description of all representations of a given group.

To prove that every representation is a direct sum of irreducibles, we will use a common technique in the theory of representations, called *averaging*. It can be applied to several objects from linear algebra, to make them behave better with respect to the group action. We will describe the averaging technique for vectors, linear maps, and inner products.

Averaging vectors. Let (V, ρ) be a representation of G. We saw that the most interesting subspaces of V are the invariant subspaces. Similarly the most interesting vectors in V are the invariant vectors. A vector $v \in V$ is said to be *invariant* if $g \cdot v = v$ for all $g \in G$. We shall write the collection of all invariant vectors as V^G.

If $v \in V$ is not invariant, then we can turn it into an invariant vector by averaging it. The average of v is defined as

$$\mathrm{Av}(v) = \frac{1}{\#G} \sum_{g \in G} g \cdot v.$$

The vector $\mathrm{Av}(v)$ is indeed invariant, because for any $h \in G$ we get

$$h \cdot \mathrm{Av}(v) = \frac{1}{\#G} \sum_{g \in G} h \cdot (g \cdot v) = \frac{1}{\#G} \sum_{g \in G} (hg) \cdot v = \frac{1}{\#G} \sum_{g' \in G} g' \cdot v = \mathrm{Av}(v).$$

For the third equality sign, we used the reindexing substitution $g' = hg$. Since g runs over all group elements, so does g', so the sums are indeed equal. Observe that averaging an invariant vector v gives back the vector itself, because we sum over $\#G$ copies of v and then divide by $\#G$. Thus averaging is a projection map $\mathrm{Av} : V \to V^G$ onto the invariant vectors.

Averaging linear maps. Let (V, ρ) and (W, σ) be two representations of G, and let $\varphi : V \to W$ be a linear map between the underlying vector spaces. The map φ need not be an intertwiner, but we can always make it into an intertwiner by averaging:

$$\mathrm{Av}(\varphi) : V \to W, \quad \mathrm{Av}(\varphi)(v) = \frac{1}{\#G} \sum_{g \in G} g \cdot \varphi(g^{-1} \cdot v)$$

In fact, averaging a linear map is a special case of averaging a vector. This is because a linear map $\varphi : V \to W$ is a vector in the space $\mathrm{Hom}(V, W)$, and we saw earlier that this vector space carries an action of G given by $(g \cdot \varphi)(v) = g \cdot \varphi(g^{-1} \cdot v)$. Therefore, if we take the average of φ as a *vector* in $\mathrm{Hom}(V, W)$, we get

$$\mathrm{Av}(\varphi)(v) = \frac{1}{\#G} \sum_{g \in G} (g \cdot \varphi)(v) = \frac{1}{\#G} \sum_{g \in G} g \cdot \varphi(g^{-1} \cdot v),$$

and this is exactly the same as what we get by taking the average of φ as a linear map.

We know that averaging a vector gives an invariant vector. An invariant vector in $\mathrm{Hom}(V, W)$ is a map $\varphi : V \to W$ satisfying $(g \cdot \varphi)(v) = \varphi(v)$ for all $v \in V$, which happens if and only if $g \cdot \varphi(g^{-1} \cdot v) = \varphi(v)$. This is in turn equivalent to $\varphi(g^{-1} \cdot v) = g^{-1} \cdot \varphi(v)$ for all g, or equivalently $\varphi(g \cdot v) = g \cdot \varphi(v)$. In other words, φ is an intertwiner. Thus we have proven that invariant vectors in $\mathrm{Hom}(V, W)$ are the same as intertwiners, and that averaging a linear map always gives an intertwiner. Furthermore, again because averaging a map is a special case of averaging a vector, the average of an intertwiner is the intertwiner itself.

Averaging inner products. Take a representation (V, ρ) of G and let $\langle -, - \rangle$ be an inner product on V. This inner product is said to be *invariant* whenever $\langle g \cdot v, g \cdot w \rangle = \langle v, w \rangle$ for each $g \in G$ and all $v, w \in V$. If the inner product is not invariant, then it can be easily converted into an invariant inner product $\langle -, - \rangle'$ by averaging:

$$\langle v, w \rangle' = \frac{1}{\#G} \sum_{g \in G} \langle g \cdot v, g \cdot w \rangle$$

This is again an inner product, because all properties required for an inner product are preserved by the construction: $\langle -, - \rangle'$ is sesquilinear because the original inner product $\langle -, - \rangle$ is sesquilinear, it is skew-symmetric because $\langle -, - \rangle$ is, and it is non-degenerate because $\langle -, - \rangle$ is. Furthermore, it is straightforward to show that the new inner product is invariant, and that the average of an invariant inner product is the inner product itself.

If V is a vector space equipped with an inner product $\langle -, - \rangle$, then every subspace $U \subseteq V$ has an orthogonal complement $U^\perp$ with respect to the inner product. Recall that $U^\perp = \{v \in V \mid \langle u, v \rangle = 0 \text{ for all } u \in U\}$. If (V, ρ) is a representation of G and U is an invariant subspace, then its orthogonal complement need not be invariant again. However, if the inner product is invariant, then the complement is in fact again invariant. This will be used to prove the following result.

Lemma 1.8. *Let (V, ρ) be a representation of G, and let $U \subseteq V$ be an invariant subspace. Then there exists an invariant subspace U' of V such that $V = U \oplus U'$.*

Proof. Take any inner product on V and take its average to obtain an invariant inner product $\langle -, - \rangle$. Then let U' be the orthogonal complement of U with respect to $\langle -, - \rangle$. To check that U' is invariant, take any $v \in U'$. Then for each $u \in U$ we have:

$$\langle u, g \cdot v \rangle = \langle gg^{-1} \cdot u, g \cdot v \rangle = \langle g^{-1} \cdot u, v \rangle = 0$$

The second equality sign holds because the inner product is invariant, and the last one because $v \in U' = U^\perp$ and U is invariant. Therefore U' is invariant. □

Now we can prove the main result alluded to in the beginning of this section: every representation decomposes into irreducibles.

Theorem 1.9. *Any representation (V, ρ) can be decomposed as a direct sum $V = U_1 \oplus \cdots \oplus U_n$, where each U_i is an irreducible subrepresentation of V.*

Proof. From the above lemma it follows that if V is not irreducible, then it can be written as $V = V_1 \oplus V_2$, where the dimension of V_1 and V_2 is strictly smaller than the dimension of V. Repeating this process with V_1 and V_2 et cetera, V can in the end be decomposed into a direct sum $U_1 \oplus \cdots \oplus U_n$, in such a way that the process cannot be repeated anymore with any of the components. In this situation, each U_i is irreducible. □

Note that this proof only works for finite-dimensional vector spaces.

1.4 Schur's Lemma

We will now start developing ways to find all irreducible representations of a given group. First we will prove Schur's Lemma, which says that intertwiners between irreducible representations are of a very simple form. We will use this lemma to describe the irreducible representations of abelian groups.

Lemma 1.10 (Schur). *Let (V, ρ) and (W, σ) be two irreducible representations of G.*

1. *Any intertwiner $\varphi : V \to W$ is either 0 or an equivalence of representations.*

2. *Any intertwiner $\varphi : V \to V$ is a scalar multiple of the identity: $\varphi = \lambda \operatorname{id}$ for some $\lambda \in \mathbb{C}$.*

Proof.

1. The kernel $\ker(\varphi)$ is an invariant subspace of V, because if $\varphi(v) = 0$, then $\varphi(g \cdot v) = g \cdot \varphi(v) = g \cdot 0 = 0$. Since V is irreducible, $\ker(\varphi)$ is either 0 or V. If $\ker(\varphi) = V$, then φ is the zero map and the assertion is proven. Otherwise $\ker(\varphi) = 0$, so φ is injective. The image $\operatorname{im}(\varphi)$ is an invariant subspace of W, since if $w = \varphi(v)$, then $g \cdot w = g \cdot \varphi(v) = \varphi(g \cdot v) \in \operatorname{im}(\varphi)$. Hence $\operatorname{im}(\varphi)$ is either 0 or all of W. The image cannot be zero, since φ is injective. Thus $\operatorname{im}(\varphi) = W$, so φ is injective and surjective, hence an equivalence of representations.

2. Let λ be an eigenvalue of φ with corresponding eigenspace $U \subseteq V$. Then U is invariant, because for all $u \in U$ we have $\varphi(g \cdot u) = g \cdot \varphi(u) = g \cdot \lambda u = \lambda g \cdot u$, hence $g \cdot u \in U$. Consequently U is 0 or V. It cannot be 0 since eigenvalues have a non-zero eigenspace. Thus $U = V$, which means that $\varphi(v) = \lambda v$ for all $v \in V$. This shows that φ is a scalar multiple of the identity. □

Remark. Schur's Lemma is sometimes phrased in the following way: let V and W be two irreducible representations of G.

1. If V and W are not equivalent, then any intertwiner from V to W is zero.

2. If V and W are equivalent, then any two intertwiners from V to W are scalar multiples of each other.

As an exercise, check that this is equivalent to the phrasing above.

In the proof of the second part, we used that every operator on a finite-dimensional vector space has an eigenvalue. This holds only for complex vector spaces, so here we implicitly used the assumption that we work with finite-dimensional complex vector spaces.

Even though Schur's Lemma is not hard to prove, it is fundamental in the theory of representations and has many useful consequences. As a first application, we will use it to find all irreducible representations of abelian groups.

Proposition 1.11. *All irreducible representations of an abelian group are one-dimensional.*

Proof. Let (V, ρ) be an irreducible representation of the abelian group G. For any fixed $g \in G$, the map $\rho(g) : V \to V$ is an intertwiner:

$$\rho(g) \circ \rho(h) = \rho(gh) = \rho(hg) = \rho(h) \circ \rho(g)$$

By Schur's Lemma, $\rho(g) = \lambda_g \,\mathrm{id}$ for some scalar λ_g. Pick any non-zero vector v in V. We will show that the subspace $\mathbb{C}v$ spanned by v is an invariant subspace of V. An element in this subspace is of the form αv for some $\alpha \in \mathbb{C}$, and we have $\rho(g)(\alpha v) = \lambda_g \alpha v \in \mathbb{C}v$. Hence $\mathbb{C}v$ is invariant and non-zero, so it is equal to the full space V. But $\mathbb{C}v$ is one-dimensional, and hence V is one-dimensional, as required. $\square$

Example 1.12. We will determine all irreducible representations of the group C_4. Such a representation ρ is one-dimensional, so it amounts to a homomorphism $\rho : C_4 \to \mathbb{C}^\times$. Since $r^4 = e$, ρ must satisy $\rho(r)^4 = 1$. Hence $\rho(r)$ can be either 1, i, -1, or $-i$. Since the representation is determined by its value on r, these four representations are all irreducible representations of C_4. (Why are these inequivalent?)

A second application of Schur's Lemma is the result that the regular representation $\mathbb{C}[G]$ of a group G contains all irreducible representations as components. The proof of this fact also requires the following lemma, whose proof is left as an exercise.

Lemma 1.13. *For every surjective intertwiner $p : (V, \rho) \to (W, \sigma)$ there exists a section $s : (W, \sigma) \to (V, \rho)$ (meaning that $p \circ s = \mathrm{id}$) that is an intertwiner itself.*

Proposition 1.14. *Every irreducible representation of G is equivalent to a subrepresentation of the regular representation $\mathbb{C}[G]$.*

Proof. Let (V, ρ) be an irreducible representation and pick any non-zero vector $v \in V$. Consider the function $G \to V$ that maps a group element g to $g \cdot v$. Since group elements of G form a basis for $\mathbb{C}[G]$, this function can be extended uniquely to a linear map $\varphi : \mathbb{C}[G] \to V$. Explicitly, this linear map looks like

$$\varphi\left(\sum_g \lambda_g e_g\right) = \sum_g \lambda_g (g \cdot v).$$

The map φ is an intertwiner, because on basis vectors it satisfies

$$\varphi(h \cdot e_g) = \varphi(e_{hg}) = hg \cdot v = h\varphi(e_g).$$

Its image $\mathrm{im}(\varphi)$ is an invariant subspace of V, so since V is irreducible and φ is non-zero, φ must be surjective. Lemma 1.13 gives a section $\psi : V \to \mathbb{C}[G]$ that is an intertwiner. Since V is irreducible, ψ is an isomorphism onto its image, which is a subrepresentation of the regular representation. $\square$

2 Characters

2.1 Basic properties

Our main goal is to classify all representations of a given group up to equivalence. For this we need techniques to check that two given representations are equivalent, to find all irreducible representations, and to decompose any representation into irreducibles. It will turn out that we can accomplish all of this by looking at the traces of the linear maps involved in a representation.

Recall that the *trace* of a square matrix is defined as the sum of its entries on the diagonal, that is $\operatorname{tr}(A) = \sum_i a_{ii}$. We can also define the trace of a linear map $\varphi : V \to V$: pick a basis for V, write φ as a matrix A with respect to this basis, and define $\operatorname{tr}(\varphi) = \operatorname{tr}(A)$. This definition is only sensible if it does not depend on the choice of basis. To prove that this is the case, we need the following crucial property of the trace, called the *cyclic property*:

$$\operatorname{tr}(AB) = \operatorname{tr}(BA)$$

This property can be proven by simply expanding the matrix multiplications on both sides. If B is an invertible matrix, the property implies that $\operatorname{tr}(A) = \operatorname{tr}(BAB^{-1})$. To show that the definition of $\operatorname{tr}(\varphi)$ is basis-independent, let A and A' be two matrix representations of φ. Then $A' = BAB^{-1}$ for some invertible matrix B, so their traces are equal.

Taking the trace of a representation gives the object that we will study in this chapter.

Definition 2.1. The *character* of a representation $\rho : G \to \mathrm{GL}(V)$ is the function

$$\chi_\rho : G \to \mathbb{C}, \quad g \mapsto \operatorname{tr}(\rho(g)).$$

If ρ is understood, this function is sometimes written as χ_V.

The cyclic property of the trace entails that the characters of equivalent representations are the same. This is because if (V, ρ) and (W, σ) are equivalent, then there exists an invertible intertwiner $\varphi : V \to W$. The intertwining property $\varphi \circ \rho(g) = \sigma(g) \circ \varphi$ is equivalent to $\sigma(g) = \varphi \circ \rho(g) \circ \varphi^{-1}$, so by the cyclic property of the trace, $\chi_\rho = \chi_\sigma$.

One of the main results of this chapter will be that the converse also holds: if two representations have the same character, then they are equivalent. Hence the character completely characterizes a representation up to equivalence, as the name "character" suggests. This gives a very easy criterion to determine whether two representations are equivalent: instead of trying to find an intertwiner or proving that none exists, it is enough to calculate the traces of the matrices in the representation and checking if these are equal.

Before proving this main result, we will establish some basic properties of characters. These are derived from the corresponding properties of the trace operation.

Proposition 2.2. *Let (V, ρ) and (W, σ) be representations of G.*

1. $\chi_{V\oplus W}(g) = \chi_V(g) + \chi_W(g)$
2. $\chi_{V\otimes W}(g) = \chi_V(g) \cdot \chi_W(g)$
3. $\chi_{V^*}(g) = \chi_V(g^{-1}) = \overline{\chi_V(g)}$
4. $\chi_{\mathrm{Hom}(V,W)}(g) = \overline{\chi_V(g)}\chi_W(g)$

Proof.

1. The matrices for $(\rho \oplus \sigma)(g)$ are obtained by putting those for $\rho(g)$ and $\sigma(g)$ as blocks on the diagonal. Therefore

$$\chi_{V\oplus W}(g) = \mathrm{tr}((\rho \oplus \sigma)(g)) = \mathrm{tr}(\rho(g)) + \mathrm{tr}(\sigma(g)) = \chi_V(g) + \chi_W(g).$$

2. If A and B are two matrices, then the formula for the Kronecker product is $(A \otimes B)_{(i,k)(j,l)} = A_{ij}B_{kl}$. This gives a formula for the trace of the tensor product:

$$\mathrm{tr}(A \otimes B) = \sum_{i,k} A_{ii}B_{kk} = \left(\sum_i A_{ii}\right)\left(\sum_k B_{kk}\right) = \mathrm{tr}(A)\,\mathrm{tr}(B)$$

   This shows that $\chi_{V\otimes W}(g) = \chi_V(g)\chi_W(g)$.

