Representation Theory of Finite Groups

Benjamin Steinberg
School of Mathematics and Statistics
Carleton University
bsteinbg@math.carleton.ca

December 15, 2009

Preface

This book arose out of course notes for a fourth year undergraduate/first year graduate course that I taught at Carleton University. The goal was to present group representation theory at a level that is accessible to students who have not yet studied module theory and who are unfamiliar with tensor products. For this reason, the Wedderburn theory of semisimple algebras is completely avoided. Instead, I have opted for a Fourier analysis approach. This sort of approach is normally taken in books with a more analytic flavor; such books, however, invariably contain material on representations of compact groups, something that I would also consider beyond the scope of an undergraduate text. So here I have done my best to blend the analytic and the algebraic viewpoints in order to keep things accessible. For example, Frobenius reciprocity is treated from a character point of view to evade use of the tensor product.

The only background required for this book is a basic knowledge of linear algebra and group theory, as well as familiarity with the definition of a ring. The proof of Burnside's theorem makes use of a small amount of Galois theory (up to the fundamental theorem) and so should be skipped if used in a course for which Galois theory is not a prerequisite. Many things are proved in more detail than one would normally expect in a textbook; this was done to make things easier on undergraduates trying to learn what is usually considered graduate level material.

The main topics covered in this book include: character theory; the group algebra; Burnside's pq-theorem and the dimension theorem; permutation representations; induced representations and Mackey's theorem; and the representation theory of the symmetric group.

It should be possible to present this material in a one semester course. Chapters 2–5 should be read by everybody; it covers the basic character theory of finite groups. The first two sections of Chapter 6 are also recommended for all readers; the reader who is less comfortable with Galois theory can then skip the last section and move on to Chapter 7 on permutation representations, which is needed for Chapters 8–10. Chapter 10, on the representation theory of the symmetric group, can be read immediately after Chapter 7.

Although this book is envisioned as a text for an advanced undergraduate or introductory graduate level course, it is also intended to be of use for mathematicians who may not be algebraists, but need group representation theory for their work.

When preparing this book I have relied on a number of classical references on representation theory, including [2–4, 6, 9, 13, 14]. For the representation theory of the symmetric group I have drawn from [4, 7, 8, 10–12]; the approach is due to James [11]. Good references for applications of representation theory to computing eigenvalues of graphs and random walks are [3,4]. Discrete Fourier analysis and its applications can be found in [1, 4].

Contents

List of Tables

Chapter 1

Introduction

The representation theory of finite groups is a subject going back to the late eighteen hundreds. The earliest pioneers in the subject were Frobenius, Schur and Burnside. Modern approaches tend to make heavy use of module theory and the Wedderburn theory of semisimple algebras. But the original approach, which nowadays can be thought of as via discrete Fourier analysis, is much more easily accessible and can be presented, for instance, in an undergraduate course. The aim of this textbook is to exposit the essential ingredients of the representation theory of finite groups over the complex numbers assuming only linear algebra and undergraduate group theory, and perhaps a minimal familiarity with ring theory.

The original purpose of representation theory was to serve as a powerful tool for obtaining information about finite groups via the methods of linear algebra, such as eigenvalues, inner product spaces and diagonalization. The first major triumph of representation theory was Burnside's pq-theorem, which states that a non-abelian group of order $p^a q^b$ with p, q prime cannot be simple, or equivalently, that every finite group of order $p^a q^b$ with p, q prime is solvable. Representation theory went on to play an indispensable role in the classification of finite simple groups.

However, representation theory is much more than just a means to study the structure of finite groups. It is also a fundamental tool with applications to many areas of mathematics and statistics, both pure and applied. For instance, sound compression is very much based on the fast Fourier transform for finite abelian groups. Fourier analysis on finite groups also plays an important role in probability and statistics, especially in the study of random walks on groups, such as card-shuffling and diffusion processes [1, 4], and in the analysis of data [5]. Applications of representation theory to graph theory, and in particular to the construction of expander graphs, can be found in [3]. Some applications along these lines, especially toward the computation of eigenvalues of Cayley graphs, is given in this text.

Chapter 2

Review of Linear Algebra

This chapter reviews the linear algebra that we shall assume throughout this text. In this book all vector spaces considered will be finite dimensional over the field $\mathbb{C}$ of complex numbers.

2.1 Notation

This section introduces our standing notation.

- If X is a set of vectors, then $\mathbb{C}X = \text{Span } X$.
- $M_{mn}(\mathbb{C}) = \{m \times n \text{ matrices with entries in } \mathbb{C}\}$.
- $M_n(\mathbb{C}) = M_{nn}(\mathbb{C})$.
- $\text{Hom}(V, W) = \{A\colon V \to W \mid A \text{ is a linear map}\}$.
- $\text{End}(V) = \text{Hom}(V, V)$ (the *endomorphism ring* of V).
- $GL(V) = \{A \in \text{End}(V) \mid A \text{ is invertible}\}$ (known as the *general linear group of* V).
- $GL_n(\mathbb{C}) = \{A \in M_n(\mathbb{C}) \mid A \text{ is invertible}\}$.
- $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$.
- $\mathbb{Z}_n = \{\overline{0}, \ldots, \overline{n-1}\}$.

Throughout we will abuse the distinction between $GL(\mathbb{C}^n)$ and $GL_n(\mathbb{C})$ by identifying an invertible transformation with its matrix with respect to the standard basis $\{e_1, \ldots, e_n\}$. Suppose $\dim V = n$ and $\dim W = m$. Then:

$$\begin{aligned} \mathrm{End}(V) &\cong M_n(\mathbb{C}); \\ GL(V) &\cong GL_n(\mathbb{C}); \\ \mathrm{Hom}(V, W) &\cong M_{mn}(\mathbb{C}). \end{aligned}$$

Notice that $GL_1(\mathbb{C}) \cong \mathbb{C}^*$ and so we shall always work with the latter. We indicate W is a subspace of V by writing $W \leq V$.

2.2 Complex inner product spaces

An *inner product*[1] on V is a map $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{C}$ such that:

(a) $\langle v, c_1w_1 + c_2w_2 \rangle = c_1\langle v, w_1 \rangle + c_2\langle v, w_2 \rangle$;

(b) $\langle w, v \rangle = \overline{\langle v, w \rangle}$;

(c) $\langle v, v \rangle \geq 0$ and $\langle v, v \rangle = 0$ if and only if $v = 0$.

A vector space equipped with an inner product is called an *inner product space*. The *norm* $\|v\|$ of a vector v in an inner product space is defined by $\|v\| = \sqrt{\langle v, v \rangle}$.