3. We will first prove that $\chi_{V^*}(g) = \chi_V(g^{-1})$. Recall that the dual representation is defined by $\rho^*(g) = \rho(g^{-1})^*$. The matrix corresponding to the dual $\rho(g^{-1})^*$ is the transpose of the matrix corresponding to $\rho(g^{-1})$. Since every matrix has the same trace as its transpose, we get

$$\chi_{V^*}(g) = \mathrm{tr}(\rho^*(g)) = \mathrm{tr}(\rho(g^{-1})^{\mathsf{T}}) = \mathrm{tr}(\rho(g^{-1})) = \chi_V(g^{-1}).$$

   For the second equality, let $\langle -, - \rangle$ be an invariant inner product on V. We will first show that $\rho(g^{-1}) = \rho(g)^\dagger$, where the dagger indicates the hermitian adjoint of the operator with respect to the inner product. By Proposition A.13, it suffices to show that the inner products with all vectors are equal. This holds because of the following computation:

$$\begin{aligned}
\langle v, \rho(g^{-1})(w)\rangle &= \langle \rho(g)(v), \rho(g)\rho(g^{-1})(w)\rangle && \text{(inner product is invariant)}\\
&= \langle \rho(g)(v), w\rangle && \text{(ρ is a homomorphism)}\\
&= \langle v, \rho(g)^\dagger(w)\rangle && \text{(hermitian adjoint)}
\end{aligned}$$

   Since this holds for all v and w, we have $\rho(g^{-1}) = \rho(g)^\dagger$, hence

$$\chi_\rho(g^{-1}) = \mathrm{tr}(\rho(g^{-1})) = \mathrm{tr}(\rho(g)^\dagger) = \mathrm{tr}(\overline{\rho(g)}^{\mathsf{T}}) = \overline{\mathrm{tr}(\rho(g))} = \overline{\chi_\rho(g)}.$$

4. This follows from $\mathrm{Hom}(V,W) \cong V^* \otimes W$ and the properties 2 and 3. $\square$

2.2 Schur's orthogonality relations

We wish to use characters as a tool to determine whether representations are equivalent, to find all irreducible representations of a group, and to decompose arbitrary representations into irreducibles. This will be achieved by proving certain equalities for characters, called Schur's orthogonality relations. To state these relations in an abstract way, we need the framework of class functions.

Recall that two elements g, h in a group G are conjugate if there exists an element $a \in G$ such that $h = aga^{-1}$. In this case we write $g \sim h$. The relation $\sim$ is an equivalence relation, and an equivalence class of $\sim$ is called a conjugacy class. The conjugacy class of an element g is written as $(g) = \{h \in G \mid \text{there exists } a \text{ such that } h = aga^{-1}\}$. The collection of all conjugacy classes in G is denoted (G).

Definition 2.3. A *class function* on a group G is a function $\varphi : G \to \mathbb{C}$ that is invariant on conjugacy classes, i.e. $\varphi(hgh^{-1}) = \varphi(g)$ for all $g, h \in G$.

A class function is the same as a function $\varphi : (G) \to \mathbb{C}$. Characters are examples of class functions, since $\chi_\rho(hgh^{-1}) = \mathrm{tr}(\rho(hgh^{-1})) = \mathrm{tr}(\rho(h)\rho(g)\rho(h)^{-1}) = \mathrm{tr}(\rho(g)) = \chi_\rho(g)$. The set of all class functions on G forms a vector space $\mathcal{C}\ell(G)$ under pointwise operations. Endow this vector space with the following inner product:

$$\langle \varphi, \psi \rangle = \frac{1}{\#G} \sum_{g \in G} \overline{\varphi(g)} \psi(g)$$

All important properties of characters will follow from the next result.

Theorem 2.4. *Let G be a finite group. The characters χ_V of the (inequivalent) irreducible representations V of G form an orthonormal basis for the space of class functions $\mathcal{C}\ell(G)$.*

We only have to take characters of inequivalent representations, since equivalent representations give the same characters.

The proof of this theorem consists of two parts. First we will prove that the characters χ_V are orthonormal. This means that

$$\langle \chi_V, \chi_W \rangle = \frac{1}{\#G} \sum_{g \in G} \overline{\chi_V(g)} \chi_W(g) = 0$$

whenever V and W are irreducible and not equivalent, and that

$$\langle \chi_V, \chi_V \rangle = \frac{1}{\#G} \sum_{g \in G} |\chi_V(g)|^2 = 1$$

for every irreducible representation V. These two formulas are called *Schur's orthogonality relations.* They can be compactly phrased as "Characters of irreducible representations are orthogonal and normalized". From general facts in linear algebra it follows that orthogonal vectors are always linearly independent.

The second step to complete the proof that the χ_V form a basis is to show that they span the whole space $\mathcal{C}\ell(G)$.

The proof of Schur's orthogonality relations requires Schur's Lemma and the following fact.

Proposition 2.5 (Dimension formula)**.** *For any representation (V, ρ) of G, the dimension of the space of invariant vectors is*

$$\dim(V^G) = \frac{1}{\#G} \sum_{g \in G} \chi_V(g).$$

Proof. Consider the averaging map

$$\mathrm{Av} : V \to V, \quad v \mapsto \frac{1}{\#G} \sum_{g \in G} g \cdot v.$$

The average of any vector is invariant, and the average of an invariant vector is the vector itself, so $\mathrm{Av}^2 = \mathrm{Av}$. This implies that the eigenvalues of the operator Av are all 0 or 1. Choose a basis in which the matrix for Av is diagonal; then it has only zeroes and ones on the diagonal. The trace of this matrix is the number of ones, and that is also the dimension of the image, so $\mathrm{tr}(\mathrm{Av}) = \dim(\mathrm{im}(\mathrm{Av}))$. Since Av maps onto V^G, $\dim(\mathrm{im}(\mathrm{Av})) = \dim(V^G)$. Furthermore, $\mathrm{tr}(\mathrm{Av}) = \frac{1}{\#G} \sum_g \chi_V(g)$, from which the result follows. □

Corollary 2.6. *Let V and W be representations of G. The dimension of the space of intertwiners from V to W is $\langle \chi_V, \chi_W \rangle$.*

Proof. By part 4 of Proposition 2.2 and the previous proposition we have

$$\begin{aligned}\langle \chi_V, \chi_W \rangle &= \frac{1}{\#G} \sum_{g \in G} \overline{\chi_V(g)} \chi_W(g) \\ &= \frac{1}{\#G} \sum_{g \in G} \chi_{\mathrm{Hom}(V,W)}(g) \\ &= \dim(\mathrm{Hom}(V, W)^G).\end{aligned}$$

The invariant vectors in $\mathrm{Hom}(V, W)$ are precisely the intertwiners from V to W, which gives the result. □

Schur's Lemma tells us that the space of intertwiners between two inequivalent irreducible representations is zero, and that the space of intertwiners from an irreducible representation to itself is one-dimensional. This gives the following two consequences of Corollary 2.6.

Corollary 2.7. *If V and W are inequivalent irreducible representations of G, then $\langle \chi_V, \chi_W \rangle = 0$.*

Corollary 2.8. *If V is an irreducible representation of G, then $\langle \chi_V, \chi_V \rangle = 1$.*

This finishes the proof of the orthogonality relations, and hence we know that the characters of irreducible representations are linearly independent. Now we will finish the proof of Theorem 2.4.

Proposition 2.9. *The class functions χ_V, where V runs over the irreducible representations of G, span the vector space $\mathcal{C}\ell(G)$.*

Proof. Since the characters χ_V are orthonormal, it suffices to show that the condition $\langle f, \chi_V\rangle = 0$ implies $f = 0$, for any class function $f : G \to \mathbb{C}$.

For any representation (V, ρ), we can define an intertwiner

$$f_V : V \to V, \quad v \mapsto \frac{1}{\#G} \sum_{g\in G} \overline{f(g)} g \cdot v.$$

This is indeed an intertwiner, because

$$\begin{aligned} f_V(h \cdot v) &= \frac{1}{\#G} \sum_{g\in G} \overline{f(g)} gh \cdot v = \frac{1}{\#G} \sum_{k\in G} \overline{f(hkh^{-1})} hk \cdot v \\ &= \frac{1}{\#G} \sum_{k\in G} \overline{f(k)} hk \cdot v = h \cdot f_V(v), \end{aligned}$$

where we used the substitution $k = h^{-1}gh$ and the fact that f is a class function.

Suppose that the representation V is irreducible. Then, by Schur's Lemma, $f_V = \lambda \,\mathrm{id}$ for some scalar λ. We can find the scalar λ by computing the trace of f_V in two different ways. Since $f_V = \lambda \,\mathrm{id}$, we obtain $\mathrm{tr}(f_V) = \lambda \dim V$. On the other hand, the definition of f_V gives $\mathrm{tr}(f_V) = \frac{1}{\#G} \sum_g \overline{f(g)} \chi_V(g) = \langle f, \chi_V\rangle$, which is zero by assumption. Therefore $\lambda = 0$ and hence f_V is the zero function.

Since any representation can be decomposed into irreducible representations, and $f_{V\oplus W} = f_V + f_W$, the function f_V is in fact zero for any representation V. In particular, this holds for the regular representation $V = \mathbb{C}[G]$ with basis $\{e_g \mid g \in G\}$. Thus we have $f_V(e_g) = 0$ for every $g \in G$. Applying this to the unit of the group gives $\sum_g \overline{f(g)} e_g = 0$, hence $\overline{f(g)} = 0$ for all g, so $f = 0$. □

2.3 Consequences of the orthogonality relations

Theorem 2.4 has many consequences that are useful to obtain properties of representations using their characters.

Corollary 2.10. *The number of irreducible representations of G is equal to the number of conjugacy classes in G.*

Proof. Since class functions can be seen as functions from (G) to $\mathbb{C}$, the dimension of $\mathcal{C}\ell(G)$ is the number of conjugacy classes. But $\mathcal{C}\ell(G)$ has characters of irreducible representations as basis, so its dimension is also equal to the number of irreducibles. □

To state the other corollaries we will simplify the notation a bit. In the following, fix a group G. We will implicitly assume that all representations are representations of this group G. Call its irreducible representations $V_1, \dots, V_k$ and write the corresponding characters as $\chi_1, \dots, \chi_k$.

Corollary 2.11. *If V is any representation of G, then the irreducible representation V_i occurs $\langle \chi_i, \chi_V \rangle$ times in the decomposition of V.*

Proof. Write the decomposition of V as $V = n_1 V_1 \oplus \cdots \oplus n_k V_k$. Then V_i occurs n_i times in the decomposition of V. By Schur's orthogonality relations we have

$$\langle \chi_i, \chi_V \rangle = \langle \chi_i, n_1\chi_1 + \cdots + n_k\chi_k \rangle = n_1\langle \chi_i, \chi_1 \rangle + \cdots n_k \langle \chi_i, \chi_k \rangle = n_i. \qquad \square$$

Corollary 2.12. *Two representations V and W are equivalent if and only if $\chi_V = \chi_W$.*

Proof. Equivalent representations have the same character because characters are class functions. Conversely, if $\chi_V = \chi_W$, then every irreducible representation occurs the same number of times in V and W. Therefore V and W have the same decomposition into irreducibles and hence they are equivalent. $\square$

Corollary 2.13. *A representation V is irreducible if and only if $\langle \chi_V, \chi_V \rangle = 1$.*

Proof. Decompose V as $n_1 V_1 \oplus \cdots \oplus n_k V_k$ into irreducibles. By the orthogonality relations,

$$\langle \chi_V, \chi_V \rangle = \langle n_1\chi_1 + \cdots + n_k\chi_k, n_1\chi_1 + \cdots + n_k\chi_k \rangle = n_1^2 + \cdots + n_k^2.$$

This is equal to 1 if and only if exactly one of the n_i is 1 and the rest is 0, which happens precisely if V is irreducible. $\square$

As an application, we will compute the decomposition of the regular representation into irreducibles. We have already seen in Proposition 1.14 that it contains all irreducibles as subrepresentations, but now we will prove that the number of times an irreducible representation occurs is equal to its dimension.

Corollary 2.14. *Let $V_1, \ldots, V_k$ be a list of all irreducible representations of G. Then the regular representation of G decomposes as*

$$\mathbb{C}[G] = (\dim V_1)V_1 \oplus \cdots \oplus (\dim V_k)V_k.$$

Proof. We will first calculate the character of the regular representation. Write $G = \{g_1, \ldots, g_n\}$. A group element g acts on a basis vector e_{g_i} via $g \cdot e_{g_i} = e_{gg_i}$. Therefore the matrix entries of $\rho(g)$ are

$$(\rho(g))_{ij} = \begin{cases} 1 & \text{if } g_i = gg_j \\ 0 & \text{otherwise} \end{cases}$$

This gives the following entries on the diagonal:

$$(\rho(g))_{ii} = \begin{cases} 1 & \text{if } g = e \\ 0 & \text{otherwise} \end{cases}$$

Hence the trace of $\rho(e)$ is $\#G$, and the trace of $\rho(g)$ is zero for $g \neq e$.

We will use this character to find the decomposition with Corollary 2.11. The number of times that V_i occurs is

$$\langle \chi_i, \chi_{\mathbb{C}[G]} \rangle = \frac{1}{\#G} \sum_{g \in G} \overline{\chi_i(g)} \chi_{\mathbb{C}[G]}(g) = \frac{1}{\#G} \overline{\chi_i(e)} \#G = \operatorname{tr}(\mathrm{id}_{V_i}) = \dim V_i,$$

which gives the desired decomposition. □

Corollary 2.15. *Let $V_1, \ldots, V_k$ be a list of all irreducible representations of G. Then*

$$\sum_{i=1}^{k} (\dim V_i)^2 = \#G.$$

Proof. We will prove this by computing the inner product $\langle \chi_{\mathbb{C}[G]}, \chi_{\mathbb{C}[G]} \rangle$ in two different ways. Since the character of the regular representation is $\chi_{\mathbb{C}[G]}(e) = \#G$ and $\chi_{\mathbb{C}[G]}(g) = 0$ for $g \neq e$, we get

$$\langle \chi_{\mathbb{C}[G]}, \chi_{\mathbb{C}[G]} \rangle = \frac{1}{\#G} \sum_{g \in G} |\chi_{\mathbb{C}[G]}(g)|^2 = \frac{1}{\#G} (\#G)^2 = \#G.$$

On the other hand, the decomposition of $\mathbb{C}[G]$ into irreducibles gives

$$\langle \chi_{\mathbb{C}[G]}, \chi_{\mathbb{C}[G]} \rangle = \sum_{i=1}^{k} (\dim V_i)^2,$$

which proves the result. □

The above formula is useful to find the dimensions of the irreducible representations. It is often easier to use if we already know the dimensions of some of them. In particular, it is good to know the number of one-dimensional representations. This number can be found using the abelianization of the group G. Recall that the commutator subgroup of G is the subgroup $[G, G]$ generated by all commutators $ghg^{-1}h^{-1}$ for $g, h \in G$. This is always a normal subgroup. The abelianization of G is then the quotient $G^{\mathrm{ab}} = G/[G, G]$.

Proposition 2.16. *The number of one-dimensional representations of G is $\#G^{\mathrm{ab}}$.*

Proof. If $\rho : G \to \mathrm{GL}(V)$ is a one-dimensional representation, then $[G, G] \subseteq \ker(\rho)$ since $\mathrm{GL}(V)$ is abelian. It follows that ρ factors uniquely through the abelianization G^{ab}. In other words, ρ can be decomposed as $\rho = \rho^{\mathrm{ab}} \circ \pi$, where π is the quotient map $\pi : G \to G^{\mathrm{ab}} = G/[G, G]$:

$$\begin{tikzcd} G \arrow[r, "\rho"] \arrow[d, "\pi"'] & \mathrm{GL}(V) \\ G^{\mathrm{ab}} \arrow[ur, "\rho^{\mathrm{ab}}"'] & \end{tikzcd}$$

Hence one-dimensional representations of G correspond to one-dimensional representations of G^{ab}. Since G^{ab} is abelian, its number of one-dimensional representations is equal to $\#G^{\mathrm{ab}}$. $\square$

2.4 Character tables

Since all representations are determined by their characters, and any representation can be decomposed into irreducible ones, we know everything about the representations of a group once we know the characters of its irreducible representations. This information can be neatly organized in a table, called the *character table* of the group. The rows of the character table are labeled by characters of irreducible representations, and the columns by conjugacy classes in the group. Thus a typical character table looks like:

$$\begin{array}{c|ccc} & (g_1) & \cdots & (g_k) \\ \hline \chi_1 & & & \\ \vdots & & & \\ \chi_k & & & \end{array}$$

Here $g_1, \dots, g_k$ are representatives of the conjugacy classes of the group, and $\chi_1, \dots, \chi_k$ are the characters of the irreducible representations $V_1, \dots, V_k$. Since characters are constant on conjugacy classes, this completely determines all characters.

We can find the character table of a given group G by using the facts from the previous section. For small groups, it is usually enough to follow this procedure for finding the characters:

1. Find the conjugacy classes of the group. The number of irreducible representations is equal to the number of classes.

2. The characters of the one-dimensional representations, which are always irreducible, can be found either by a direct calculation or by using Proposition 2.16.

3. The size of the other irreducible representations can be found by using the formula $\sum_i (\dim V_i)^2 = \#G$, where we sum over all irreducible representations.

4. Schur's orthogonality relations give equations that the unknown representations have to satisfy.

Example 2.17. We will construct the character table of the dihedral group D_3. The conjugacy classes of D_3 are $(e) = \{e\}$, $(r) = \{r, r^2\}$, and $(s) = \{s, rs, r^2s\}$.

To find the one-dimensional representations, we will use that such a representation ρ is determined by its values on the generators r and s. It should satisfy $\rho(r)^3 = 1$, $\rho(s)^2 = 1$, and $\rho(r)\rho(s) = \rho(s)\rho(r)^{-1}$. From these equations it follows that $\rho(r) = 1$ and $\rho(s) \in \{\pm 1\}$, so there are two 1-dimensional representations.

Alternatively, we could have used Proposition 2.16 to find the one-dimensional representations. The abelianization of D_3 is $D_3/[D_3, D_3] \cong C_2$, so there are two one-dimensional representations. Since the trivial representation and the representation determined by $\rho(r) = 1$, $\rho(s) = -1$ are both well-defined representations, it follows that these are all one-dimensional representations.

Next we will find the number of irreducible representations and their dimensions. The group D_3 has three conjugacy classes, hence three irreducible representations. Two of these are 1-dimensional. Call the dimension of the last one x, then the formula $\sum_i (\dim V_i)^2 = \#G$ gives $1^2 + 1^2 + x^2 = 6$, so $x = 2$.

So far we know the characters χ_1, χ_2 of the 1-dimensional representations, and we know that $\chi_3(e) = \dim V_3 = 2$ for the remaining character χ_3. Thus the character table is of the form

	(e)	(r)	(s)
χ_1	1	1	1
χ_2	1	1	-1
χ_3	2	a	b

where the numbers a and b are still unknown. These can be found using Schur's orthogonality relations:

$$\langle \chi_1, \chi_3 \rangle = \frac{1}{\#D_3} \sum_{g \in D_3} \overline{\chi_1(g)} \chi_3(g) = \frac{1}{6}(2 + 2a + 3b) = 0$$

$$\langle \chi_2, \chi_3 \rangle = \frac{1}{6}(2 + 2a - 3b) = 0$$

Note that we have to sum over all elements of the group and not just the representatives from the conjugacy classes, so when reading the numbers from the character table, we have to multiply by the number of elements in each conjugacy class. Solving these equations for a and b gives $a = -1$, $b = 0$, which completes the character table.

3 The symmetric group

3.1 Idempotents in the group ring

The goal of this chapter is to describe all irreducible representations of the symmetric group S_n. In particular, we will define an explicit bijection between the conjugacy classes of S_n and its irreducible representations. This requires a connection between representations and ideals in a certain ring, so we will start by presenting this connection.

Let G be a group, and let $\mathbb{C}[G]$ be the vector space with basis G. Then each element of $\mathbb{C}[G]$ is a formal linear combination

$$\sum_{g \in G} \alpha_g g.$$

The vector space $\mathbb{C}[G]$ also forms a ring, whose multiplication is determined on basis vectors by $g \cdot h = gh$. Explicitly, the product of two elements $\sum_g \alpha_g g$ and $\sum_g \beta_g g$ is

$$\left(\sum_{g \in G} \alpha_g g\right)\left(\sum_{h \in G} \beta_h h\right) = \sum_{g \in G}\sum_{h \in G} \alpha_g \beta_h (gh) = \sum_{k \in G} \gamma_k k$$

where

$$\gamma_k = \sum_{g,h:\ gh=k} \alpha_g \beta_h = \sum_{g \in G} \alpha_g \beta_{g^{-1}k}.$$

The ring $\mathbb{C}[G]$ is called the *group ring*.

The group ring can be used to translate representation theoretic terms into ring theoretic terms. For example:

1. Subrepresentations of the regular representation $\mathbb{C}[G]$ are the same as left ideals in the group ring. This is because a subrepresentation $V \subseteq \mathbb{C}[G]$ is by definition closed under left multiplication by any $g \in G$, hence also by any element of the group ring $\mathbb{C}[G]$.

2. Every subrepresentation V of $\mathbb{C}[G]$ has a complement $V^\perp$ that is also a subrepresentation. In terms of ideals, this says that every left ideal $I \subseteq \mathbb{C}[G]$ has a complementary ideal I' with $I + I' = \mathbb{C}[G]$ and $I \cap I' = \{0\}$.

3. It follows that the unit $1 \in \mathbb{C}[G]$ can be written uniquely as $1 = e + e'$, where $e \in I$ and $e' \in I'$. To avoid trivial cases, we will assume from now on that $e, e' \neq 1$.

Lemma 3.1. *Let I be left ideal in the group ring $\mathbb{C}[G]$ and let I' be its complementary ideal. Write $1 \in \mathbb{C}[G]$ as $1 = e + e'$, where $e \in I$ and $e' \in I'$. Then:*

1. *e and e' are idempotent: $e^2 = e$ and $(e')^2 = e'$.*

2. *e and e' are disjoint: $ee' = e'e = 0$.*

3. *The ideals I and I' are generated by e and e', respectively: $I = (e)$ and $I' = (e')$. The notation (e) stands for $(e) = \mathbb{C}[G]e = \{xe \mid x \in \mathbb{C}[G]\}$.*

Proof. Every element in $\mathbb{C}[G]$ can be decomposed uniquely as $x+y$, where $x \in I$ and $y \in I'$. Because $1 = e + e'$ we have $e = e^2 + ee'$. Here $e^2 \in I$ and $ee' \in I'$, since I and I' are left ideals. At the same time we have $e = e + 0$ with $e \in I$ and $0 \in I'$. Because the decompositions are unique, it follows that $e^2 = e$ and $ee' = 0$.

The above argument is symmetric in e and e', so 1 and 2 have been proven.

If $x \in I$, then $x = x \cdot 1 = x(e+e') = xe + xe'$. But also $x = x + 0$, so $xe = x$ and $xe' = 0$. This shows that $I \subseteq (e)$. The reverse inclusion is obvious, and hence 3 holds. □

Corollary 3.2. *There is a one-to-one correspondence between left ideals in the group ring $\mathbb{C}[G]$ and idempotents $e \in \mathbb{C}[G]$. In fact, if e is the idempotent belonging to a left ideal I, then $I = (e)$ and*

$$x \in I \iff x = xe$$

for any $x \in \mathbb{C}[G]$.

When I, considered as a representation, is not irreducible, then we can write $I = J \oplus J'$ for certain left ideals J, J'. Let e be an idempotent for which $I = (e)$. Then e can be split as $e = f + f'$ for $f \in J$, $f' \in J'$. These satisfy $f = fe = f(f+f') = f^2 + ff'$, so $f^2 = f$ and $ff' = 0$. Thus any idempotent generating a reducible representation can be decomposed into "smaller" disjoint idempotents. This gives the following characterization of irreducible representations in terms of idempotents in the group ring.

Corollary 3.3. *The representation $I = (e)$ is irreducible if and only if e cannot be split into disjoint idempotents. This means that if $e = f + f'$ with $f^2 = f$, $(f')^2 = f'$, and $ff' = f'f = 0$, then either $f = 0$ or $f' = 0$.*

An idempotent element e with the property from the corollary is called a *primitive* idempotent.

Corollary 3.4. *In $\mathbb{C}[G]$, the unit 1 can be written as $1 = e_1 + \cdots + e_k$, where the e_i are pairwise disjoint primitive idempotents. Here "pairwise disjoint" means that $e_ie_j = 0$ whenever $i \neq j$.*

We have seen how to characterize subrepresentations of the regular representation in terms of idempotents in the group ring. We also know how to characterize the irreducible representations among these. Now we will continue with intertwiners.

Proposition 3.5. *Let $I = (e)$ and $I' = (e')$ be subrepresentations of the regular representation $\mathbb{C}[G]$. Intertwiners $\varphi : I \to I'$ correspond to $x_0 \in I'$ with the property that $ex_0e' = x_0$.*

Proof. Given an intertwiner φ, define $x_0 = \varphi(e)$. We will show that this x_0 satisfies the required property. Since $x_0 \in I'$, we have $x_0 e' = x_0$ by Corollary 3.2. Furthermore, φ is an intertwiner and e is idempotent, so $x_0 = \varphi(e) = \varphi(e^2) = e\varphi(e) = ex_0$. Combining these two facts gives $ex_0e' = x_0$.

Conversely, given x_0, define an intertwiner φ by $\varphi(y) = yx_0$. This is clearly an intertwiner, and from $ex_0 = x_0$ it follows that both constructions are inverses. □

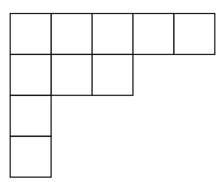
Corollary 3.6. *A representation $I = (e)$ is irreducible if and only if for each x_0 with $ex_0e' = x_0$ we have $x_0 = \lambda e$ for some scalar $\lambda \in \mathbb{C}$.*

Proof. The representation I is irreducible if and only if every intertwiner $I \to I'$ is an isomorphism onto its image. Such an isomorphism is a scalar multiplication by Schur's Lemma. □

Corollary 3.7. *Two irreducible representations $I = (e)$ and $I' = (e')$ are equivalent if and only if there exists $x_0 \neq 0$ such that $ex_0e' = x_0$.*

3.2 Young diagrams

In this section we will construct an explicit bijection between conjugacy classes and irreducible representations of S_n. Two permutations in S_n are conjugate if and only if they have the same cycle type. Cycle types are partitions of n as a sum $n = n_1 + \cdots + n_k$, where the numbers n_i represent the lengths of the cycles. In such a partition $n = n_1 + \cdots + n_k$, we may assume that $n_1 \geq n_2 \geq \cdots \geq n_k$. Then the partitions can be drawn as a diagram with k rows, where the i^{th} row consists of n_i squares. For example, the partition $10 = 5 + 3 + 1 + 1$ becomes the diagram



Such a diagram is called a *Young diagram.*

If we put the numbers $1, \ldots, n$ in the squares of a Young diagram, we call it a *Young tableau.* For example, we can make the above diagram into a Young tableau

5	1	6	4	3
2	7	8		
10				
9				

The group S_n acts on the collection of Young tableaus by permuting the numbers in the squares. That is, if $p \in S_n$ and T is a Young tableau, then $p \cdot T$ changes the number i in T into $p(i)$.

Call two tableaus T and T' *equivalent* if $T = p \cdot T'$ for some permutation p. This happens if and only if T and T' have the same underlying diagram. We know that there is a one-to-one correspondence between:

- Irreducible representations of S_n;
- Conjugacy classes of elements in S_n;
- Equivalence classes of Young tableaus;
- Young diagrams.

Furthermore, in the previous section we saw that there is an equivalence between irreducible representations of S_n and equivalence classes of primitive idempotents in the group ring $\mathbb{C}[S_n]$. We wish to construct an explicit bijection from equivalence classes of tableaus to equivalence classes of primitive idempotents, which will establish the goal of this chapter.

Construction of primitive idempotent. Given a tableau T, we define two subgroups of S_n:

$$\mathbf{P} = \mathbf{P}_T = \{p \in S_n \mid p \text{ fixes the rows of } T\}$$

$$\mathbf{Q} = \mathbf{Q}_T = \{q \in S_n \mid q \text{ fixes the columns of } T\}$$

Note that, if $p \in \mathbf{P}$, then p need not fix every element in every row; it only maps each element to a number in the same row. Observe that $\mathbf{P}_{rT} = r\mathbf{P}_T r^{-1}$ and $\mathbf{Q}_{rT} = r\mathbf{Q}_T r^{-1}$.

Using the subgroups $\mathbf{P}$ and $\mathbf{Q}$, define two elements of the group algebra $\mathbb{C}[S_n]$:

$$P = P_T = \sum_{p \in \mathbf{P}_T} p$$

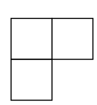
$$Q = Q_T = \sum_{q \in \mathbf{Q}_T} \operatorname{sgn}(q) q$$

Finally, define $e = e_T = P_T Q_T$.

Claim. The element e is idempotent up to scalar multiplication. This means that $e^2 = \lambda e$ for some $\lambda \in \mathbb{C}$, $\lambda \neq 0$. Then $\frac{e}{\lambda}$ is idempotent, and this gives the bijection from tableaus to primitive idempotents.

The proof of this claim is long and will be postponed to the next section. Here we will look at an example of the above construction.

Example 3.8. We wish to construct the irreducible representation of S_3 associated to the diagram



First we have to fill in numbers in the diagram to obtain a tableau. Equivalent tableaus will turn out to give equivalent representations, so it does not matter which Young tableau we choose. Let T be the tableau

1	2
3	

The subgroup $\mathbf{P}_T \subseteq S_3$ consists of those permutations that keep all entries in the same row. For the tableau T, this means that 3 stays at its place, while 1 and 2 may be swapped or not. Therefore

$$\mathbf{P}_T = \{\mathrm{id}, (1\ 2)\}.$$

Similarly,

$$\mathbf{Q}_T = \{\mathrm{id}, (1\ 3)\}.$$

From this we compute

$$P_T = \sum_{p \in \mathbf{P}_T} p = \mathrm{id} + (1\ 2) \in \mathbb{C}[S_n]$$

$$Q_T = \sum_{q \in \mathbf{Q}_T} \mathrm{sgn}(q) q = \mathrm{id} - (1\ 3) \in \mathbb{C}[S_n]$$

$$e_T = P_T Q_T = \mathrm{id} + (1\ 2) - (1\ 3) - (1\ 2)(1\ 3) = \mathrm{id} + (1\ 2) - (1\ 3) - (1\ 3\ 2)$$

The ideal $(e_T) = \mathbb{C}[S_3]$ is a representation of S_3. We can find an explicit description of this representation by finding a basis for (e_T). Write $e = e_T$ and define

$$f = (1\ 3)e = (1\ 3) + (1\ 2\ 3) - \mathrm{id} - (2\ 3) \in (e).$$

If we let any permutation in S_n act on e, then it always gives a linear combination of e and f. Furthermore e and f are linearly independent, so together they form a basis for (e). To describe the representation, it suffices to write down the action of $(1\ 2)$ and $(1\ 3)$ on the basis vectors, since S_3 is generated by these two transpositions. These actions are given by:

$$\begin{array}{ll} (1\ 2) \cdot e = e & \qquad (1\ 3) \cdot e = f \\ (1\ 2) \cdot f = -e - f & \qquad (1\ 3) \cdot f = e \end{array}$$

For example, the equation $(1\ 2) \cdot f = -e - f$ can be verified by the computation

$$(1\,2) \cdot f = (1\,2)((1\,3) + (1\,2\,3) - \mathrm{id} - (2\,3)) = (1\,3\,2) + (2\,3) - (1\,2) - (1\,2\,3) = -e - f$$

and similarly for the other equations.

We can check directly that this representation is irreducible using Corollary 2.13. First we rewrite the representation in terms of matrices, as a map $\rho : S_3 \to \mathrm{GL}_2(\mathbb{C})$:

$$\rho((1\ 2)) = \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix}, \quad \rho((1\ 3)) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Since ρ is a homomorphism and $(1\ 3)(1\ 2) = (1\ 2\ 3)$, it follows that

$$\rho((1\ 2\ 3)) = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$$

Hence the character of ρ is determined by

$$\chi_\rho(\mathrm{id}) = 2, \quad \chi_\rho((1\ 2)) = 0, \quad \chi_\rho((1\ 2\ 3)) = -1.$$

Thus ρ is irreducible because

$$\langle \chi_\rho, \chi_\rho \rangle = \frac{1}{\# S_3} \sum_{\sigma \in S_3} |\chi_\rho(\sigma)|^2 = \frac{1}{6}(2^2 + 3 \cdot 0 + 2 \cdot (-1)^2) = 1.$$

3.3 Young tableaus and primitive idempotents

In this section we will prove all claims made in the previous section. We will show that the element $e_T \in \mathbb{C}[S_n]$ assigned to a tableau T is idempotent up to scalar multiplication. Furthermore we will show that the resulting representation (e_T) is irreducible, or, equivalently, that e_T is a primitive idempotent. Finally we will show that the construction of e_T gives a bijection from equivalence classes of tableaus to equivalence classes of primitive idempotents. This means that T and T' are equivalent if and only if (e_T) and $(e_{T'})$ are.

We start with a lemma about decompositions of permutations with respect to any tableau T.

Lemma 3.9.

1. *An $r \in S_n$ can be written as $pq = r$ for some $p \in \mathbf{P}_T$, $q \in \mathbf{Q}_T$ in at most one way.*

2. *This can be done precisely when the following condition holds: for any two numbers i and j in T with $i \neq j$, if i and j are in the same row of T, then i and j are* not *in the same column of rT.*

Proof.

1. Suppose that $r = pq = p'q'$. Then $(p')^{-1}p = (q')^{-1}q$, and this is a permutation that fixes rows as well as columns. Hence it is the identity, so $p = p'$ and $q = q'$.

2. First suppose that $r = pq$ with $p \in \mathbf{P}_T$ and $q \in \mathbf{Q}_T$. Also let i and j be numbers in the same row of T. Then i and j are also in the same row of pT, hence in distinct columns. Since $pqp^{-1} \in \mathbf{Q}_{pT}$ and $r = pqp^{-1}p$, this shows that i and j are also not in the same column of rT.

   Conversely, suppose that $r \in S_n$ has the property that any i, j in the same *row* of T occur in distinct *columns* of rT, for $i \neq j$. Look at the numbers in the first column of rT. They occur in different rows of T. Let p_1 be any permutation in $\mathbf{P}_T$ moving these numbers to the far left. In this way, p_1T and rT have the same numbers in the first column, although they are possibly in a different order.

Now look at the numbers in the second column of rT. They occur in distinct rows of T, or equivalently in distinct rows of p_1T, but not in the first column of p_1T. Let $p_2 \in \mathbf{P}_{p_1T} = \mathbf{P}_T$ be any permutation moving them to the second column of p_1T while keeping the first column untouched. Now p_1p_2T has the same first two columns as rT.