Example 2.2.1. The *standard inner product* on $\mathbb{C}^n$ is given by

$$\langle (a_1, \ldots, a_n), (b_1, \ldots, b_n) \rangle = \sum_{i=1}^{n} \overline{a_i} b_i.$$

Recall that two vectors v, w in an inner product space V are said to be *orthogonal* if $\langle v, w \rangle = 0$. A subset of V is called *orthogonal* if the elements of V are pairwise orthogonal. If in addition, the norm of each vector is 1, the set is termed *orthonormal*. An orthogonal set of non-zero vectors is always linearly independent, in particular any orthonormal set is linearly independent. If $\{e_1, \ldots, e_n\}$ is an orthonormal basis for an inner product space V and $v \in V$, then $v = \langle e_1, v \rangle e_1 + \cdots + \langle e_n, v \rangle e_n$.

[1] Our choice to make the second variable linear is typical in physics; many mathematicians use the opposite convention.

Example 2.2.2. For a finite set X, the set $\mathbb{C}^X = \{f\colon X \to \mathbb{C}\}$ is a vector space with pointwise operations. Namely, one defines

$$\begin{aligned}(f+g)(x) &= f(x) + g(x);\\ (cf)(x) &= cf(x).\end{aligned}$$

For each $x \in X$, define a function $\delta_x\colon X \to \mathbb{C}$ by

$$\delta_x(y) = \begin{cases} 1 & x = y \\ 0 & x \neq y. \end{cases}$$

There is a natural inner product on $\mathbb{C}^X$ given by

$$\langle f, g\rangle = \sum_{x\in X} \overline{f(x)}g(x).$$

The set $\{\delta_x \mid x \in X\}$ is an orthonormal basis with respect to this inner product. If $f \in \mathbb{C}^X$, then its unique expression as a linear combination of the δ_x is given by

$$f = \sum_{x\in X} f(x)\delta_x.$$

Consequently, $\dim \mathbb{C}^X = |X|$.

If $W_1, W_2 \leq V$, then $W_1 + W_2 = \{w_1 + w_2 \mid w_1 \in W_1, w_2 \in W_2\}$. This is the smallest subspace containing W_1 and W_2. If in addition $W_1 \cap W_2 = \{0\}$, then $W_1 + W_2$ is called a *direct sum*, written $W_1 \oplus W_2$. As vector spaces, $W_1 \oplus W_2 \cong W_1 \times W_2$. In fact, if V and W are any two vector spaces, one can form their *external direct sum* by setting $V \oplus W = V \times W$. Note that

$$\dim(W_1 \oplus W_2) = \dim W_1 + \dim W_2.$$

More precisely, if B_1 is a basis for W_1 and B_2 is a basis for W_2, then $B_1 \cup B_2$ is a basis for $W_1 \oplus W_2$.

Direct sum decompositions are easy to obtain in inner product spaces. If $W \leq V$, then the *orthogonal complement* of W is the subspace

$$W^\perp = \{v \in V \mid \langle v, w\rangle = 0 \text{ for all } w \in W\}.$$

Proposition 2.2.3. *If V is an inner product space and $W \leq V$, then there results a direct sum decomposition $V = W \oplus W^\perp$.*

Proof. First, if $w \in W \cap W^\perp$ then $\langle w, w\rangle = 0$ implies $w = 0$; so $W \cap W^\perp = \{0\}$. Let $\mathrm{proj}_W \colon V \to W$ be the orthogonal projection to W. Then, for $v \in V$, we have $\mathrm{proj}_W(v) \in W$, $v - \mathrm{proj}_W(v) \in W^\perp$ and

$$v = \mathrm{proj}_W(v) + (v - \mathrm{proj}_W(v)).$$

This completes the proof. □

A linear map $U \in GL(V)$ is said to be *unitary* if $\langle Uv, Uw\rangle = \langle v, w\rangle$ for all $v, w \in V$. The set $U(V)$ of unitary maps is a subgroup of $GL(V)$.

Example 2.2.4. If $U = (u_{ij})$, then U^* is the conjugate transpose of U, i.e., $U^* = (\overline{u_{ji}})$. For the standard inner product on $\mathbb{C}^n$, $U \in GL_n(\mathbb{C})$ is unitary if and only if $U^{-1} = U^*$. We denote by $U_n(\mathbb{C})$ the set of all $n \times n$ unitary matrices. A matrix $A \in M_n(\mathbb{C})$ is called *self-adjoint* if $A^* = A$.

2.3 Further notions from linear algebra

If $X \subseteq \mathrm{End}(V)$ and $W \leq V$, then W is called X*-invariant* if, for any $A \in X$ and any $w \in W$, one has $Aw \in W$, i.e., $XW \subseteq W$.

A key example comes from the theory of eigenvalues and eigenvectors. Recall that $\lambda \in \mathbb{C}$ is an *eigenvalue* of $A \in \mathrm{End}(V)$ if $\lambda I - A$ is not invertible; in other words, if $Av = \lambda v$ for some $v \neq 0$. The *eigenspace* corresponding to λ is the set $V_\lambda = \{v \in V \mid Av = \lambda v\}$, which is a subspace of V. Note that if $v \in V_\lambda$, then $A(Av) = A(\lambda v) = \lambda Av$, so $Av \in V_\lambda$. Thus V_λ is A-invariant. Conversely, if $W \leq V$ is A-invariant with $\dim W = 1$ (that is, W is a line), then $W \subseteq V_\lambda$ for some λ. In fact, if $w \in W \setminus \{0\}$, then $\{w\}$ is a basis for W. Since $Aw \in W$, we have that $Aw = \lambda w$ for some $\lambda \in \mathbb{C}$. So w is an eigenvector with eigenvalue λ, whence $w \in V_\lambda$; thus $W \subseteq V_\lambda$.

Recall that the *characteristic polynomial* $p_A(x)$ of a linear operator A on an n-dimensional vector space V is given by $p_A(x) = \det(xI - A)$. This is a monic polynomial of degree n and the roots of $p_A(x)$ are exactly the eigenvalues of A.

Theorem 2.3.1 (Cayley-Hamilton)**.** *Let $p_A(x)$ be the characteristic polynomial of A. Then $p_A(A) = 0$.*

If $A \in \mathrm{End}(V)$, *the minimal polynomial* of A, denoted $m_A(x)$, is the smallest degree monic polynomial $f(x)$ such that $f(A) = 0$.