Continue in this way. In the end we find $p_1, \dots, p_k$ so that $p_kp_{k-1}\cdots p_2p_1T$ and rT have the same numbers in each column. Let $p = p_k \cdots p_1 \in \mathbf{P}_T$, and let $q \in \mathbf{Q}_{pT}$ rearrange the columns in such a way that $qpT = rT$; then $qp = r$. Write $q = pq'p^{-1}$ for some $q' \in \mathbf{Q}_T$. Then $r = qp = pq'$. □

Order on Young tableaus. We will define a partial order on the collection of all Young diagrams. Let D and D' be Young diagrams. Then D is nothing but a sequence of numbers $m_1 \geq m_2 \geq \cdots \geq m_k$ with $m_1 + \cdots + m_k = n$, where m_i is the length of the i^{th} row. Similarly D' can be seen as a sequence of numbers $m'_1 \geq m'_2 \geq \cdots \geq m'_l$. Say that $D > D'$ if and only if $m_i > m'_i$ for the first i where $m_i \neq m'_i$. We will use the same notation for tableaus, and write $T > T'$ if they are ordered in this way when viewed as diagrams. Thus for any permutations $r, s \in S_n$, $T > T'$ if and only if $rT > sT$.

Lemma 3.10. *If T and T' are tableaus such that any two i, j in the same row of T occur in different columns of T', then $T \leq T'$.*

Proof. The numbers in the first row of T occur in different columns of T', so $m_1 \leq m'_1$. If $m_1 < m'_1$ then we are done. Otherwise $m_1 = m'_1$; in that case let $q' \in \mathbf{Q}_{T'} \subseteq S_n$ bring these numbers to the first row of $q'T'$. Then T and $q'T'$ have the same numbers in their first rows, and $T \leq T'$ if and only if $T \leq q'T'$. Now delete the first rows from T and $q'T'$ and repeat the argument. □

Lemma 3.11. *Let T and T' be tableaus, and suppose there are $i \neq j$ in the same row of T and the same column of T'. (This occurs e.g. if $T > T'$.) Then $Q_{T'}P_T = 0$.*

Proof. Write $\mathbf{Q}' = \mathbf{Q}_{T'}$ and $\mathbf{P} = \mathbf{P}_T$. Let τ be the permutation $(i\ j)$ in $\mathbf{P} \cap \mathbf{Q}'$. Then $P = \tau P = P\tau$ and $Q' = -\tau Q' = -Q'\tau$, so

$$Q'P = -(Q'\tau)P = -Q'(\tau P) = -Q'P,$$

whence $Q'P = 0$. □

Now let us return to the elements $e = e_T \in \mathbb{C}[S_n]$ assigned to tableaus T:

$$e = e_T = P_TQ_T = \sum_{p \in \mathbf{P}_T, q \in \mathbf{Q}_T} \operatorname{sgn}(q)pq$$

Observe the following facts about the element e:

1. The permutations $pq \in S_n$ occuring in the above sum are all different, by Lemma 3.9.
2. If $T > T'$, then $e'e = 0$, by Lemmas 3.10 and 3.11.
3. For a tableau T and $p_0 \in \mathbf{P}$, $q_0 \in \mathbf{Q}$,

$$p_0 e q_0 = \operatorname{sgn}(q_0) e.$$

We take up this last property, for a fixed tableau T and associated P, Q, and e, and show that it is characteristic of the element e, in the following sense.

Lemma 3.12. *Let $x \in \mathbb{C}[S_n]$ have the property that*

$$p_0 x q_0 = \operatorname{sgn}(q_0) x \textit{ for all } p_0 \in \mathbf{P},\ q_0 \in \mathbf{Q}.$$

Then, writing $x = \sum_s x_s s$, we have $x = x_{\mathrm{id}} e$.

Proof. Write

$$x = \sum_{s \in S_n} x_s s = \sum_{s \in \mathbf{PQ}} x_s s + \sum_{s \notin \mathbf{PQ}} x_s s = A + B,$$

and notice that the assumption holds for A and B separately:

$$p_0 A q_0 = \operatorname{sgn}(q_0) A, \quad p_0 B q_0 = \operatorname{sgn}(q_0) B,$$

for any $p_0 \in \mathbf{P}$, $q_0 \in \mathbf{Q}$. For A, this gives

$$\sum_{p,q} x_{pq} pq = \sum_{p,q} \operatorname{sgn}(q_0) x_{pq} p_0 pq q_0,$$

so $x_{p_0 pq q_0} = \operatorname{sgn}(q_0) x_{pq}$ for any p_0, p, q_0, q. In particular, $x_{\mathrm{id}} = \operatorname{sgn}(q) x_{pq}$, whence

$$A = \sum_{p,q} x_{pq} pq = \sum_{p,q} \operatorname{sgn}(q) x_{\mathrm{id}} pq = x_{\mathrm{id}} e.$$

Next, we show that $B = 0$. If $s \notin \mathbf{PQ}$ then there are $i \neq j$ in the same row of T and the same column of sT, by Lemma 3.9. Let $\tau = (i\ j) \in \mathbf{P}_T \cap \mathbf{Q}_{sT}$, so $\tau = s\sigma s^{-1}$ for some $\sigma \in \mathbf{Q}_T$. Then $x = -\tau x \sigma$ by assumption, so $x_s = -x_{\tau s \sigma}$ for this particular s. But $\tau = s\sigma s^{-1}$, so $\tau s \sigma = s$, hence $x_s = 0$. This holds for any $s \notin \mathbf{PQ}$, so $B = 0$. $\square$

We are now ready to prove the claim made before.

Proposition 3.13. *For any tableau T, the associated element $e = e_T$ satisfies $e^2 = \lambda e$ for some non-zero $\lambda \in \mathbb{C}$.*

Proof. The element e^2 clearly satisfies the hypothesis of Lemma 3.12, so $e^2 = \lambda e$, where λ is the coordinate of $e^2 \in \mathbb{C}[S_n]$ at the unit. To calculate λ, consider the right multiplication map

$$R_e : \mathbb{C}[S_n] \to \mathbb{C}[S_n], \quad x \mapsto xe.$$

Then R_e projects $\mathbb{C}[S_n]$ onto the ideal $I = (e)$, on which the map R_e is right multiplication by λ. Hence the trace of the map R_e is

$$\operatorname{tr}(R_e) = \lambda \dim(I) = \lambda \dim((e)).$$

On the other hand, we can easily compute the trace of R_e, since for a basis vector $s \in \mathbb{C}[S_n]$,

$$R_e(s) = se = \sum \operatorname{sgn}(q) spq,$$

and $spq = s$ if and only if $p = q = \mathrm{id}$, so $\operatorname{tr}(R_e) = \#S_n = n!$. Therefore

$$\lambda = \frac{n!}{\dim((e))}. \qquad \square$$

So now we have found our idempotent e/λ assigned to a tableau T. It remains to show that:

1. $(e) = (e/\lambda)$ is irreducible.
2. For tableaus T and T', the representations (e) and (e') are equivalent if and only if T and T' are.

Proof.

1. We know that any intertwiner $\varphi : (e) \to (e)$ must be of the form $\varphi(y) = yx_0$ for some x_0 with $(e/\lambda)x_0(e/\lambda) = x_0$, or in other words $ex_0e = \lambda^2 x_0$. But such an x_0 satisfies the hypothesis of Lemma 3.12, so x_0 is a multiple of e, and hence φ is an isomorphism. This proves that (e) is irreducible.

2. Suppose that T and T' are equivalent. Then $T = sT'$ for some permutation s, so $e_T = se_{T'}s^{-1}$. But then (e_T) and $(se_{T'}s^{-1})$ are obviously equivalent.

   For the other direction, assume that T and T' have different shapes, say that $T > T'$. Let $\varphi : (e') \to (e)$ be an intertwiner. Then φ is of the form $\varphi(y) = yx_0$ for some $x_0 \in \mathbb{C}[S_n]$ with $e'x_0e = \lambda\lambda' x_0$. We claim that $x_0 = 0$. Write $x_0 = \sum_s x_s s$. Then $e'x_0e = \sum_s x_s e'se$. We already know from Lemma 3.11 that $e'e = 0$ because $T > T'$. But then also $T > s^{-1}T'$ for each s, so $(s^{-1}e's)e = 0$, and hence $e'se = 0$. Thus $e'x_0e = 0$, and hence $x_0 = 0$ because λ and λ' are non-zero. $\square$

4 Restriction and induction

4.1 Modules over a ring

Modules over a ring are a generalization of representations of a group. Some constructions on representations are most naturally formulated in the language of modules, so we will give a brief introduction to modules here.

Definition 4.1. Let R be a ring. A *module* over R, sometimes called an R-*module*, is an abelian group M (write $+$ for the group operation) together with a linear action of the ring R on M. This means that there is a map

$$R \times M \to M, \quad (r, m) \mapsto r \cdot m$$

with the following properties:

- $r \cdot (m + n) = r \cdot m + r \cdot n$
- $(r + s) \cdot m = r \cdot m + s \cdot m$
- $(rs) \cdot m = r \cdot (s \cdot m)$
- $1m = m$

A *homomorphism of R-modules* is a group homomorphism $\varphi : M \to N$ that is compatible with the action of R in the sense that $\varphi(r \cdot m) = r \cdot \varphi(m)$.

Examples 4.2.

1. If R is a field, then a module over R is exactly a vector space over R. Homomorphisms of modules over a field are linear maps between the vector spaces. Thus modules generalize vector spaces to the setting of rings.
2. Let G be a group. A module over the group ring $\mathbb{C}[G]$ is the same as a representation of G. Indeed, a map $G \times V \to V$ can be extended linearly to a map $\mathbb{C}[G] \times V \to V$ in a unique way, and conversely, a map $\mathbb{C}[G] \times V \to V$ can be restricted to obtain a representation of G. The homomorphisms of $\mathbb{C}[G]$-modules are precisely the intertwiners.
3. If I is a left ideal in the ring R, then it is a module over R, because it is closed under addition and left multiplication with elements in R.

Modules over different rings can be related by means of ring homomorphisms. Let $f : S \to R$ be a ring homomorphism. Then any R-module M can be transformed into an S-module denoted $f^*(M)$. The underlying abelian group is again M, and the action of S is given by $s \cdot m = f(s) \cdot m$. This construction is called *restriction* along f. A common special case is when S is a subring of R, and f is the inclusion homomorphism. In this case, $f^*(M)$ is the actual restriction of the R-action to an S-action.

In the other direction, given a ring homomorphism $f : S \to R$ and an S-module N, it is possible to construct an R-module denoted $R \otimes_S N$ or $f_!(N)$. The construction is similar to the tensor product of representations. Elements of $R \otimes_S N$ are linear combinations of the form

$$\sum_{i=1}^{n} k_i r_i \otimes x_i,$$

where $k_i \in \mathbb{Z}$, $r_i \in R$, and $x_i \in N$. The tensor symbol $\otimes$ is again a formal symbol to indicate the generators of the tensor product. The elements of $R \otimes_S N$ are subject to the following relations:

$$(r + r') \otimes x = r \otimes x + r' \otimes x$$

$$r \otimes (x + x') = r \otimes x + r \otimes x'$$

$$(rf(s)) \otimes x = r \otimes (s \cdot x)$$

In words, the tensor operation $\otimes$ is linear in both variables separately, and it is compatible with the action of S on N.

Because $R \otimes_S N$ was defined in terms of linear combinations with coefficients in $\mathbb{Z}$, it is an abelian group. We make it into an R-module by defining $t \cdot (r \otimes x) = (tr) \otimes x$, for $r, t \in R$ and $x \in N$. The R-module $R \otimes_S N$ is characterized by the following analogue of Propositions A.10 and 1.5.

Proposition 4.3. *There is a one-to-one correspondence between:*

- *Homomorphisms of S-modules $N \to f^*(M)$;*
- *Homomorphisms of R-modules $R \otimes_S N \to M$.*

Proof. If $\alpha : N \to f^*(M)$ is a homomorphism of S-modules, define a homomorphism $\beta : R \otimes_S N \to M$ on generators by $\beta(r \otimes x) = r \cdot \alpha(n)$. The map β is well-defined if α respects the S-action, since

$$\beta(rf(s) \otimes x) = rf(s) \cdot \alpha(x) = r\alpha(s \cdot x) = \beta(r \otimes s \cdot x).$$

It is easy to check that β preserves the R-action.

For the other direction, given $\beta : R \otimes_S N \to M$, define $\alpha : N \to f^*(M)$ by $\alpha(x) = \beta(1 \otimes x)$. Then α is a homomorphism of S-modules, and the construction are mutually inverse. $\square$

4.2 Induced representations

If $f : H \to G$ is a group homomorphism and V is a representation of G, then we can obtain a representation of H by restriction along f. The underlying vector space of the restriction is again V, and the action of H on V is $h \cdot v = f(h)v$. The resulting representation is written as $f^*(V)$, or as $\operatorname{Res}_H^G V$, or simply $\operatorname{Res} V$ if the homomorphism f is clear from context. This is in fact a special case of the construction in the previous section, since representations of G are the same as $\mathbb{C}[G]$-modules.

In the same way, we can specialize the construction of the tensor product of modules to the case of representations. This provides a method to construct G-representations from H-representations, given a group homomorphism $f : H \to G$. The resulting G-representation is called the *induced representation*, and it is given by

$$f_!(W) = \mathbb{C}[G] \otimes_{\mathbb{C}[H]} W$$

for a H-representation W. Since $f_!(W)$ is a $\mathbb{C}[G]$-module, it is indeed a representation of G. The induced representation $f_!(W)$ is often written as $\operatorname{Ind}_H^G W$ or $\operatorname{Ind} W$. It is especially interesting in the case where H is a subgroup of G, and f is the inclusion from H into G, since then it enables us to produce representations of a large group from representations of a smaller group, which are often easier to find.

The fundamental mapping property from Proposition 4.3 immediately translates into a bijective correspondence between intertwiners of G-representations $f_!(W) \to V$ and intertwiners of H-representations $W \to f^*(V)$. This correspondence can also be written as

$$\operatorname{Hom}(\operatorname{Ind} W, V)^H \cong \operatorname{Hom}(W, \operatorname{Res} V)^G.$$

Here we used that the invariant vectors in a Hom-representation are exactly the intertwiners. By the dimension formula, we obtain the following formula for the characters.

Proposition 4.4 (Frobenius reciprocity)**.** *Let $f : H \to G$ be a group homomorphism, V a representation of G, and W a representation of H. Then*

$$\langle \chi_{\operatorname{Ind} W}, \chi_V \rangle_G = \langle \chi_W, \chi_{\operatorname{Res} V} \rangle_H.$$

The subscripts G and H are used here to emphasize that we use the inner product on $\mathcal{C\ell}(G)$ on the left, and the one on $\mathcal{C\ell}(H)$ on the right.

Example 4.5. Let $f : C_3 \to D_3$ be the inclusion homomorphism, and define a one-dimensional representation (W, ρ) of C_3 by $\rho(r) = \omega = e^{2\pi i/3}$. This gives an induced representation $V = \operatorname{Ind} W$ of D_3. We wish to find this induced representation. For this we can compute its decomposition into irreducible representations of D_3. Let V_1, V_2, V_3 be the irreducible representations of D_3, and recall that its character table is:

	(e)	(r)	(s)
χ_1	1	1	1
χ_2	1	1	-1
χ_3	2	-1	0

The representation V_i occurs $\langle \chi_V, \chi_i \rangle_{D_3}$ times in V, which equals $\langle \chi_W, \chi_{\operatorname{Res} V_i} \rangle_{C_3}$ by Frobenius reciprocity. For $i = 1, 2$ this is 0 by orthogonality. For $i = 3$ we compute this number using the character table:

$$\langle \chi_W, \chi_{\operatorname{Res} V_3} \rangle = \frac{1}{3}\left(2 - \omega - \omega^2\right) = 1.$$

Thus the induced representation V is V_3, the standard representation of D_3.

Frobenius reciprocity is also useful to establish basic properties of induced representations. As an example application, we will show that induction is transitive.

Proposition 4.6. *If $K \subseteq H \subseteq G$, then $\operatorname{Ind}_K^G W \cong \operatorname{Ind}_H^G(\operatorname{Ind}_K^H W)$.*

Proof. We will show that each irreducible representation of G occurs the same number of times in $\operatorname{Ind}_K^G W$ and in $\operatorname{Ind}_H^G(\operatorname{Ind}_K^H W)$; then it will follow that these two representations are equivalent. Let V be an irreducible representation of G. It occurs $\langle \chi_V, \chi_{\operatorname{Ind}_H^G(\operatorname{Ind}_K^H W)} \rangle$ times in $\operatorname{Ind}_H^G(\operatorname{Ind}_K^H W)$, which is equal to $\langle \chi_{\operatorname{Res}_K^H(\operatorname{Res}_H^G V)}, \chi_W \rangle$, as can be seen by applying Frobenius reciprocity twice. Since restriction is transitive, this is the same as $\langle \chi_{\operatorname{Res}_K^G V}, \chi_W \rangle$, which is in turn equal to $\langle \chi_V, \chi_{\operatorname{Ind}_K^G W} \rangle$ by Frobenius reciprocity. Hence V occurs the same number of times in both representations. □

4.3 Characters of induced representations

As we noticed before, induced representations are especially interesting to obtain representations of a large group from representations from a subgroup. Let H be a subgroup of G, and let $i : H \to G$ be the inclusion homomorphism. In this section we will take a closer look at the induced representations obtained in this setting. In particular, we will give a new description of these induced representations, and compute their characters.

It will be convenient to choose a set $R \subseteq G$ of representatives for the set G/H of left cosets. More precisely, $G/H = \{\xi H \mid \xi \in R\}$, and every coset gH occurs as ξH for exactly one $\xi \in R$. Thus, for every $g \in G$ there is exactly one $\xi \in R$ such that $g^{-1}\xi \in H$. Of course, we may represent the coset H itself by the unit $e \in G$, and take e to be an element of R. So we will assume that $e \in R$ from now on.

For a representation W of H, any element of the G-representation $\operatorname{Ind} W = i_!(W)$ is a linear combination of elements of the form $g \otimes w$, where $g \in G$ is identified with the corresponding basis vector of $\mathbb{C}[G]$. Since i is an inclusion, the relations for the tensor product simplify to

$$g \otimes (w + w') = g \otimes w + g \otimes w',$$

$$gh \otimes w = g \otimes (h \cdot w)$$

for $g \in G$, $h \in H$, and $w, w' \in W$.

We can rewrite a generating tensor $g \otimes w$ in terms of the representatives in R. Let ξ be the unique element in R for which $gH = \xi H$; then

$$g \otimes w = g(g^{-1}\xi)(g^{-1}\xi)^{-1} \otimes w = gg^{-1}\xi \otimes (g^{-1}\xi)^{-1}w = \xi \otimes (g^{-1}\xi)^{-1}w.$$

Therefore any element of $\operatorname{Ind} W = \mathbb{C}[G] \otimes_{\mathbb{C}[H]} W$ is a linear combination of elements of the form $\xi \otimes w'$ for $\xi \in R$, $w' \in W$.

This can be used to obtain an alternative description of the induced representation $\mathbb{C}[G] \otimes_{\mathbb{C}[H]} W$. Consider the vector space $\bigoplus_{\xi \in R} W$. Elements of this vector space are sequences of vectors in W, indexed by representatives in R. Write the sequence $(0, \ldots, 0, w, 0, \ldots, 0)$, with the only non-zero element at the position indexed by $\xi \in R$, as (ξ, w). In this way all elements of $\bigoplus_{\xi \in R} W$ are linear combinations of elements of the form (ξ, w).

We can turn $\bigoplus_{\xi \in R} W$ into a representation of G in the following way. The group G acts on the set of left cosets G/H, via $g \cdot (g'H) = (gg')H$. Hence G also acts on the set R. To distinguish this action from the product in G, we denote it by $\star$. Explicitly, $g \star \xi$ is the unique ξ' for which $\xi' H = g\xi H$. In other words, the equation

$$(g \star \xi)^{-1} g\xi \in H$$

determines $g \star \xi \in R$ completely. Define the action of G on $\bigoplus_{\xi \in R} W$ as follows:

$$g \cdot (\xi, w) = (g \star \xi, (g \star \xi)^{-1} g\xi w)$$

Proposition 4.7. *The representation $\bigoplus_{\xi \in R} W$ is equivalent to the induced representation $\mathbb{C}[G] \otimes_{\mathbb{C}[H]} W$.*

Proof. Define a map

$$\sigma : \bigoplus_{\xi \in R} W \to \mathbb{C}[G] \otimes_{\mathbb{C}[H]} W, \quad (\xi, w) \mapsto \xi \otimes w.$$

Then σ is an intertwiner of G-representations, because

$$\sigma(g \cdot (\xi, w)) = \sigma(g \star \xi, (g \star \xi)^{-1} g\xi w) = (g \star \xi) \otimes (g \star \xi)^{-1} g\xi w = g\xi \otimes w = g \cdot \sigma(\xi, w).$$

We will define an inverse τ for σ. For $g \otimes w \in \mathbb{C}[G] \otimes_{\mathbb{C}[H]} W$, let $\tau(g \otimes w)$ be $(\xi, (g^{-1}\xi)^{-1}w)$, where ξ is the unique element of R with $gH = \xi H$. To check that τ is well-defined, note that gh has the same representative in R as g, since $ghH = gH$. Call this representative ξ, then

$$\tau(gh \otimes w) = (\xi, ((gh)^{-1}\xi)^{-1}w) = (\xi, \xi^{-1}ghw) = (\xi, (g^{-1}\xi)^{-1}hw) = \tau(g \otimes hw).$$

Furthermore, τ is the inverse of σ because of the following two calculations.

$$\tau\sigma(\xi, w) = \tau(\xi \otimes w) = (\xi, (\xi^{-1}\xi)^{-1}w) = (\xi, w)$$

$$\sigma\tau(g \otimes w) = \sigma(\xi, (g^{-1}\xi)^{-1}w) = \xi \otimes (g^{-1}\xi)^{-1}w = \xi(g^{-1}\xi)^{-1} \otimes w = g \otimes w \quad \square$$

The above description can be used to compute the character of an induced representation.

Proposition 4.8. *If $H \subseteq G$, then the character of the induced representation* $\operatorname{Ind}_H^G W$ *is given by*

$$\chi_{\operatorname{Ind} W}(g) = \sum_{\xi \in R:\ g \star \xi = \xi} \chi_W(\xi^{-1} g \xi).$$

Proof. If $e_1, \ldots, e_n$ is a basis for W, then all elements of the form (ξ, e_i) with $\xi \in R$ form a basis for $\operatorname{Ind} W \cong \bigoplus_{\xi \in R} W$. Taking the trace of $\rho_{\operatorname{Ind} W}(g)$ in this basis involves only those $\xi \in R$ for which $g \star \xi = \xi$, because of the form of the action of G on $\bigoplus_\xi W$. Therefore

$$\chi_{\operatorname{Ind} W}(g) = \sum_{\xi \in R:\ g \star \xi = \xi} \chi_W((g \star \xi)^{-1} g \xi) = \sum_{\xi \in R:\ g \star \xi = \xi} \chi_W(\xi^{-1} g \xi). \qquad \square$$

Remark. The expression $\chi_W(\xi^{-1} g \xi)$ in the character formula makes sense because $\xi^{-1} g \xi \in H$. Be careful, even though χ_W is a class function, we cannot conclude something like "$\chi_W(\xi^{-1} g \xi) = \chi_W(g)$" since this involves elements outside the group H, hence it does not make sense.

4.4 Mackey's irreducibility criterion

The formula for the character of an induced representation provides a tool to investigate when this induced representation is irreducible. As before, let H be a subgroup of G, let W be a representation of H, and let R be a set of representatives for the left cosets in G/H.

Character theory provides an easy criterion for irreducibility. The induced representation $\operatorname{Ind} W$ is irreducible if and only if

$$\langle \chi_{\operatorname{Ind} W}, \chi_{\operatorname{Ind} W} \rangle_G = 1.$$

By Frobenius reciprocity, this is equivalent to

$$\langle \chi_W, \chi_{\operatorname{Res} \operatorname{Ind} W} \rangle_H = 1.$$

According to Proposition 4.8, the character of $\operatorname{Res} \operatorname{Ind} W$ is given by

$$\chi_{\operatorname{Res} \operatorname{Ind} W}(h) = \chi_{\operatorname{Ind} W}(h) = \sum_{\xi \in R:\ h \star \xi = \xi} \chi_W(\xi^{-1} h \xi).$$

Use this to expand the inner product:

$$\langle \chi_W, \chi_{\operatorname{Res} \operatorname{Ind} W} \rangle = \frac{1}{\# H} \sum_{h \in H} \sum_{\xi:\ h \star \xi = \xi} \overline{\chi_W(h)} \chi_W(\xi^{-1} h \xi)$$

The condition $h \star \xi = \xi$ holds if and only if $\xi^{-1}h\xi \in H$, which is in turn equivalent to $h \in \xi H \xi^{-1}$. Thus we can rewrite the sum to obtain

$$\begin{aligned}\langle \chi_W, \chi_{\mathrm{Res}\,\mathrm{Ind}\,W} \rangle &= \frac{1}{\#H} \sum_{\xi \in R} \sum_{h \in H \cap \xi H \xi^{-1}} \overline{\chi_W(h)} \chi_W(\xi^{-1}h\xi) \\ &= \langle \chi_W, \chi_W \rangle + \sum_{\xi \neq e} \frac{1}{\#H} \sum_{h \in H \cap \xi H \xi^{-1}} \overline{\chi_W(h)} \chi_W(\xi^{-1}h\xi).\end{aligned}$$

The first term in this sum is always a positive integer. We will now analyze the second term. For any $\xi \in R$ with $\xi \neq e$, $H \cap \xi H \xi^{-1}$ is a subgroup of H. We can obtain representations of this group by restricting representations of H. There are two interesting homomorphisms from $H \cap \xi H \xi^{-1}$ into H, along which we can restrict. The first one is the inclusion

$$i_\xi : H \cap \xi H \xi^{-1} \to H, \quad i_\xi(h) = h,$$

and the second one is the conjugation

$$c_\xi : H \cap \xi H \xi^{-1} \to H, \quad c_\xi(h) = \xi^{-1}h\xi.$$

The representation W can be restricted along both homomorphisms to obtain representations $i_\xi^*(W)$ and $c_\xi^*(W)$ of $H \cap \xi H \xi^{-1}$. The inner product of these two representations is

$$\langle i_\xi^*(W), c_\xi^*(W) \rangle = \frac{1}{\#(H \cap \xi H \xi^{-1})} \sum_{h \in H \cap \xi H \xi^{-1}} \overline{\chi_W(h)} \chi_W(\xi^{-1}h\xi).$$

Combining this with the earlier computation yields

$$\langle \chi_{\mathrm{Ind}\,W}, \chi_{\mathrm{Ind}\,W} \rangle = \langle \chi_W, \chi_W \rangle + \sum_{\xi \neq e} \frac{\#(H \cap \xi H \xi^{-1})}{\#H} \langle \chi_{i_\xi^*(W)}, \chi_{c_\xi^*(W)} \rangle.$$

We want to know when this expression is equal to 1. Since $\langle \chi_W, \chi_W \rangle \geq 1$ and all numbers in the sum are non-negative, this happens if and only if $\langle \chi_{i_\xi^*(W)}, \chi_{c_\xi^*(W)} \rangle = 0$ for each $\xi \neq e$. We capture this property in a definition.

Definition 4.9. Two representations V and V' of G are called *disjoint* if no irreducible representation of G occurs in both V and V'.

In other words, the representations V and V' are disjoint if and only if all irreducible representations occuring in the decomposition of V are different from all irreducible representations occuring in the decomposition of V'. By Schur's orthogonality relations, this is equivalent to $\langle \chi_V, \chi'_V \rangle = 0$. Thus we have proven the following criterion for irreducibility of induced representations.

Theorem 4.10 (Mackey)**.** *Let H be a subgroup of G, and let W be a representation of H. Then the induced representation $\mathrm{Ind}_H^G W$ is irreducible if and only if W itself is irreducible, and for each left coset $\xi H \in G/H$ with $\xi \notin H$, the representations $i_\xi^*(W)$ and $c_\xi^*(W)$ of $H \cap \xi H \xi^{-1}$ are disjoint.*

Example 4.11. Let (W, ρ) be the standard representation of C_n, i.e. $W = \mathbb{C}$ and $\rho(r) = e^{2\pi i/n}$. It induces a representation $\operatorname{Ind} W$ of D_n, via the inclusion $C_n \subseteq D_n$. We wish to know for which n this representation is irreducible. The representation W itself is irreducible since it is one-dimensional, so the first condition of Mackey's criterion is fulfilled. For the second condition, notice that the only non-trivial left coset of C_n in D_n is sC_n, where s is the reflection in D_n. Since C_n is a normal subgroup of D_n, the representations $i_s^*(W)$ and $c_s^*(W)$ are representations of $C_n \cap sC_ns^{-1} = C_n \cap C_n = C_n$. The restriction $i_s^*(W)$ is simply W. The restriction $c_s^*(W)$ has underlying vector space $\mathbb{C}$, and the corresponding homomorphism $\rho_s : C_n \to \mathbb{C}$ is determined by $\rho_s(r) = \rho(s^{-1}rs) = \rho(r^{-1}) = e^{-2\pi i/n}$. Hence $\operatorname{Ind} W$ is irreducible if and only if $e^{2\pi i/n} \neq e^{-2\pi i/n}$, which holds if and only if $n \neq 1, 2$.

5 Category theory

5.1 Categories

Algebra provides insight into mathematical structures by viewing these in an abstract way. That is, we are not interested in what the elements of a structure look like, but in the relations between those elements. It is possible to achieve an even higher level of abstraction by not considering single algebraic structures, but focusing on the connections or homomorphisms between structures. This is the main idea behind *category theory.* Informally, a category is a collection of objects, regarded as mathematical structures, together with morphisms connecting the objects. Many aspects of algebra can be phrased in terms of categories. In this way, category theory gives a language that makes it easier to see similarities and differences between various structures. Here we shall introduce several basic notions of this theory. In particular, we will concentrate on examples of categorical concepts that are useful in representation theory.

Definition 5.1. A *category* $\mathbf{C}$ consists of:

- A class of *objects* $\mathrm{Ob}(\mathbf{C})$;
- For each two objects X and Y there is a set $\mathrm{Hom}(X,Y)$, whose objects are called *morphisms* from X to Y;
- For each $X,Y,Z \in \mathrm{Ob}(\mathbf{C})$ there is a *composition operation* $\circ : \mathrm{Hom}(X,Y) \times \mathrm{Hom}(Y,Z) \to \mathrm{Hom}(X,Z)$, written as $(f,g) \mapsto g \circ f$;
- For each object X there is an *identity morphism* $\mathrm{id}_X \in \mathrm{Hom}(X,X)$.

These satisfy the following requirements:

- Composition is associative: for $f \in \mathrm{Hom}(W,X)$, $g \in \mathrm{Hom}(X,Y)$, and $h \in \mathrm{Hom}(Y,Z)$, we have $(h \circ g) \circ f = h \circ (g \circ f)$.
- The identity morphism behaves as a unit for composition: for any $f \in \mathrm{Hom}(X,Y)$ we have $\mathrm{id}_Y \circ f = f = f \circ \mathrm{id}_X$.

We usually write $X \in \mathbf{C}$ instead of $X \in \mathrm{Ob}(\mathbf{C})$, and $f : X \to Y$ instead of $f \in \mathrm{Hom}(X,Y)$.

Examples 5.2.

1. The category **Sets** has sets as objects, and a morphism from a set X to Y is a function $f : X \to Y$. The composition operation is simply composition of functions. Checking that this operation is associative and that the identity function is a unit is easy.
2. All well-known algebraic structures form a category with their corresponding homomorphisms. For example, there is a category **Grp** of groups together with group homomorphisms, a category **Ab** of abelian groups and group homomorphisms, and a category **Ring** of rings with ring homomorphisms. If we fix a field k, then we can form the category $\mathbf{Vect}_k$, with vector spaces over k as objects and linear maps as morphisms.

3. Let G be a group. Representations of G are the objects in a category $\mathbf{Rep}_G$. The morphisms in this category are intertwiners. As before, we only consider complex representations.

4. In all of the above examples, morphisms are certain functions between the objects, and the composition operation is actual function composition. This is not the case in all categories. For instance, we can view a group G as a category in the following way. The category has one object, denoted by $*$. Morphisms from $*$ to $*$ are group elements $g \in G$, and composition is given by group multiplication in G. The group axioms guarantee that composition is associative and has a unit.

Categories themselves also form a mathematical structure, so we may ask what the associated morphisms are. These are called functors.

Definition 5.3. Let $\mathbf{C}$ and $\mathbf{D}$ be categories. A *functor* from $\mathbf{C}$ to $\mathbf{D}$ is a map F that assigns to each object X of $\mathbf{C}$ an object $F(X)$ of $\mathbf{D}$, and to each morphism $f : X \to Y$ in $\mathbf{C}$ a morphism $F(f) : F(X) \to F(Y)$. The assignment should satisfy $F(\mathrm{id}_X) = \mathrm{id}_{F(X)}$ and $F(g \circ f) = F(g) \circ F(f)$.

Functors compose, and the identity map Id on a category is always a functor. Therefore there exists a category $\mathbf{Cat}$ with categories as objects and functors as morphisms. Readers who are worried about set-theoretic size issues occuring when considering the category of categories may restrict themselves to categories whose class of objects forms a set.

Examples 5.4.

1. Every group has an underlying set, and every group homomorphism is in particular a function between the underlying sets. Hence there is a functor U from $\mathbf{Grp}$ to $\mathbf{Sets}$ sending a group to its underlying set, and satisfying $U(f) = f$ on morphisms. This functor is called *forgetful*, since it does nothing but forgetting the group structure.

   There are many more examples of forgetful functors: the functor $\mathbf{Ab} \to \mathbf{Grp}$ forgetting that a group is abelian; the functor $\mathbf{Vect}_k \to \mathbf{Ab}$ forgetting the scalar multiplication of a vector space, but retaining the addition; and the functor $\mathbf{Rep}_G \to \mathbf{Vect}_{\mathbb{C}}$ mapping a representation (V, ρ) to V.

2. We will define a functor $(-)^{\mathrm{ab}} : \mathbf{Grp} \to \mathbf{Ab}$, called 'abelianization'. On an object G in $\mathbf{Grp}$, define $G^{\mathrm{ab}} = G/[G,G]$, where $[G,G]$ is the commutator subgroup of G. A homomorphism $f : G \to H$ gives rise to a map $f^{\mathrm{ab}} : G/[G,G] \to H/[H,H]$, $f^{\mathrm{ab}}(g[G,G]) = f(g)[H,H]$. This map is well-defined since a group homomorphism maps commutators to commutators. The assignment $(-)^{\mathrm{ab}}$ clearly fulfills the conditions for a functor.