Fact 2.3.2. *If $q(A) = 0$ then $m_A(x) \mid q(x)$.*

Proof. Write $q(x) = m_A(x)f(x) + r(x)$ with either $r(x) = 0$, or $\deg(r(x)) < \deg(m_A(x))$. Then

$$0 = q(A) = m_A(A)f(A) + r(A) = r(A).$$

By minimality of $m_A(x)$, we conclude that $r(x) = 0$. □

Corollary 2.3.3. *If $p_A(x)$ is the characteristic polynomial of A, then $m_A(x)$ divides $p_A(x)$.*

The relevance of the minimal polynomial is that it provides a criterion for diagonalizability of a matrix, amongst other things.

Theorem 2.3.4. *A matrix $A \in M_n(\mathbb{C})$ is diagonalizable if and only if $m_A(x)$ has no repeated roots.*

Example 2.3.5. For the matrix

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$m_A(x) = (x-1)(x-3)$, whereas $p_A(x) = (x-1)^2(x-3)$. On the other hand, the matrix

$$B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

has $m_B(x) = (x-1)^2 = p_B(x)$ and so is not diagonalizable.

One of the main results from linear algebra is the spectral theorem for matrices.

Theorem 2.3.6 (Spectral Theorem)**.** *Let $A \in M_n(\mathbb{C})$ be self-adjoint. Then there is a unitary matrix $U \in U_n(\mathbb{C})$ such that UAU^* is diagonal. Moreover, the eigenvalues of A are real.*

The *trace* of a matrix $A = (a_{ij})$ is defined by

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Some basic facts concerning the trace function $\mathrm{Tr}\colon M_n(\mathbb{C}) \to \mathbb{C}$ are that Tr is linear and $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. Consequently $\mathrm{Tr}(PAP^{-1}) = \mathrm{Tr}(P^{-1}PA) = \mathrm{Tr}(A)$. In particular, this shows that $\mathrm{Tr}(A)$ does not depend on the basis and so if $T \in \mathrm{End}(V)$, then $\mathrm{Tr}(T)$ makes sense: choose any basis and compute Tr of the associated matrix. Similar remarks apply to the determinant.

Chapter 3

Group Representations

The goal of group representation theory is to study groups via their actions on vector spaces. Consideration of groups acting on sets leads to such important results as the Sylow theorems. By acting on vector spaces even more detailed information about a group can be obtained. This is the subject of representation theory. As byproducts emerge Fourier analysis on finite groups and the study of complex-valued functions on a group.

3.1 Basic definitions and first examples

An action of a group G on a set X is the same thing as a homomorphism $\varphi\colon G \to S_X$, where S_X is the symmetric group on X. This motivates the following definition.

Definition 3.1.1 (Representation)**.** A *representation* of a group G is a homomorphism $\varphi\colon G \to GL(V)$ for some (finite-dimensional) non-zero vector space V. The dimension of V is called the *degree of* φ.

We usually write φ_g for $\varphi(g)$ and $\varphi_g(v)$, or simply $\varphi_g v$, for the action of φ_g on $v \in V$. Suppose that dim $V = n$. To a basis B for V, we can associate a vector space isomorphism $T\colon V \to \mathbb{C}^n$ by taking coordinates. More precisely, if $B = \{b_1, \ldots, b_n\}$, then $T(b_i) = e_i$ where e_i is the i^{th} standard unit vector. We can then define a representation $\psi\colon G \to GL_n(\mathbb{C})$ by setting $\psi_g = T\varphi_g T^{-1}$ for $g \in G$. If B' is another basis, we have another isomorphism $S\colon V \to \mathbb{C}^n$, and hence a representation $\psi'\colon G \to GL_n(\mathbb{C})$ given by $\psi'_g = S\varphi_g S^{-1}$. The representations ψ and ψ' are related via the formula $\psi'_g = ST^{-1}\psi_g TS^{-1} = (ST^{-1})\psi_g(ST^{-1})^{-1}$. We want to think of φ, ψ and ψ' as all being the same representation. This leads us to the important notion of equivalence.

Definition 3.1.2 (Equivalence)**.** Two representations $\varphi\colon G \to GL(V)$ and $\psi\colon G \to GL(W)$ are *equivalent* if there exists an isomorphism $T\colon V \to W$ such that $\psi_g = T\varphi_g T^{-1}$ for all $g \in G$, i.e., $\psi_g T = T\varphi_g$ for all $g \in G$. In this case, we write $\varphi \sim \psi$. In pictures, we have that the diagram

$$\begin{array}{ccc} V & \xrightarrow{\varphi_g} & V \\ {\scriptstyle T}\downarrow & & \downarrow{\scriptstyle T} \\ W & \xrightarrow[\psi_g]{} & W \end{array}$$

commutes, meaning that either of the two ways of going from the upper left to the lower right corner of the diagram give the same answer.

Example 3.1.3. Define $\varphi\colon \mathbb{Z}_n \to GL_2(\mathbb{C})$ by

$$\varphi_{\overline{m}} = \begin{bmatrix} \cos\left(\frac{2\pi m}{n}\right) & -\sin\left(\frac{2\pi m}{n}\right) \\ \sin\left(\frac{2\pi m}{n}\right) & \cos\left(\frac{2\pi m}{n}\right) \end{bmatrix},$$

which is the matrix for rotation by $2\pi m/n$, and $\psi\colon \mathbb{Z}_n \to GL_2(\mathbb{C})$ by

$$\psi_{\overline{m}} = \begin{bmatrix} e^{\frac{2\pi mi}{n}} & 0 \\ 0 & e^{\frac{-2\pi mi}{n}} \end{bmatrix}.$$

Then $\varphi \sim \psi$. To see this, let

$$A = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix},$$

and so

$$A^{-1} = \frac{1}{2i}\begin{bmatrix} 1 & i \\ -1 & i \end{bmatrix}.$$

Then direct computation shows

$$\begin{aligned} A^{-1}\varphi_{\overline{m}}A &= \frac{1}{2i}\begin{bmatrix} 1 & i \\ -1 & i \end{bmatrix}\begin{bmatrix} \cos\left(\frac{2\pi m}{n}\right) & -\sin\left(\frac{2\pi m}{n}\right) \\ \sin\left(\frac{2\pi m}{n}\right) & \cos\left(\frac{2\pi m}{n}\right) \end{bmatrix}\begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \\ &= \frac{1}{2i}\begin{bmatrix} e^{\frac{2\pi mi}{n}} & ie^{\frac{2\pi mi}{n}} \\ -e^{\frac{-2\pi mi}{n}} & ie^{\frac{-2\pi mi}{n}} \end{bmatrix}\begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \\ &= \frac{1}{2i}\begin{bmatrix} 2ie^{\frac{2\pi mi}{n}} & 0 \\ 0 & 2ie^{\frac{-2\pi mi}{n}} \end{bmatrix} \\ &= \psi_{\overline{m}}. \end{aligned}$$

The following representation of the symmetric group is very important.