3. Let U be a vector space. Tensoring with U yields a functor $U \otimes (-) : \mathbf{Vect}_{\mathbb{C}} \to \mathbf{Vect}_{\mathbb{C}}$. If $f : V \to W$ is a morphism in $\mathbf{Vect}_{\mathbb{C}}$, then $U \otimes f : U \otimes V \to U \otimes W$ is defined by $(U \otimes f)(u \otimes v) = u \otimes f(v)$. This is easily

checked to be a functor. In the special case where U is a representation of a group G, the tensor product $U \otimes V$ is again a representation of G, with action $g \cdot (u \otimes v) = (g \cdot u) \otimes v$. Thus tensoring becomes a functor $U \otimes (-) : \mathbf{Vect}_{\mathbb{C}} \to \mathbf{Rep}_G$.

4. Again, let U be a vector space. The functor $\mathrm{Hom}(U, -) : \mathbf{Vect}_{\mathbb{C}} \to \mathbf{Vect}_{\mathbb{C}}$ maps an object V to the space of linear maps from U to V. For a linear map $f : V \to W$, define $\mathrm{Hom}(U, f) : \mathrm{Hom}(U, V) \to \mathrm{Hom}(U, W)$ by $\mathrm{Hom}(U, f)(\varphi) = f \circ \varphi$, as indicated in the following diagram:

$$\begin{tikzcd} U \arrow[r, "\varphi"] \arrow[dr, "\mathrm{Hom}(U,f)(\varphi)"'] & V \arrow[d, "f"] \\ & W \end{tikzcd}$$

   Just like in the previous example, this functor can also be seen as a functor $\mathbf{Vect}_{\mathbb{C}} \to \mathbf{Rep}_G$, provided U is a representation of G. The action of G on $\mathrm{Hom}(U, V)$ is given by $(g \cdot \varphi)(u) = \varphi(g^{-1} \cdot u)$. This is a special case of the action of G on a space of linear maps, where V is the trivial representation.

Some functors reverse the direction of the morphisms. More precisely, a *contravariant functor* is an assignment F that sends objects X to objects $F(X)$, and morphisms $f : X \to Y$ to morphisms $F(f) : F(Y) \to F(X)$, in such a way that $F(\mathrm{id}_X) = \mathrm{id}_{F(X)}$ and $F(g \circ f) = F(f) \circ F(g)$. An "ordinary" functor is sometimes called a *covariant* functor to emphasize the distinction with contravariant ones.

Example 5.5. Restriction of a representation along a group homomorphism can be described in the language of categories. There exists a contravariant functor $\mathbf{Rep} : \mathbf{Grp} \to \mathbf{Cat}$ that assigns to a group G the category $\mathbf{Rep}_G$ of complex representations. For a homomorphism $f : H \to G$ we have to define a functor $f^* : \mathbf{Rep}_G \to \mathbf{Rep}_H$. Keep in mind that H and G are swapped, since we want the functor $\mathbf{Rep}$ to be contravariant, but that f^* is covariant, since morphisms in $\mathbf{Cat}$ are covariant functors. The functor f^* is defined on objects of $\mathbf{Rep}_G$ via $f^*(V, \rho) = (V, \rho \circ f)$. Furthermore, f^* maps an intertwiner $\varphi : (V, \rho) \to (W, \sigma)$ to itself, considered as an intertwiner $\varphi : (V, \rho \circ f) \to (W, \sigma \circ f)$.

We have seen that morphisms between categories are functors. It is possible to take the abstraction to an even higher level by defining morphisms between functors.

Definition 5.6. Suppose that F and G are both functors from $\mathbf{C}$ to $\mathbf{D}$. A *natural transformation* σ from F to G, notation $\sigma : F \Rightarrow G$, is a family of morphisms $\sigma_X : F(X) \to G(X)$ in $\mathbf{D}$ indexed by objects $X \in \mathbf{C}$, subject to the following requirement: for each morphism $f : X \to Y$ in $\mathbf{C}$ we have $\sigma_Y \circ F(f) = G(f) \circ \sigma_X$. In other words, the following diagram commutes:

$$\begin{tikzcd}
F(X) \arrow[r, "\sigma_X"] \arrow[d, "F(f)"'] & G(X) \arrow[d, "G(f)"] \\
F(Y) \arrow[r, "\sigma_Y"'] & G(Y)
\end{tikzcd}$$

Natural transformations are the morphisms in the category of functors from $\mathbf{C}$ to $\mathbf{D}$. This category is written as $\mathbf{Fun}(\mathbf{C}, \mathbf{D},)$.

Examples 5.7.

1. Composing the functor $(-)^{\mathrm{ab}} : \mathbf{Grp} \to \mathbf{Ab}$ with the forgetful functor $\mathbf{Ab} \to \mathbf{Grp}$ yields a functor $\mathbf{Grp} \to \mathbf{Grp}$, again called $(-)^{\mathrm{ab}}$. There is a natural transformation $\pi : \mathrm{Id} \Rightarrow (-)^{\mathrm{ab}}$ whose components $\pi_G : G \to G/[G, G]$ are the projections onto the quotient group. To show that this transformation is natural, we have to prove that the diagram

$$\begin{tikzcd}
G \arrow[r, "\pi"] \arrow[d, "f"'] & G/[G, G] \arrow[d, "f^{\mathrm{ab}}"] \\
H \arrow[r, "\pi"'] & H/[H, H]
\end{tikzcd}$$

   commutes. But this follows immediately from the definition of f^{ab}.

2. Let G be a group, considered as a category with one object. We will take a look at natural transformations from the identity functor $\mathrm{Id} : G \to G$ to itself. Such a transformation consists of a family of morphisms from $*$ to $*$ indexed by objects of G, but since G has only one object as a category, this is the same as a single morphism from $*$ to $*$. Such a morphism is an element g of G. The naturality condition tells that $hg = gh$ for all $h \in G$. It follows that natural transformations from Id to Id are exactly elements in the center of G.

5.2 Adjunctions

We have seen that there is a nice connection between tensor products and spaces of linear maps. For vector spaces this connection is elaborated in Proposition A.10, and for representations in Proposition 1.5. We will rephrase this connection in terms of functors. A vector space U gives a tensor product functor $F = U \otimes (-) : \mathbf{Vect}_{\mathbb{C}} \to \mathbf{Vect}_{\mathbb{C}}$, and a Hom functor $G = \mathrm{Hom}(U, -) : \mathbf{Vect}_{\mathbb{C}} \to \mathbf{Vect}_{\mathbb{C}}$. There is a bijective correspondence between maps $f : F(V) \to W$ and maps $g : V \to G(W)$, in which the maps f and g determine each other via the equation $f(u \otimes v) = g(v)(u)$. Analogous situations for other functors are abundant in mathematics, and are called adjunctions. Loosely speaking, an adjunction consists of functors $F : \mathbf{C} \to \mathbf{D}$ and $G : \mathbf{D} \to \mathbf{C}$, in such a way that there is a bijection between morphisms $F(X) \to Y$ and morphisms $X \to F(Y)$. This bijection is subject to a naturality condition, which looks complicated but is often trivial to check in practice.

Definition 5.8. A functor $F : \mathbf{C} \to \mathbf{D}$ is said to be *left adjoint* to a functor $G : \mathbf{D} \to \mathbf{C}$ if for each $X \in \mathbf{C}$ and $Y \in \mathbf{D}$ there is a bijection $\Phi_{X,Y} : \mathrm{Hom}(F(X), Y) \to \mathrm{Hom}(X, G(Y))$ that is natural in X and Y. The latter requirement means the following: if $f : X' \to X$ and $g : Y \to Y'$ are morphisms in $\mathbf{C}$ and $\mathbf{D}$, respectively, then for all $h : F(X) \to Y$ in $\mathbf{D}$ we have that

$$\Phi_{X',Y'}(g \circ h \circ F(f)) = G(g) \circ \Phi_{X,Y}(h) \circ f.$$

Instead of "F is left adjoint to G" one sometimes says "G is right adjoint to F" or "F and G form an adjunction between $\mathbf{C}$ and $\mathbf{D}$", and one writes $F \dashv G$.

Not every functor has a left or a right adjoint, but whenever an adjoint exists, it is unique. There are several equivalent definitions of an adjunction, but here we will only use the above one.

Examples 5.9.

1. We will show that the abelianization functor $(-)^{\mathrm{ab}} : \mathbf{Grp} \to \mathbf{Ab}$ is left adjoint to the forgetful functor $\mathbf{Ab} \to \mathbf{Grp}$. To this end, let A be an abelian group, and G an arbitrary group. We have to establish a bijective correspondence between homomorphisms $G/[G,G] \to A$ and homomorphisms $G \to A$. In one direction, a morphism $h : G \to A$ gives rise to $\Phi(h) : G/[G,G] \to A$ defined by $\Phi(h)(g[G,G]) = h(g)$. This map is well-defined since A is abelian. In the other direction, a map $k : G/[G,G] \to A$ induces a map $\Phi^{-1}(k) : G \to A$ by composing with the canonical projection, i.e. $\Phi^{-1}(k)(g) = k(g[G,G])$. These constructions are clearly mutually inverse, and the naturality requirement follows from the definition of Φ.

2. For a fixed vector space U there is an adjunction

$$\begin{tikzcd} \mathbf{Vect}_{\mathbb{C}} \arrow[d, bend left, "\mathrm{Hom}(U,-)"] \arrow[d, phantom, "\dashv"] \\ \mathbf{Vect}_{\mathbb{C}} \arrow[u, bend left, "U\otimes(-)"] \end{tikzcd}$$

   The proof that this is indeed an adjunction has been sketched above, the details are left to the reader.

3. The forgetful functor $U : \mathbf{Rep}_G \to \mathbf{Vect}_{\mathbb{C}}$ has both a left and a right adjoint:

$$\begin{tikzcd} \mathbf{Rep}_G \arrow[d, "U" description] \\ \mathbf{Vect}_{\mathbb{C}} \arrow[u, bend left=60, "\mathbb{C}[G]\otimes(-)", "\dashv"'] \arrow[u, bend right=60, "\mathrm{Hom}(\mathbb{C}[G],-)"', "\dashv"] \end{tikzcd}$$

The functors are defined as in items 3 and 4 of Examples 5.4, with $U = \mathbb{C}[G]$.

4. Let $f : H \to G$ be a group homomorphism. Apply the functor **Rep** from Example 5.5 to obtain a functor $f^* : \mathbf{Rep}_G \to \mathbf{Rep}_H$. The left adjoint of f^* is the induced representation functor $V \mapsto \mathbb{C}[G] \otimes_{\mathbb{C}[H]} V$, and the right adjoint of f^* is the coinduced representation functor $V \mapsto \mathrm{Hom}_{\mathbb{C}[H]}(\mathbb{C}[G], V)$. The space $\mathrm{Hom}_{\mathbb{C}[H]}(\mathbb{C}[G], V)$ consists of linear maps $f : \mathbb{C}[G] \to V$ for which $f(ax) = af(x)$ for each $a \in \mathbb{C}[H]$ and $x \in \mathbb{C}[G]$. The statement that the induced representation is left adjoint to the restriction functor f^* is a categorical reformulation of Frobenius reciprocity.

   Often these constructions are considered in the situation where H is a subgroup of G, and f is the inclusion. In the special case where H is the one-element subgroup of G, we recover the previous example.

5.3 Tannaka duality

For any group G, we can form its category of representations $\mathbf{Rep}_G$. We now turn to the reverse problem: is it possible to reconstruct a group from its category of representations? It turns out that this is indeed possible, using some of the categorical concepts we have developed before.

If G is a group, then the category of representations comes equipped with a forgetful functor $U : \mathbf{Rep}_G \to \mathbf{Vect}_{\mathbb{C}}$. We will show that G can be reconstructed from the collection of natural transformations from U to itself. Explicitly, such a natural transformation $\sigma : U \Rightarrow U$ consists of a natural family of linear maps $\sigma_{(V,\rho)} : U(V) \to U(V)$ indexed by representations (V, ρ). Sometimes $\sigma_{(V,\rho)}$ is abbreviated to σ_V. These linear maps need not be intertwiners, because we applied the forgetful functor to the representations, so they are simply maps from the underlying vector space to itself. Naturality means that for each *intertwiner* $\varphi : V \to W$ the following diagram commutes:

$$\begin{tikzcd} V \arrow[r, "\sigma_V"] \arrow[d, "\varphi"'] & V \arrow[d, "\varphi"] \\ W \arrow[r, "\sigma_W"'] & W \end{tikzcd}$$

The collection of natural transformations from U to itself will be denoted by $\mathrm{End}(U)$.

We will look at some basic properties of natural transformations in $\mathrm{End}(U)$. These transformations behave well with respect to subrepresentations and direct sums of representations.

Lemma 5.10. *Let $\sigma : U \Rightarrow U$ be a natural transformation.*

1. *If V is a subrepresentation of W, then $(\sigma_W)|_V = \sigma_V$. Here $(\sigma_W)|_V$ indicates the restriction of $\sigma_W : U(W) \to U(W)$ to $U(V) \subseteq U(W)$.*

2. *The transformation σ preserves direct sums, i.e. $\sigma_{V\oplus W} = \sigma_V \oplus \sigma_W$.*

Proof.

1. Since V is a subrepresentation of W, the inclusion map $i : V \to W$ is an intertwiner. Hence the diagram

$$\begin{tikzcd} V \arrow[r, "\sigma_V"] \arrow[d, "i"'] & V \arrow[d, "i"] \\ W \arrow[r, "\sigma_W"'] & W \end{tikzcd}$$

commutes by naturality of σ. This means that $\sigma_V(v) = \sigma_W(v)$ for each $v \in V$, or equivalently $(\sigma_W)|_V = \sigma_V$.

2. The projection map $\pi_V : V \oplus W \to V$ given by $\pi_V(v, w) = v$ is an intertwiner, so it fits in a commutative diagram

$$\begin{tikzcd} V \oplus W \arrow[r, "\sigma_{V\oplus W}"] \arrow[d, "\pi_V"'] & V \oplus W \arrow[d, "\pi_V"] \\ V \arrow[r, "\sigma_V"'] & V \end{tikzcd}$$

Commutativity of this diagram says that $\pi_V(\sigma_{V\oplus W}(v, w)) = \sigma_V(v)$, in other words, the first component of $\sigma_{V\oplus W}(v, w)$ is $\sigma_V(v)$. Likewise one proves that its second component is $\sigma_W(w)$, which gives $\sigma_{V\oplus W}(v, w) = (\sigma_V(v), \sigma_W(w))$. An alternative way of writing this is $\sigma_{V\oplus W} = \sigma_V \oplus \sigma_W$. □

All natural transformations preserve subrepresentations and direct sums, but unfortunately not all of them preserve tensor products. Therefore it is often good to restrict the class of transformations under consideration to the tensor-preserving ones.

Definition 5.11. A natural transformation $\sigma : U \Rightarrow U$ is *monoidal* if:

- σ preserves tensor products: $\sigma_{V\otimes W} = \sigma_V \otimes \sigma_W$.
- The value of σ on the trivial representation $\mathbb{C}$ is $\sigma_{\mathbb{C}} = \mathrm{id}_{\mathbb{C}}$.

Note that on the left-hand side in the first condition we take a tensor product of representations V and W, while on the right-hand side we take a tensor product of a linear map $U(V) \to U(V)$ with a map $U(W) \to U(W)$. The second condition ensures that σ also preserves the empty tensor product, which is the trivial representation. The collection of monoidal natural transformations $\sigma : U \Rightarrow U$ is denoted $\mathrm{End}_{\otimes}(U)$. We will now show that the group G can be recovered as the group of monoidal natural transformations from U to itself.

Theorem 5.12 (Tannaka)**.** *Let G be a finite group, and $U : \mathbf{Rep}_G \to \mathbf{Vect}_{\mathbb{C}}$ the corresponding forgetful functor. The set $\mathrm{End}_{\otimes}(U)$ of monoidal natural transformations $\sigma : U \Rightarrow U$ forms a group under composition, and this group is isomorphic to G.*

Proof. Define a map $f : G \to \mathrm{End}_{\otimes}(U)$ by letting $f(g)$ be the natural transformation with components $f(g)_V : U(V) \to U(V)$, $f(g)_V(v) = g \cdot v$. In other words, if we write the representation V as (V, ρ), then the component $f(g)_{(V,\rho)}$ is the map $\rho(g)$. Then $f(g)$ is a natural transformation, because any intertwiner $\varphi : V \to W$ satisfies $\varphi \circ f(g)_V = f(g)_W \circ \varphi$. From the definition of a tensor product of representations it follows that $f(g)$ is always a monoidal transformation, so f is well-defined. Furthermore, the map f maps products to compositions because ρ is a homomorphism. We shall prove that f is a bijection; then it will follow that $\mathrm{End}_{\otimes}(U)$ is a group and that f is an isomorphism between G and $\mathrm{End}_{\otimes}(U)$.

To show that f is injective, suppose that $f(g) = f(h)$. Then all components of the natural transformations $f(g)$ and $f(h)$ are equal, so for each representation V, we have $f(g)_V = f(h)_V$. This holds in particular for the regular representation $V = \mathbb{C}[G]$. Writing out the definition of f then gives $g \cdot v = h \cdot v$ for every $v \in \mathbb{C}[G]$. Plugging in the unit of the group for v shows that $g = h$, which establishes injectivity.

The proof of surjectivity is more involved and will be presented using a sequence of claims. Let $\sigma \in \mathrm{End}_{\otimes}(U)$ be arbitrary; we seek a group element $g \in G$ such that $f(g) = \sigma$. First we define an element of the group ring instead: let $a = \sigma_{\mathbb{C}[G]}(e) \in \mathbb{C}[G]$, that is, a is the result of applying the component at the regular representation $\sigma_{\mathbb{C}[G]} : \mathbb{C}[G] \to \mathbb{C}[G]$ to the unit $e \in G \subseteq \mathbb{C}[G]$.

Claim 1. a is nonzero.

Proof. Let $\mathbb{C}$ be the trivial representation and define an intertwiner $\varepsilon : \mathbb{C}[G] \to \mathbb{C}$ that maps every basis vector g to 1. Hence on a linear combination it acts as

$$\varepsilon(a_1 g_1 + \cdots + a_n g_n) = a_1 + \cdots + a_n.$$

Naturality of σ gives a commutative diagram

$$\begin{tikzcd} \mathbb{C}[G] \arrow[r, "\sigma_{\mathbb{C}[G]}"] \arrow[d, "\varepsilon"'] & \mathbb{C}[G] \arrow[d, "\varepsilon"] \\ \mathbb{C} \arrow[r, "\mathrm{id}"'] & \mathbb{C} \end{tikzcd}$$

where we used that $\sigma_{\mathbb{C}} = \mathrm{id}$, because σ is monoidal. Commutativity of the diagram implies that $\varepsilon(\sigma_{\mathbb{C}[G]}(g)) = \varepsilon(g) = 1$ for each $g \in G$. In particular, for $g = e$ we get $\varepsilon(a) = 1$, which is only possible if $a \neq 0$.

Claim 2. The element $a \in \mathbb{C}[G]$ is actually an element of G. That means, if we express a in terms of basis vectors as $a = a_1 g_1 + \cdots + a_n g_n$, then exactly one a_i is one and the other coefficients are zero.

Proof. Define a map $\Delta : \mathbb{C}[G] \to \mathbb{C}[G] \otimes \mathbb{C}[G]$ on basis vectors by $\Delta(g) = g \otimes g$. Extending this linearly to an arbitrary vector of $\mathbb{C}[G]$ gives

$$\Delta(a_1 g_1 + \cdots + a_n g_n) = a_1(g_1 \otimes g_1) + \cdots + a_n(g_n \otimes g_n).$$

Beware that we cannot conclude that $\Delta(x) = x \otimes x$ for an arbitrary $x \in \mathbb{C}[G]$, since this latter equation also involves cross terms.

Since Δ is an intertwiner, naturality of σ gives a commutative diagram

$$\begin{tikzcd}
\mathbb{C}[G] \arrow[r, "\sigma_{\mathbb{C}[G]}"] \arrow[d, "\Delta"'] & \mathbb{C}[G] \arrow[d, "\Delta"] \\
\mathbb{C}[G] \otimes \mathbb{C}[G] \arrow[r, "\sigma_{\mathbb{C}[G] \otimes \mathbb{C}[G]}"'] & \mathbb{C}[G] \otimes \mathbb{C}[G]
\end{tikzcd}$$

Hence for any $g \in G$ we have

$$\Delta(\sigma_{\mathbb{C}[G]}(g)) = \sigma_{\mathbb{C}[G] \otimes \mathbb{C}[G]}(g \otimes g).$$

Because σ is monoidal,

$$\Delta(\sigma_{\mathbb{C}[G]}(g)) = \sigma_{\mathbb{C}[G]}(g) \otimes \sigma_{\mathbb{C}[G]}(g).$$

Plug in $g = e$ and use $a = \sigma_{\mathbb{C}[G]}(e)$ to obtain

$$\Delta(a) = a \otimes a.$$

Write $a = a_1 g_1 + \cdots + a_n g_n$, then this equation becomes

$$\begin{aligned}
& a_1(g_1 \otimes g_1) + \cdots + a_n(g_n \otimes g_n) \\
= \quad & a_1^2(g_1 \otimes g_1) + a_1 a_2(g_1 \otimes g_2) + \cdots + a_1 a_n(g_1 \otimes g_n) \\
& + \cdots \\
& + a_n a_1(g_n \otimes g_1) + a_n a_2(g_n \otimes g_2) + \cdots + a_n^2(g_n \otimes g_n).
\end{aligned}$$

Comparing coefficients at the basis vectors $g_i \otimes g_i$ gives $a_i^2 = a_i$ for all i, so each a_i is either 0 or 1. Comparing coefficients at the other basis vectors shows that no two different a_i can be 1, so at most one a_i is 1 and all others are zero. But by our previous claim a is not identically zero, hence exactly one a_i is 1.

Since we know that a is an element of G, we can now try to prove that $f(a) = \sigma$.

Claim 3. The transformations $f(a)$ and σ are equal at the regular representation, i.e. $f(a)_{\mathbb{C}[G]} = \sigma_{\mathbb{C}[G]}$.

Proof. For any $y \in \mathbb{C}[G]$, let $r_y : \mathbb{C}[G] \to \mathbb{C}[G]$ be the map "right multiplication by y", given by $r_y(x) = xy$ for each $x \in \mathbb{C}[G]$. Because this map is an intertwiner, naturality of σ gives $\sigma_{\mathbb{C}[G]} \circ r_y = r_y \circ \sigma_{\mathbb{C}[G]}$, whence $\sigma_{\mathbb{C}[G]}(xy) = \sigma_{\mathbb{C}[G]}(x)y$ for all $x, y \in \mathbb{C}[G]$. It follows that

$$f(a)_{\mathbb{C}[G]}(y) = ay = \sigma_{\mathbb{C}[G]}(e)y = \sigma_{\mathbb{C}[G]}(ey) = \sigma_{\mathbb{C}[G]}(y)$$

for every $y \in \mathbb{C}[G]$, and hence $f(a)_{\mathbb{C}[G]}$ and $\sigma_{\mathbb{C}[G]}$ are equal.

Claim 4. If two natural transformations $\sigma, \tau : U \Rightarrow U$ are equal at the regular representation, then $\sigma = \tau$.

Proof. Every irreducible representation is a subrepresentation of the regular representation by Proposition 1.14. Hence σ and τ are equal at all irreducible representations by part 1 of Lemma 5.10. But every representation is a direct sum of irreducibles, so part 2 of the same lemma shows that the transformations are equal everywhere.

The final two claims together prove that f is surjective. $\square$

A Useful linear algebra

A.1 Linear maps

Most concepts in linear algebra have an abstract and a concrete version. The abstract version uses vector spaces and linear maps, while the concrete version uses vectors in a euclidean space $\mathbb{C}^n$ and matrices. Abstract version are usually better for proving theorems, and concrete versions for doing computations. These two versions can be connected using bases. A basis for a vector space V is a collection of vectors $e_1, \ldots, e_n$ such that the following properties hold:

- Linear independence: if $\lambda_1 e_1 + \cdots + \lambda_n e_n = 0$, then $\lambda_1 = \cdots = \lambda_n = 0$.
- Completeness: every $v \in V$ can be written as a linear combination $v = \lambda_1 e_1 + \cdots + \lambda_n e_n$.

Thus any vector in v can be expressed in terms of basis vectors in a unique way, and a choice of basis provides an isomorphism from V onto $\mathbb{C}^n$. For an explicit description of this isomorphism, write a vector $v \in V$ as a linear combination $v = \lambda_1 e_1 + \cdots + \lambda_n e_n$; then the isomorphism maps v to $(\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^n$.

This shows that a basis connects vectors in an abstract vector space with concrete vectors in $\mathbb{C}^n$. Similarly bases connect linear maps with matrices. Let $\varphi : V \to W$ be a linear map and take bases $e_1, \ldots, e_n$ for V and $f_1, \ldots, f_m$ for W. Then the map φ is completely determined by the values $\varphi(e_i)$. Expanding these values in the basis for W gives

$$\varphi(e_i) = a_{i1} f_1 + \cdots + a_{im} f_m.$$

The coefficients a_{ij} can be collected into an $n \times m$ matrix $A = (a_{ij})$, which characterizes the map φ. Thus a choice of basis provides an isomorphism between linear maps and matrices.

We will look at several ways to construct new vector spaces from old ones. Each of these constructions can be applied to both vector spaces and linear maps. For each construction we will describe an abstract version and a version using bases and matrices.

Our first construction is the vector space of linear maps. The collection of all linear maps $\varphi : V \to W$ forms a vector space under pointwise operations. This vector space is written as $\mathrm{Hom}(V, W)$. Bases for the spaces V and W give a basis for $\mathrm{Hom}(V, W)$:

Proposition A.1. *Let V and W be vector spaces with bases $e_1, \ldots, e_n$ and $f_1, \ldots, f_m$, respectively. Then the maps $\alpha_{ij} : V \to W$ defined by $\alpha_{ij}(e_i) = f_j$ and $\alpha_{ij}(e_{i'}) = 0$ for $i' \neq i$ form a basis for* $\mathrm{Hom}(V, W)$.

Proof. The matrix representation of α_{ij} has a 1 on position (i, j) and zeroes everywhere else. These matrices clearly form a basis for the space of all $n \times m$ matrices. $\square$

There is a special case of the construction of the vector space of linear maps $\mathrm{Hom}(V, W)$, where we take the one-dimensional space $\mathbb{C}$ for W. The *dual space* of a vector space V consists of all linear maps from V to $\mathbb{C}$ and is denoted V^*. Specializing Proposition A.1 to the dual space gives the following.

Proposition A.2. *If V has basis $e_1, \ldots, e_n$, then the dual space V^* has basis $e_1^*, \ldots, e_n^*$, where $e_i^* : V \to \mathbb{C}$ is defined on basis vectors by*

$$e_i^*(e_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Also linear maps have a dual. If $f : V \to W$ is linear, then it defines a dual map $f^* : W^* \to V^*$ by $f^*(\phi) = \phi \circ f$ for $\phi : W \to \mathbb{C}$. Dualizing reverses the order of composition, in the sense that $(f \circ g)^* = g^* \circ f^*$. In matrix form, the dual map corresponds to the transpose matrix:

Proposition A.3. *Let $\varphi : V \to W$ be linear, $e_1, \ldots, e_n$ a basis for V, and $f_1, \ldots, f_m$ a basis for W. If A is the matrix representing φ with respect to these bases, then its transpose A^T represents $\varphi^* : W^* \to V^*$ with respect to the dual bases $f_1^*, \ldots, f_m^*$ and $e_1^*, \ldots, e_n^*$.*

Proof. Since the matrix $A = (a_{ij})$ represents φ, we can write $\varphi(e_i) = \sum_j a_{ij} f_j$. We wish to show that the entry at position (j, i) of φ^* is a_{ij}. To achieve this, it suffices to show that $\varphi^*(f_j^*) = \sum_i a_{ij} e_i^*$. Both sides of this equation are maps from V to $\mathbb{C}$, so we can show that they are equal by evaluating both at all basis vectors e_i. For the left-hand side we obtain

$$\varphi^*(f_j^*)(e_i) = f_j^*(\varphi(e_i)) = f_j^*\left(\sum_k a_{ik} f_k\right) = \sum_k a_{ik} f_j^*(f_k) = a_{ij},$$

while the right-hand side becomes

$$\left(\sum_k a_{kj} e_k^*\right)(e_i) = \sum_k a_{kj} e_k^*(e_i) = a_{ij},$$

proving the desired. □

A.2 Direct sums

Let V and W be two vector spaces. Their *direct sum* $V \oplus W$ is the vector space of pairs (v, w) for $v \in V$ and $w \in W$ with pointwise operations. Thus the direct sum is the same as the product $V \times W$. We write this vector space as a direct sum, since $(v, w) = (v, 0) + (0, w)$, so each vector in $V \oplus W$ can be expressed uniquely as a sum of a vector in V and a vector in W. Another reason for writing the product as a direct sum is that we get a basis for $V \oplus W$ by taking the union of the bases for V and W, not by taking their product. More precisely:

Proposition A.4. *Let* $e_1, \ldots, e_n$ *be a basis for* V *and* $f_1, \ldots, f_m$ *a basis for* W*. Then* $(e_1, 0), \ldots, (e_n, 0), (0, f_1,), \ldots, (0, f_m)$ *is a basis for* $V \oplus W$*. As a consequence,* $\dim(V \oplus W) = \dim V + \dim W$*.*

Two linear maps $\varphi : V \to V'$ and $\psi : W \to W'$ can be combined into one map $\varphi \oplus \psi : V \oplus W \to V' \oplus W'$, defined by $(\varphi \oplus \psi)(v, w) = (\varphi(v), \psi(w))$. Thus we can not only take the direct sum of vector spaces, but also of linear maps. If we write φ as a matrix A and ψ as a matrix B, then the matrix of $\varphi \oplus \psi$ is obtained by putting A and B as blocks on the diagonal, yielding the matrix $\left[\begin{smallmatrix} A & 0 \\ 0 & B \end{smallmatrix}\right]$.

The direct sum interacts with the space of linear maps in a nice way.

Proposition A.5. *Let* U*,* V*, and* W *be vector spaces.*

1. *The vector spaces* $\mathrm{Hom}(U \oplus V, W)$ *and* $\mathrm{Hom}(U, W) \oplus \mathrm{Hom}(V, W)$ *are isomorphic.*

2. *The vector spaces* $\mathrm{Hom}(U, V \oplus W)$ *and* $\mathrm{Hom}(U, V) \oplus \mathrm{Hom}(U, W)$ *are isomorphic.*

Proof. We will only prove the first point; the second one is similar. Given a map $\varphi : U \oplus V \to W$, form the maps $\varphi_1 : U \to W$ and $\varphi_2 : V \to W$ by $\varphi_1(u) = \varphi(u, 0)$ and $\varphi_2(v) = \varphi(0, v)$. Then the function $\Phi : \mathrm{Hom}(U \oplus V, W) \to \mathrm{Hom}(U, W) \oplus \mathrm{Hom}(V, W)$ that sends φ to the pair (φ_1, φ_2) is linear. It has an inverse $\Psi : \mathrm{Hom}(U, W) \oplus \mathrm{Hom}(V, W) \to \mathrm{Hom}(U \oplus V, W)$ that maps a pair (φ_1, φ_2) to the function $(u, v) \mapsto \varphi_1(u) + \varphi_2(v)$.

To show that Ψ is indeed an inverse, take a function $\varphi : U \oplus V \to W$. Applying Φ gives the pair (φ_1, φ_2), and applying Ψ to this pair gives the function $(u, v) \mapsto \varphi_1(u) + \varphi_2(v) = \varphi(u, 0) + \varphi(0, v) = \varphi((u, 0) + (0, v)) = \varphi(u, v)$, hence $\Psi \circ \Phi = \mathrm{id}$. For the other direction, start with a pair (φ_1, φ_2). Applying Ψ gives the single function $(u, v) \mapsto \varphi_1(u) + \varphi_2(v)$. Applying Φ to this function gives a pair again. The first component of this pair is $u \mapsto \varphi_1(u) + \varphi_2(0) = \varphi_1(u)$, and the second component is $v \mapsto \varphi_1(0) + \varphi_2(v) = \varphi_2(v)$. Thus $\Psi(\Psi(\varphi_1, \varphi_2)) = (\varphi_1, \varphi_2)$, completing the proof. $\square$

The matrix version of the above result states that we can combine several matrices into one large block matrix. For example, the proposition implies that $\mathrm{Hom}(V \oplus W, V' \oplus W') \cong \mathrm{Hom}(V, V') \oplus \mathrm{Hom}(W, V') \oplus \mathrm{Hom}(V, W') \oplus \mathrm{Hom}(W, W')$. We will describe the isomorphism Φ from right to left. Take linear maps

$$\begin{array}{ll} \varphi_{11} : V \to V', & \varphi_{12} : W \to V', \\ \varphi_{21} : V \to W', & \varphi_{22} : W \to W' \end{array}$$

and represent φ_{ij} by the matrix A_{ij}. Then we obtain a map $V \oplus W \to V' \oplus W'$ by putting these matrices as blocks in a larger matrix

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

This block acts on a vector (v, w) by mapping it to $(A_{11}v + A_{12}w, A_{21}v + A_{22}w) \in V' \oplus W'$. This is precisely what we get if we write (v, w) as a column vector and apply block multiplication to it. Saying that the map Φ is an isomorphism means that any linear map between the direct sums can be decomposed into blocks.

A.3 Tensor products

A map $\varphi : U \oplus V \to W$ out of a direct sum is linear if and only if $\varphi(u+u', v+v') = \varphi(u, v) + \varphi(u', v')$ and $\varphi(\lambda u, \lambda v) = \lambda\varphi(u, v)$. In other words, it has to be linear in both variables simultaneously. One often encounters maps out of a direct sum that are linear in each variable separately instead. Such a map is called bilinear.