Example 3.1.4 (Standard representation of S_n)**.** Define $\varphi\colon S_n \to GL_n(\mathbb{C})$ on basis elements by $\varphi_\sigma(e_i) = e_{\sigma(i)}$. One obtains the matrix for φ_σ by permuting the rows of the identity matrix according to σ. So, for instance, when $n = 3$ we have

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

Notice in Example 3.1.4 that

$$\varphi_\sigma(e_1 + e_2 + \cdots + e_n) = e_{\sigma(1)} + e_{\sigma(2)} + \cdots + e_{\sigma(n)} = e_1 + e_2 + \cdots + e_n$$

where the last equality holds since σ is a permutation and addition is commutative. Thus $\mathbb{C}(e_1 + \cdots + e_n)$ is invariant under all the φ_σ with $\sigma \in S_3$. This leads to the following definition.

Definition 3.1.5 (G-invariant subspace)**.** Let $\varphi\colon G \to GL(V)$ be a representation. A subspace $W \leq V$ is *G-invariant* if, for all $g \in G$ and $w \in W$, one has $\varphi_g w \in W$.

For ψ from Example 3.1.3, $\mathbb{C}e_1, \mathbb{C}e_2$ are both $\mathbb{Z}_n$-invariant and $\mathbb{C}^2 = \mathbb{C}e_1 \oplus \mathbb{C}e_2$. This is the kind of situation we would like to always happen.

Definition 3.1.6 (Direct sum of representations)**.** Suppose that representations $\varphi^{(1)}\colon G \to GL(V_1)$ and $\varphi^{(2)}\colon G \to GL(V_2)$ are given. Then their direct sum

$$\varphi^{(1)} \oplus \varphi^{(2)}\colon G \to GL(V_1 \oplus V_2)$$

is given by

$$(\varphi^{(1)} \oplus \varphi^{(2)})_g(v_1, v_2) = (\varphi^{(1)}_g(v_1), \varphi^{(2)}_g(v_2)).$$

Let's try to understand direct sums in terms of matrices. Suppose that $\varphi^{(1)}\colon G \to GL_m(\mathbb{C})$ and $\varphi^{(2)}\colon G \to GL_n(\mathbb{C})$ are representations. Then

$$\varphi^{(1)} \oplus \varphi^{(2)}\colon G \to GL_{m+n}(\mathbb{C})$$

has block matrix form

$$(\varphi^{(1)} \oplus \varphi^{(2)})_g = \begin{bmatrix} \varphi^{(1)}_g & 0 \\ 0 & \varphi^{(2)}_g \end{bmatrix}.$$

Example 3.1.7. Define $\varphi^{(1)}\colon \mathbb{Z}_n \to \mathbb{C}^*$ by $\varphi^{(1)}_{\overline{m}} = e^{\frac{2\pi i m}{n}}$, and $\varphi^{(2)}\colon \mathbb{Z}_n \to \mathbb{C}^*$ by $\varphi^{(2)}_{\overline{m}} = e^{\frac{-2\pi i m}{n}}$. Then

$$(\varphi^{(1)} \oplus \varphi^{(2)})_{\overline{m}} = \begin{bmatrix} e^{\frac{2\pi i m}{n}} & 0 \\ 0 & e^{\frac{-2\pi i m}{n}} \end{bmatrix}.$$

Since representations are a special kind of homomorphism, if a group G is generated by a set X, then a representation φ of G is determined by its values on X; of course, not any assignment of matrices to the generators gives a valid representation!

Example 3.1.8. Let $\rho\colon S_3 \to GL_2(\mathbb{C})$ be specified on the generators (1 2) and (1 2 3) by

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

and let $\psi\colon S_3 \to \mathbb{C}^*$ be defined by $\psi_\sigma = 1$. Then

$$(\rho \oplus \psi)_{(12)} = \begin{bmatrix} -1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ (\rho \oplus \psi)_{(123)} = \begin{bmatrix} -1 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We shall see later that $\rho \oplus \psi$ is equivalent to the representation of S_3 considered in Example 3.1.4.

Let $\varphi\colon G \to GL(V)$ be a representation. If $W \leq V$ is a G-invariant subspace, we may restrict φ to obtain a representation $\varphi|_W\colon G \to GL(W)$ by setting $(\varphi|_W)_g(w) = \varphi_g(w)$ for $w \in W$. Precisely because W is G-invariant, we have $\varphi_g(w) \in W$. Sometime one says $\varphi|_W$ is a *subrepresentation* of φ. If $W_1, W_2 \leq V$ are G-invariant and $V = W_1 \oplus W_2$, then one easily verifies $\varphi \sim \varphi|_{W_1} \oplus \varphi|_{W_2}$.

A particularly simple example of a representation is the trivial representation.

Example 3.1.9 (Trivial representation)**.** The trivial representation of a group G is the homomorphism $\varphi\colon G \to \mathbb{C}^*$ given by $\varphi(g) = 1$ for all $g \in G$.

If $n > 1$, then the representation $\rho\colon G \to GL_n(\mathbb{C})$ given by $\rho_g = I$ all $g \in G$ is *not* equivalent to the trivial representation; rather, it is equivalent to the direct sum of n copies of the trivial representation.

In mathematics, it is often the case that one has some sort of unique factorization into primes, or irreducibles. This is the case for representation theory. The notion of irreducible is modeled on the notion of a simple group.

Definition 3.1.10 (Irreducible)**.** A representation $\varphi\colon G \to GL(V)$ is said to be *irreducible* if the only G-invariant subspaces of V are $\{0\}$ and V.

Example 3.1.11. Any degree one representation $\varphi\colon G \to \mathbb{C}^*$ is irreducible, since $\mathbb{C}$ has no proper non-zero subspaces.

Table 3.1 exhibits some analogies between the concepts we have seen so far with ones from Group Theory and Linear Algebra.

Groups	Vector spaces	Representations
subgroup	subspace	G-invariant subspace
simple group	one-dimensional subspace	irreducible representation
direct product	direct sum	direct sum
isomorphism	isomorphism	equivalence

Table 3.1: Analogies between groups, vector spaces and representations

If $G = \{1\}$ is the trivial group and $\varphi\colon G \to GL(V)$ is a representation, then necessarily $\varphi_1 = I$. So a representation of the trivial group is the same datum as a vector space. For the trivial group, a G-invariant subspace is nothing more than a subspace. A representation of $\{1\}$ is irreducible if and only if it has degree one. So the middle column of the above table is a special case of the third column.

Example 3.1.12. The representations from Example 3.1.3 are not irreducible. For instance,

$$\mathbb{C}\begin{bmatrix} i \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbb{C}\begin{bmatrix} -i \\ 1 \end{bmatrix}$$

are $\mathbb{Z}_n$-invariant subspaces for φ, while the coordinate axes $\mathbb{C}e_1$ and $\mathbb{C}e_2$ are invariant subspaces for ψ.

Not surprisingly, after the one-dimensional representations, the next easiest class to analyze consists of the two-dimensional representations.

Example 3.1.13. The representation $\rho\colon S_3 \to GL_2(\mathbb{C})$ from Example 3.1.8 is irreducible.

Proof. Since $\dim\ \mathbb{C}^2 = 2$, any non-zero proper S_3-invariant subspace W is one-dimensional. Let v be a non-zero vector in W; so $W = \mathbb{C}v$. Then $\rho_\sigma(v) = \lambda v$ for some $\lambda \in \mathbb{C}$, since by S_3-invariance of W we have $\rho_\sigma(v) \in W = \mathbb{C}v$. It follows that v must be an eigenvector for all the ρ_σ, $\sigma \in S_3$.

Claim. $\rho_{(1\ 2)}$ and $\rho_{(1\ 2\ 3)}$ do not have a common eigenvector.

Indeed, direct computation reveals $\rho_{(1\ 2)}$ has eigenvalues 1 and -1 with $V_{-1} = \mathbb{C}e_1$ and $V_1 = \mathbb{C}\begin{bmatrix} -1 \\ 2 \end{bmatrix}$. Clearly e_1 is not an eigenvector of $\rho_{(1\ 2\ 3)}$, since $\rho_{(1\ 2\ 3)}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Also, $\rho_{(1\ 2\ 3)}\begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$, so $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ is not an eigenvector of $\rho_{(1\ 2\ 3)}$. Thus $\rho_{(1\ 2)}$ and $\rho_{(1\ 2\ 3)}$ have no common eigenvector, which implies that ρ is irreducible by the discussion above. □

Let us summarize as a proposition the idea underlying this example.

Proposition 3.1.14. *If $\varphi\colon G \to GL(V)$ is a representation of degree 2 (i.e., $\dim V = 2$), then φ is irreducible if and only if there is no common eigenvector v to all φ_g with $g \in G$.*

Notice that this trick of using eigenvectors only works for degree 2 representations.

Example 3.1.15. Let r be rotation by $\pi/2$ and s be reflection over the x-axis. These permutations generate the dihedral group D_4. Let the representation $\varphi\colon D_4 \to GL_2(\mathbb{C})$ be defined by

$$\varphi(r^k) = \begin{bmatrix} i^k & 0 \\ 0 & (-i)^k \end{bmatrix},\ \varphi(sr^k) = \begin{bmatrix} 0 & (-i)^k \\ i^k & 0 \end{bmatrix}.$$

Then one can apply the proposition to check that φ is an irreducible representation.

Our eventual goal is to show that each representation is equivalent to a direct sum of irreducible representations. Let us define some terminology to that effect.

Definition 3.1.16 (Completely reducible)**.** Let G be a group. A representation $\varphi\colon G \to GL(V)$ is said to be *completely reducible* if $V = V_1 \oplus V_2 \oplus \cdots \oplus V_n$ where the V_i are non-zero G-invariant subspaces and $\varphi|_{V_i}$ is irreducible for all $i = 1, \ldots, n$.

Equivalently, φ is completely reducible if $\varphi \sim \varphi^{(1)} \oplus \varphi^{(2)} \oplus \cdots \oplus \varphi^{(n)}$ where the $\varphi^{(i)}$ are irreducible representations.

Definition 3.1.17 (Decomposable)**.** We say that φ is *decomposable* if $V = V_1 \oplus V_2$ with V_1, V_2 non-zero G-invariant subspaces. Otherwise, V is called *indecomposable*.

If $T\colon V \to V$ is a linear transformation and B is a basis for V, then we shall use $[T]_B$ to denote the matrix for T in the basis B. Let $\varphi\colon G \to GL(V)$ be a decomposable representation, say with $V = V_1 \oplus V_2$ where V_1, V_2 are non-trivial G-invariant subspaces. Let $\varphi^{(i)} = \varphi_{V_i}$. Choose bases B_1 and B_2 for V_1 and V_2, respectively. Then it follows from the definition of a direct sum that $B = B_1 \cup B_2$ is a basis for V. Since V_i is G-invariant, we have $\varphi_g(B_i) \subseteq V_i = \mathbb{C}B_i$. Thus we have in matrix form

$$[\varphi_g]_B = \begin{bmatrix} [\varphi^{(1)}]_{B_1} & 0 \\ 0 & [\varphi^{(2)}]_{B_2} \end{bmatrix}$$

and so $\varphi \sim \varphi^{(1)} \oplus \varphi^{(2)}$.

Complete reducibility is the analogue of diagonalizability in representation theory. Our goal is to show that any representation of a finite group is completely reducible. To do this we show that any representation is either irreducible or decomposable, and then proceed by induction on the degree. First we must show that these notions depend only on the equivalence class of a representation.

Lemma 3.1.18. *Let $\varphi\colon G \to GL(V)$ be equivalent to a decomposable representation. Then φ is decomposable.*

Proof. Let $\psi\colon G \to GL(W)$ be a decomposable representation with $\psi \sim \varphi$ and $T\colon V \to W$ a vector space isomorphism with $\varphi_g = T^{-1}\psi_g T$. Suppose that W_1 and W_2 are non-zero invariant subspaces of W with $W = W_1 \oplus W_2$. Since T is an equivalence we have

$$\begin{array}{ccc} V & \xrightarrow{\varphi_g} & V \\ {\scriptstyle T}\downarrow & & \downarrow{\scriptstyle T} \\ W & \xrightarrow[\psi_g]{} & W \end{array}$$

commutes, i.e., $T\varphi_g = \psi_g T$, all $g \in G$. Let $V_1 = T^{-1}(W_1)$ and $V_2 = T^{-1}(W_2)$. First we claim $V = V_1 \oplus V_2$. Indeed, if $v \in V_1 \cap V_2$, then $Tv \in W_1 \cap W_2 = \{0\}$ and so $Tv = 0$. But T is injective so this implies $v = 0$. Next, if $v \in V$, then $Tv = w_1 + w_2$ some $w_1 \in W_1$ and $w_2 \in W_2$. Then $v = T^{-1}w_1 + T^{-1}w_2 \in V_1 + V_2$. Thus $V = V_1 \oplus V_2$.