Definition A.6. If U, V, and W are vector spaces, then a *bilinear map* from $U \oplus V$ to W is a map φ for which

$$\begin{aligned} \varphi(u + u', v) &= \varphi(u, v) + \varphi(u', v) & \qquad \varphi(\lambda u, v) &= \lambda\varphi(u, v) \\ \varphi(u, v + v') &= \varphi(u, v) + \varphi(u, v') & \qquad \varphi(u, \lambda v) &= \lambda\varphi(u, v) \end{aligned}$$

Bilinear maps are often more difficult to work with than linear maps. Therefore it would be nice to have a procedure to convert bilinear maps into linear maps. For this we will use a construction on vector spaces called the tensor product. For any two vector space U and V, we will construct a new vector space $U \otimes V$, in such a way that bilinear maps $U \oplus V \to W$ correspond bijectively to linear maps $U \otimes V \to W$.

There are two descriptions of the tensor product $V \otimes W$ of two vector spaces V and W. We will first discuss an abstract definition. The vector space $V \otimes W$ is generated by all vectors of the form $v \otimes w$, for $v \in V$ and $w \in W$. The tensor symbol $\otimes$ in $v \otimes w$ is just a formal symbol to denote a generating vector of the tensor product $V \otimes W$, so the vector $v \otimes w$ can also be thought of as the pair (v, w). Since the vector space $V \otimes W$ is generated by the elements $v \otimes w$, an arbitrary element of $V \otimes W$ looks like

$$\sum_{i=1}^{n} \lambda_i v_i \otimes w_i$$

for certain $\lambda_i \in \mathbb{C}$, $v_i \in V$, and $w_i \in W$. We stipulate that these elements are subject to the following relations:

$$\begin{aligned} (v + v') \otimes w &= v \otimes w + v' \otimes w & \qquad (\lambda v) \otimes w &= \lambda(v \otimes w) \\ v \otimes (w + w') &= v \otimes w + v \otimes w' & \qquad v \otimes (\lambda w) &= \lambda(v \otimes w) \end{aligned}$$

These relations can be seen as "rules" to perform calculations with the vectors in $V \otimes W$. They say that the symbol $\otimes$ behaves as a bilinear operation: it is linear in each variable separately.

The second, more concrete description of the tensor product involves a choice of basis. Let $e_1, \ldots, e_n$ be a basis for V and $f_1, \ldots, f_m$ a basis for W. Then

$V \otimes W$ is the vector space with basis vectors $e_i \otimes f_j$. In particular, if V has dimension n and W has dimension m, the tensor product $V \otimes W$ has dimension nm. This description can be used to give a more concrete picture of a vector $v \otimes w$, by expressing it in terms of basis vectors. First write v and w as a linear combination of basis vectors:

$$v = \sum_{i=1}^{n} \lambda_i e_i, \quad w = \sum_{j=1}^{m} \mu_j f_j$$

Then the relations for the tensor product give:

$$\begin{aligned} v \otimes w &= \left(\sum_{i=1}^{n} \lambda_i e_i \right) \otimes \left(\sum_{j=1}^{m} \mu_j f_j \right) \\ &= \sum_{i=1}^{n} \left(\lambda_i e_i \otimes \left(\sum_{j=1}^{m} \mu_j f_j \right) \right) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} (\lambda_i e_i) \otimes (\mu_j f_j) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} \lambda_i \mu_j (e_i \otimes f_j) \end{aligned}$$

and this is an expression of $v \otimes w$ in terms of basis vectors.

Example A.7. We take a look at the tensor product $\mathbb{C}^2 \otimes \mathbb{C}^3$. Since the dimension of the tensor product is the product of the dimensions of the constituent spaces, $\mathbb{C}^2 \otimes \mathbb{C}^3$ is 6-dimensional, and hence it is isomorphic to $\mathbb{C}^6$. An arbitrary vector in $\mathbb{C}^2$ is of the form (α, β), and in $\mathbb{C}^3$ of the form $(\gamma, \delta, \varepsilon)$. We wish to express their tensor product $(\alpha, \beta) \otimes (\gamma, \delta, \varepsilon)$ as a vector in $\mathbb{C}^6$. For this we can use the above calculation. Write the standard bases of $\mathbb{C}^2$ and $\mathbb{C}^3$ as e_1, e_2 and f_1, f_2, f_3 respectively. Then $(\alpha, \beta) = \alpha e_1 + \beta e_2$ and $(\gamma, \delta, \varepsilon) = \gamma f_1 + \delta f_2 + \varepsilon f_3$. Therefore

$$\begin{aligned} (\alpha, \beta) \otimes (\gamma, \delta, \varepsilon) &= \alpha\gamma(e_1 \otimes f_1) + \alpha\delta(e_1 \otimes f_2) + \alpha\varepsilon(e_1 \otimes f_3) \\ &\quad + \beta\gamma(e_2 \otimes f_1) + \beta\delta(e_2 \otimes f_2) + \beta\varepsilon(e_2 \otimes f_3) \\ &= (\alpha\gamma, \alpha\delta, \alpha\varepsilon, \beta\gamma, \beta\delta, \beta\varepsilon), \end{aligned}$$

since the vectors $e_i \otimes f_j$ form the standard basis of $\mathbb{C}^6$.

Tensor products can be used to transform bilinear maps into linear maps, since bilinear maps from $U \oplus V$ correspond to linear maps from the tensor product $U \otimes V$:

Proposition A.8. *There is a one-to-one correspondence between bilinear maps $\varphi : U \oplus V \to W$ and linear maps $\psi : U \otimes V \to W$. On the generating vectors of the tensor product, this correspondence is given by $\psi(u \otimes v) = \varphi(u, v)$.*

This proposition holds because the generating relations for the tensor product amount to the same as the conditions for a bilinear map. It provides a way of constructing linear maps out of a tensor product by finding a suitable bilinear map. Therefore we can use this result to establish some basic properties of tensor products.

Proposition A.9.

1. *The tensor product is commutative:* $U \otimes V \cong V \otimes U$.
2. *The tensor product is associative:* $U \otimes (V \otimes W) \cong (U \otimes V) \otimes W$.
3. *The ground field is a unit for the tensor product:* $\mathbb{C} \otimes V \cong V \otimes \mathbb{C} \cong V$.
4. *The tensor product distributes over the direct sum:* $U \otimes (V \oplus W) \cong (U \otimes V) \oplus (U \otimes W)$.

Proof. We will only prove commutativity here; the other properties can be obtained in a similar way. To produce a linear map $U \otimes V \to V \otimes U$, we have to find a bilinear map $\varphi : U \oplus V \to V \otimes U$. For this we take $\varphi(u, v) = v \otimes u$. Then φ is bilinear because

$$\varphi(u + u', v) = v \otimes (u + u') = v \otimes u + v \otimes u' = \varphi(u, v) + \varphi(u', v),$$

and similarly for the other variable and for scalar multiplications. Therefore φ gives a unique linear map ψ for which $\psi(u \otimes v) = \varphi(u, v) = v \otimes u$. To show that ψ is an isomorphism, it suffices to show that it has a linear inverse $\chi : V \otimes U \to U \otimes V$. This χ is obtained via a bilinear map in the same way as ψ, so it satisfies $\chi(v \otimes u) = u \otimes v$. To check that χ is an inverse for ψ, note that

$$\psi\chi(v \otimes u) = \psi(u \otimes v) = v \otimes u.$$

Since the vectors $v \otimes u$ generate the space $V \otimes U$ and the map $\psi\chi$ is equal to id on all generators, it is equal to id everywhere. Similarly $\chi\psi = \mathrm{id}$, so ψ is indeed an isomorphism. $\square$

So far we have only considered tensor products of vector spaces. It is also possible to take the tensor product of two linear maps. Again there is an abstract definition and a definition in terms of matrices. Abstractly, if $\varphi : V \to V'$ and $\psi : W \to W'$ are two linear maps, then their tensor product $\varphi \otimes \psi : V \otimes W \to V' \otimes W'$ is defined on generators by $(\varphi \otimes \psi)(v \otimes w) = \varphi(v) \otimes \psi(w)$. This extends uniquely to a linear map defined on all vectors in $V \otimes W$.

If we write the maps φ and ψ as matrices, then the construction of their tensor product is known as the *Kronecker product*. Suppose that $\dim V = n$, $\dim V' = n'$, $\dim W = m$, and $\dim W' = m'$. Then φ can be written as an $n \times n'$ matrix A with indices A_{ij}, and ψ as an $m \times m'$ matrix B with indices B_{kl}. The space $V \otimes W$ has dimension mn and the space $V' \otimes W'$ has dimension $m'n'$, so $\varphi \otimes \psi$ is represented by an $mn \times m'n'$ matrix $A \otimes B$. Its entries can be indexed by pairs $(A \otimes B)_{(i,k)(j,l)}$, where $1 \leq i \leq n$, $1 \leq k \leq m$, $1 \leq j \leq n'$, and $1 \leq l \leq m'$. In this notation, the entries are given by

$$(A \otimes B)_{(i,k)(j,l)} = A_{ij}B_{kl}.$$

This can be checked by applying both sides of the equation to all basis vectors; the computation is similar to the tensor product of two vectors. For example, the Kronecker product of two 2×2 matrices becomes:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \otimes \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{bmatrix}$$

We will now present two results that relate tensor products to spaces of linear maps and dual spaces.

Proposition A.10. *The vector spaces* $\mathrm{Hom}(U \otimes V, W)$ *and* $\mathrm{Hom}(U, \mathrm{Hom}(V, W))$ *are isomorphic.*

Expanding the definitions, we see that a linear map from U to $\mathrm{Hom}(V, W)$ is the same as a bilinear map $U \oplus V \to W$. Thus the above result follows from Proposition A.8. It can be used to translate between maps from a tensor product and maps into a space of linear maps, so it is sometimes called the *Hom-tensor adjunction*.

Proposition A.11. *If* W *is finite-dimensional, then the vector spaces* $V^* \otimes W$ *and* $\mathrm{Hom}(V, W)$ *are isomorphic.*

Proof. We will provide two proofs. For the first proof, we will construct an explicit isomorphism $\Phi : V^* \otimes W \to \mathrm{Hom}(V, W)$. There is a bilinear map $\Phi' : V^* \oplus W \to \mathrm{Hom}(V, W)$ given by

$$\Phi'(\varphi, w)(v) = \varphi(v)w.$$

This gives a unique $\Phi : V^* \otimes W \to \mathrm{Hom}(V, W)$ for which $\Phi(\varphi \otimes w)(v) = \varphi(v)w$. To check that it is an isomorphism, we will first prove that it is surjective. Let $\psi \in \mathrm{Hom}(V, W)$ be any linear map and let $e_1, \ldots, e_n$ be a basis for W. Expand ψ in this basis as

$$\psi(v) = \sum_i \psi_i(v)e_i;$$

then $\sum_i \psi_i \otimes e_i$ is a pre-image of ψ because

$$\Phi\left(\sum_i \psi_i \otimes e_i\right)(v) = \sum_i \psi_i(v)e_i = \psi(v).$$

Since the spaces $V^* \otimes W$ and $\mathrm{Hom}(V, W)$ have the same dimension, it follows that they are isomorphic.

Alternatively, we can use that if W is finite-dimensional, then it is isomorphic to $\mathbb{C}^n$. Since $\mathbb{C}^n$ is the direct sum of n copies of $\mathbb{C}$, Proposition A.5 and the fact that the tensor product distributes over direct sums give

$$\mathrm{Hom}(V, \mathbb{C}^n) \cong \mathrm{Hom}(V, \mathbb{C})^n = (V^*)^n \cong V^* \otimes \mathbb{C}^n. \qquad \square$$

A.4 Inner products

Many vector spaces are equipped with an inner product. This is a map $\langle -,- \rangle : V \times V \to \mathbb{C}$ subject to the following conditions.

- Sesquilinearity: $\langle u, \lambda v + \mu w \rangle = \lambda \langle u, v \rangle + \mu \langle u, w \rangle$ for all $u, v, w \in V$ and $\lambda, \mu \in \mathbb{C}$.
- Skew-symmetry: $\langle v, w \rangle = \overline{\langle w, v \rangle}$ for all $v, w \in V$.
- Non-degeneracy: $\langle v, v \rangle \geq 0$ with equality if and only if $v = 0$.

The first two conditions together imply that the inner product is antilinear in the first variable, that is, $\langle \lambda u + \mu v, w \rangle = \overline{\lambda} \langle u, w \rangle + \overline{\mu} \langle v, w \rangle$. A vector space together with an inner product is called an inner product space.

When working with an inner product space, we would like other objects to interact well with the inner product. For instance, we would like a basis for the space to be orthonormal. An *orthonormal basis* for V is a basis $e_1, \ldots, e_n$ such that $\langle e_i, e_i \rangle = 1$ and $\langle e_i, e_j \rangle = 0$ for all $i \neq j$. When expressing a vector $v \in V$ in terms of such an orthonormal basis, the coefficient for the basis vector e_i is $\langle e_i, v \rangle$. In other words, every $v \in V$ can be written as

$$v = \sum_{i=1}^{n} \langle e_i, v \rangle e_i.$$

Every inner product space possesses an orthonormal basis, since an arbitrary basis can be made orthonormal by applying the Gram-Schmidt process.

The following result gives a quick way to prove that a given set of vectors forms an orthonormal basis.

Proposition A.12. *Let $e_1, \ldots, e_n$ be a set of vectors in V satisfying the following requirements:*

- *For all i we have $\langle e_i, e_i \rangle = 1$.*
- *For all $i \neq j$ we have $\langle e_i, e_j \rangle = 0$.*
- *If $\langle v, e_i \rangle = 0$ for all i, then $v = 0$.*

Then $e_1, \ldots, e_n$ form an orthonormal basis for V.

Inner products and orthonormal bases also give an easy method to check if two linear maps are equal.

Proposition A.13. *Let $\varphi : V \to W$ be a linear map, let $e_1, \ldots, e_n$ be an orthonormal basis for V, and $f_1, \ldots, f_m$ an orthonormal basis for W. The following are equivalent:*

1. *$\varphi = \psi$.*
2. *$\langle w, \varphi(v) \rangle = \langle w, \psi(v) \rangle$ for all $v \in V$ and $w \in W$.*

3. $\langle f_j, \varphi(e_i)\rangle = \langle f_j, \psi(e_i)\rangle$ *for all* i *and* j*.*

Proof. The implications $1 \Rightarrow 2 \Rightarrow 3$ are evident. For the remaining implication $3 \Rightarrow 1$, fix a number i and assume that $\langle f_j, \varphi(e_i)\rangle = \langle f_j, \psi(e_i)\rangle$ for all j. Express φ as a matrix $A = (a_{ij})$ with respect to the bases, and express ψ as a matrix $B = (b_{ij})$. Then $\varphi(e_i) = \sum_j a_{ij} f_j$ and $\psi(e_i) = \sum_j b_{ij} f_j$. Since our bases are orthonormal, it follows that $a_{ij} = \langle f_j, \varphi(e_i)\rangle = \langle f_j, \psi(e_i)\rangle = b_{ij}$ for all j. Since i was arbitrary, all coefficients of A and B are equal, hence $\varphi = \psi$. □

An inner product also gives rise to a bijection between maps $V \to W$ and maps $W \to V$. This bijection involves the construction of the hermitian adjoint of an operator.

Proposition A.14. *Let* V *and* W *be inner product spaces. For any linear* $\varphi : V \to W$ *there exists a unique* $\psi : W \to V$ *such that* $\langle v, \psi(w)\rangle = \langle \varphi(v), w\rangle$ *for all* $v \in V$ *and* $w \in W$*.*

Proof. Let $e_1, \ldots, e_n$ be an orthonormal basis for V and define ψ by $\psi(w) = \sum_i \langle \varphi(e_i), w\rangle e_i$. Then

$$\begin{aligned}\langle v, \psi(w)\rangle = \langle v, \sum_i \langle \varphi(e_i), w\rangle e_i\rangle &= \sum_i \langle \varphi(e_i), w\rangle \langle v, e_i\rangle = \sum_i \overline{\langle e_i, v\rangle} \langle \varphi(e_i), w\rangle \\ &= \langle \varphi\left(\sum_i \langle e_i, v\rangle e_i\right), w\rangle = \langle \varphi(v), w\rangle,\end{aligned}$$

where we used that the inner product is antilinear in the first variable ($\langle \lambda v, w\rangle = \overline{\lambda}\langle v, w\rangle$) and that any $v \in V$ can be expressed as $v = \sum_i \langle e_i, v\rangle e_i$. Uniqueness of ψ follows from Proposition A.13. □

The unique map ψ from the above proposition is written as $\varphi^\dagger$, and called the *hermitian adjoint* of φ. From the formula $\langle v, \varphi^\dagger(w)\rangle = \langle \varphi(v), w\rangle$ we can derive a number of properties of the hermitian adjoint:

$$\begin{array}{ll} (\varphi + \psi)^\dagger = \varphi^\dagger + \psi^\dagger & \qquad (\varphi \circ \psi)^\dagger = \psi^\dagger \circ \varphi^\dagger \\ (\lambda\varphi)^\dagger = \overline{\lambda}\varphi^\dagger & \qquad \varphi^{\dagger\dagger} = \varphi \end{array}$$

Proposition A.15. *If* A *is the matrix associated to* $\varphi : V \to W$ *(with respect to fixed bases of* V *and* W*), then the conjugate transpose* $\overline{A}^{\mathsf{T}}$ *is the matrix associated to* $\varphi^\dagger$*.*

Proof. Exercise. □

Index