Next we show that V_1, V_2 are G-invariant. If $v \in V_i$, then $\varphi_g v = T^{-1}\psi_g Tv$. But $Tv \in W_i$ implies $\psi_g Tv \in W_i$ since W_i is G-invariant. Therefore, we conclude that $\varphi_g v = T^{-1}\psi_g Tv \in T^{-1}(W_i) = V_i$, as required. □

Similarly, we have the following results, whose proofs we omit.

Lemma 3.1.19. *Let $\varphi\colon G \to GL(V)$ be equivalent to an irreducible representation. Then φ is irreducible.*

Lemma 3.1.20. *Let $\varphi\colon G \to GL(V)$ be equivalent to a completely reducible representation. Then φ is completely reducible.*

3.2 Maschke's theorem and complete reducibility

In order to effect direct sum decompositions of representations, we take advantage of the tools of inner products and orthogonal decompositions.

Definition 3.2.1 (Unitary representation)**.** Let V be an inner product space. A representation $\varphi\colon G \to GL(V)$ is called *unitary* if φ_g is unitary for all $g \in G$, i.e., $\langle \varphi_g(v), \varphi_g(w)\rangle = \langle v, w\rangle$ for all $v, w \in W$. In other words, $\varphi\colon G \to U(V)$.

Identifying $GL_1(\mathbb{C})$ with $\mathbb{C}^*$, we see that a complex number z is unitary if and only if $\overline{z} = z^{-1}$, that is $z\overline{z} = 1$. But this says exactly that $|z| = 1$, so $U_1(\mathbb{C})$ is exactly the unit circle S^1 in $\mathbb{C}$. Hence a one-dimensional unitary representation is a homomorphism $\varphi\colon G \to S^1$.

Example 3.2.2. Define $\varphi\colon \mathbb{R} \to S^1$ by $\varphi(t) = e^{2\pi it}$. Then φ is a unitary representation of $\mathbb{R}$ since $\varphi(t+s) = e^{2\pi i(t+s)} = e^{2\pi it}e^{2\pi is} = \varphi(t)\varphi(s)$.

A crucial fact is that every indecomposable unitary representation is irreducible as the following proposition shows.

Proposition 3.2.3. *Let $\varphi\colon G \to GL(V)$ be a unitary representation of a group. Then φ is either irreducible or decomposable.*

Proof. Suppose φ is not irreducible. Then there is a non-zero proper G-invariant subspace W of U. Its orthogonal complement $W^\perp$ is then also non-zero and $V = W \oplus W^\perp$. So it remains to prove that $W^\perp$ is G-invariant. If $v \in W^\perp$ and $w \in W$, then

$$\begin{aligned} \langle w, \varphi_g(v)\rangle &= \langle \varphi_{g^{-1}}(w), \varphi_{g^{-1}}\varphi_g(v)\rangle && (3.1)\\ &= \langle \varphi_{g^{-1}}(w), v\rangle && (3.2)\\ &= 0 && (3.3) \end{aligned}$$

where (3.1) follows since φ is unitary, (3.2) follows since $\varphi_{g^{-1}}\varphi_g = \varphi_1 = I$ and (3.3) follows since $\varphi_{g^{-1}}w \in W$, as W is G-invariant, and $v \in W^\perp$. □

It turns out that for finite groups every representation is equivalent to a unitary one. This is not true for infinite groups, as we shall see momentarily.

Proposition 3.2.4. *Every representation of a finite group G is equivalent to a unitary representation.*

Proof. Let $\varphi\colon G \to GL(V)$ be a representation where $\dim V = n$. Choose a basis B for V, and let $T\colon V \to \mathbb{C}^n$ be the isomorphism taking coordinates with respect to B. Then setting $\rho_g = T\varphi_g T^{-1}$, for $g \in G$, yields a representation $\rho\colon G \to GL_n(\mathbb{C})$ equivalent to φ_g. Let $\langle \cdot, \cdot \rangle$ be the standard inner product on $\mathbb{C}^n$. We define a new inner product $(\cdot, \cdot)$ on $\mathbb{C}^n$ using the crucial "averaging trick." It will be a frequent player throughout the course. Without further ado, define

$$(v, w) = \sum_{g \in G} \langle \rho_g v, \rho_g w \rangle.$$

This summation over G of course requires that G is finite. It can be viewed as a "smoothing" process.

Let us check that this is indeed an inner product. First we check:

$$\begin{aligned}(v, c_1 w_1 + c_2 w_2) &= \sum_{g \in G} \langle \rho_g v, \rho_g (c_1 w_1 + c_2 w_2) \rangle \\ &= \sum_{g \in G} \left(c_1 \langle \rho_g v, \rho_g w_1 \rangle + c_2 \langle \rho_g v, \rho_g w_2 \rangle \right) \\ &= c_1 \sum_{g \in G} \langle \rho_g v, \rho_g w_1 \rangle + c_2 \sum_{g \in G} \langle \rho_g v, \rho_g w_2 \rangle \\ &= c_1 (v, w_1) + c_2 (v, w_2).\end{aligned}$$

Next we verify:

$$\begin{aligned}(w, v) &= \sum_{g \in G} \langle \rho_g w, \rho_g v \rangle \\ &= \sum_{g \in G} \overline{\langle \rho_g v, \rho_g w \rangle} \\ &= \overline{(v, w)}.\end{aligned}$$

Finally, observe that

$$(v, v) = \sum_{g \in G} \langle \rho_g v, \rho_g v \rangle \geq 0$$

because each term $\langle \rho_g v, \rho_g v\rangle \geq 0$. If $(v,v) = 0$, then

$$0 = \sum_{g\in G} \langle \rho_g v, \rho_g v\rangle$$

which implies $\langle \rho_g v, \rho_g v\rangle = 0$ for all $g \in G$ since we are adding non-negative numbers. Hence $0 = \langle \rho_1 v, \rho_1 v\rangle = \langle v, v\rangle$, and so $v = 0$. We have now established that $(\cdot,\cdot)$ is an inner product.

To verify that the representation is unitary with respect to this inner product, we compute

$$(\rho_h v, \rho_h w) = \sum_{g\in G} \langle \rho_g \rho_h v, \rho_g \rho_h v\rangle = \sum_{g\in G} \langle \rho_{gh} v, \rho_{gh} w\rangle.$$

We now apply a change of variables, by setting $x = gh$. As g ranges over all G, x ranges over all elements of G since if $k \in G$, then when $g = kh^{-1}$, $x = k$. Therefore,

$$(\rho_h v, \rho_h w) = \sum_{x\in G} \langle \rho_x v, \rho_x w\rangle = (v, w).$$

This completes the proof. □

As a corollary we obtain that every indecomposable representation of a finite group is irreducible.

Corollary 3.2.5. *Let $\varphi\colon G \to GL(V)$ be a representation of a finite group. Then φ is either irreducible or decomposable.*

Proof. By Proposition 3.2.4, φ is equivalent to a unitary representation ρ. Proposition 3.2.3 then implies that ρ is either irreducible or decomposable. Lemmas 3.1.18 and 3.1.19 then yield that φ is either irreducible or decomposable, as was desired. □

The following example shows that Corollary 3.2.5 fails for infinite groups and hence Proposition 3.2.4 must also fail for infinite groups.

Example 3.2.6. We provide an example of an indecomposable representation of $\mathbb{Z}$, which is not irreducible. Define $\varphi\colon \mathbb{Z} \to GL_2(\mathbb{C})$ by

$$\varphi(n) = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}.$$

It is straightforward to verify that φ is a homomorphism. The vector e_1 is an eigenvector of $\varphi(n)$ for all $n \in \mathbb{Z}$ and so $\mathbb{C}e_1$ is a $\mathbb{Z}$-invariant subspace. This shows that φ is not irreducible. On the other hand, if φ were decomposable, it would be equivalent to a direct sum of one-dimensional representations. Such a representation is diagonal. But we saw in Example 2.3.5 that $\varphi(1)$ is not diagonalizable. It follows that φ is indecomposable.

Remark 3.2.7. Observe that any irreducible representation is indecomposable. The previous example shows that the converse fails.

The next theorem is the pinnacle of this chapter. Its proof is quite analogous to the proof of the existence of a prime factorization of an integer or of a factorization of polynomials into irreducibles.

Theorem 3.2.8 (Maschke)**.** *Every representation of a finite group is completely reducible.*

Proof. Let $\varphi\colon G \to GL(V)$ be a representation of a finite group G. The proof proceeds by induction on the degree of φ, that is $\dim V$. If $\dim V = 1$, then φ is irreducible since V has no non-zero proper subspaces. Assume the statement is true for $\dim V \leq n$. Let $\varphi\colon G \to GL(V)$ with $\dim V = n+1$. If φ is irreducible, then we are done. Otherwise, φ is decomposable by Corollary 3.2.5, so $V = V_1 \oplus V_2$ where $0 \neq V_1, V_2$ are G-invariant subspaces. Since $\dim V_1, \dim V_2 < \dim V$, by induction, $\varphi|_{V_1}$ and $\varphi|_{V_2}$ are completely reducible. Therefore, $V_1 = U_1 \oplus \cdots \oplus U_s$ and $V_2 = W_1 \oplus \cdots \oplus W_r$ where the U_i, W_j are G-invariant and the subrepresentations $\varphi|_{U_i}, \varphi|_{W_j}$ are irreducible for all $1 \leq i \leq s, 1 \leq j \leq r$. Then $V = U_1 \oplus \cdots U_s \oplus W_1 \oplus \cdots \oplus W_r$ and hence φ is completely irreducible. □

Remark 3.2.9. If one follows the details of the proof carefully, one can verify that if φ is a unitary matrix representation, then φ is equivalent to a direct sum of irreducible unitary representations via an equivalence implemented by a unitary matrix T.

In conclusion if $\varphi\colon G \to GL_n(\mathbb{C})$ is any representation of a finite group, then

$$\varphi \sim \begin{bmatrix} \varphi^{(1)} & 0 & \cdots & 0 \\ 0 & \varphi^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \varphi^{(m)} \end{bmatrix}$$

where the $\varphi^{(i)}$ are irreducible for all i. This is analogous to the spectral theorem stating that all self-adjoint matrices are diagonalizable.

There still remains the question as to whether the decomposition into irreducible representations is unique. This will be resolved in the next chapter.

Exercises

Exercise 3.1. Let $\varphi\colon D_4 \to \mathrm{GL}_2(\mathbb{C})$ be the representation given by

$$\varphi(r^k) = \begin{bmatrix} i^k & 0 \\ 0 & (-i)^k \end{bmatrix},\ \varphi(sr^k) = \begin{bmatrix} 0 & (-i)^k \\ i^k & 0 \end{bmatrix}$$

where r is rotation counterclockwise by $\pi/2$ and s is reflection over the x-axis. Prove that φ is irreducible. You may assume φ is a representation.

Exercise 3.2. Prove Lemma 3.1.19.

Exercise 3.3. Let $\varphi, \psi\colon G \to \mathbb{C}^*$ be one-dimensional representations. Show that φ is equivalent to ψ if and only if $\varphi = \psi$.

Exercise 3.4. Let $\varphi\colon G \to \mathbb{C}^*$ be a representation. Suppose $g \in G$ has order n.

1. Show that $\varphi(g)$ is an n^{th}-root of unity (i.e. a solution to the equation $z^n = 1$).
2. Construct n inequivalent one-dimensional representations $\mathbb{Z}_n \to \mathbb{C}^*$.
3. Explain why your representations are the only possible one-dimensional representations.

Exercise 3.5. Let $\varphi\colon G \to GL(V)$ be a representation of a finite group G. Define the fixed subspace

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

1. Show that V^G is a G-invariant subspace.
2. Show that
$$\frac{1}{|G|}\sum_{h\in G} \varphi_h v \in V^G$$
for all $v \in V$.
3. Show that if $v \in V^G$, then
$$\frac{1}{|G|}\sum_{h\in G} \varphi_h v = v.$$

Exercise 3.6. Let $\varphi\colon G \to GL_n(\mathbb{C})$ be a representation.

1. Show that setting $\psi_g = \overline{\varphi_g} = (\overline{\varphi_{ij}(g)})$ results in a representation $\psi\colon G \to GL_n(\mathbb{C})$ called the conjugate representation. Provide an example showing that φ and ψ do not have to be equivalent.

2. Let $\chi\colon G \to \mathbb{C}^*$ be a degree 1 representation of G. Define a map $\varphi^\chi\colon G \to GL_n(\mathbb{C})$ by $\varphi^\chi_g = \chi(g)\varphi_g$. Show that φ^χ is a representation. Give an example showing that φ and φ^χ do not have to be equivalent.

Chapter 4

Character Theory and the Orthogonality Relations

This chapter gets to the heart of group representation theory: the character theory. In particular, we establish the various orthogonality relations and use them to prove the uniqueness of the decomposition of a representation into irreducibles. An application to graph theory is presented in this chapter. In the next chapter, we use the results of this chapter to develop Fourier analysis on finite groups.

4.1 Homomorphisms of representations

To proceed, we shall need a notion of homomorphism of representations. The idea is the following. Let $\varphi\colon G \to GL(V)$ be a representation. We can think of elements of G as scalars via $g \cdot v = \varphi_g v$ for $v \in V$. A homomorphism between $\varphi\colon G \to GL(V)$ and $\rho\colon G \to GL(W)$ should be a linear transformation $T\colon V \to W$ such that $Tgv = gTv$ for all $g \in G$ and $v \in V$. Formally, this means $T\varphi_g v = \rho_g Tv$ all $v \in V$, i.e., $T\varphi_g = \rho_g T$ for all $g \in G$.

Definition 4.1.1 (Homomorphism)**.** Let $\varphi\colon G \to GL(V)$, $\rho\colon G \to GL(W)$ be representations. A *homomorphism*[1] from φ to ρ is by definition a linear

[1]Some authors use the term *intertwiner* for what we call homomorphism.

map $T\colon V \to W$ such that $T\varphi_g = \rho_g T$ for all $g \in G$, that is, the diagram

$$\begin{CD} V @>{\varphi_g}>> V \\ @V{T}VV @VV{T}V \\ W @>>{\rho_g}> W \end{CD}$$

commutes for all $g \in G$.

The set of all homomorphisms from φ to ρ is denoted $\mathrm{Hom}_G(\varphi, \rho)$. Notice that $\mathrm{Hom}_G(\varphi, \rho) \subseteq \mathrm{Hom}(V, W)$.

Remark 4.1.2. If $T \in \mathrm{Hom}_G(\varphi, \rho)$ is invertible, then $\varphi \sim \rho$ and T is an equivalence (or isomorphism).

Remark 4.1.3. Observe that $T\colon V \to V$ belongs to $\mathrm{Hom}_G(\varphi, \varphi)$ if and only if $T\varphi_g = \varphi_g T$ for all $g \in G$, i.e., T commutes with (or centralizes) $\varphi(G)$. In particular, the identity map $I\colon V \to V$ is always an element of $\mathrm{Hom}_G(\varphi, \varphi)$.

As is typical for homomorphisms in algebra, the kernel and the image of a homomorphism of representations are subrepresentations.

Proposition 4.1.4. *Let $T\colon V \to W$ be in $\mathrm{Hom}_G(\varphi, \rho)$. Then $\ker T$ is a G-invariant subspace of V and $T(V) = \mathrm{Im}\, T$ is a G-invariant subspace of W.*

Proof. Let $v \in \ker T$ and $g \in G$. Then $T\varphi_g v = \rho_g T v = 0$ since $v \in \ker T$. Hence $\varphi_g v \in \ker T$. We conclude $\ker T$ is G-invariant.

Now let $w \in \mathrm{Im}\, T$, say $w = Tv$ with $v \in V$. Then $\rho_g w = \rho_g T v = T\varphi_g v \in \mathrm{Im}\, T$, establishing that $\mathrm{Im}\, T$ is G-invariant. □

The set of homomorphisms from φ to ρ has the additional structure of a vector space, as the following proposition reveals.

Proposition 4.1.5. *Let $\varphi\colon G \to GL(V)$ and $\rho\colon G \to GL(W)$ be representations. Then $\mathrm{Hom}_G(\varphi, \rho)$ is a subspace of $\mathrm{Hom}(V, W)$.*

Proof. Let $T_1, T_2 \in \mathrm{Hom}_G(\varphi, \rho)$ and $c_1, c_2 \in \mathbb{C}$. Then

$$(c_1T_1 + c_2T_2)\varphi_g = c_1T_1\varphi_g + c_2T_2\varphi_g = c_1\rho_g T_1 + c_2\rho_g T_2 = \rho_g(c_1T_1 + c_2T_2)$$

and hence $c_1T_1 + c_2T_2 \in \mathrm{Hom}_G(\varphi, \rho)$, as required. □

Fundamental to all of representation theory is the important observation, due to I. Schur, that roughly speaking homomorphisms between irreducible representations are very limited. This is the first place that we seriously use that we are working over the field of complex numbers and not the field of real numbers. Namely, we use that every linear operator on a finite-dimensional complex vector space has an eigenvalue. This is a consequence of the fact that every polynomial over $\mathbb{C}$ has a root, in particular the characteristic polynomial of the operator has a root.

Lemma 4.1.6 (Schur's lemma)**.** *Let φ, ρ be irreducible representations of G, and $T \in \mathrm{Hom}_G(\varphi, \rho)$. Then either T is invertible or $T = 0$. Consequently:*

(a) If $\varphi \nsim \rho$, then $\mathrm{Hom}_G(\varphi, \rho) = 0$;

(b) If $\varphi = \rho$, then $T = \lambda I$ with $\lambda \in \mathbb{C}$ (i.e., T is a scalar matrix).

Proof. Let $\varphi\colon G \to GL(V)$, $\rho\colon G \to GL(W)$, and let $T\colon V \to W$ be in $\mathrm{Hom}_G(\varphi, \rho)$. If $T = 0$, we are done; so assume that $T \neq 0$. Proposition 4.1.4 implies that $\ker T$ is G-invariant and hence either $\ker T = V$ or $\ker T = 0$. Since $T \neq 0$, the former does not happen; thus $\ker T = 0$ and so T is injective. Also, according to Proposition 4.1.4, $\mathrm{Im}\, T$ is G-invariant, so $\mathrm{Im}\, T = W$ or $\mathrm{Im}\, T = 0$. If $\mathrm{Im}\, T = 0$ then again $T = 0$. So it must be $\mathrm{Im}\, T = W$, that is, T is surjective. We conclude that T is invertible.

For (a), assume $\mathrm{Hom}_G(\varphi, \rho) \neq 0$. That means there exists $T \neq 0$ in $\mathrm{Hom}_G(\varphi, \rho)$. Then T is invertible, by the above, and so $\varphi \sim \rho$. This is the contrapositive of what we wanted to show.

To establish (b), let λ be an eigenvalue of T (here is where we use that we are working over $\mathbb{C}$ and not $\mathbb{R}$). Then $\lambda I - T$ is not invertible by definition of an eigenvalue. Since $I \in \mathrm{Hom}_G(\varphi, \varphi)$, Proposition 4.1.5 tells us that $\lambda I - T$ belongs to $\mathrm{Hom}_G(\varphi, \varphi)$. Since all non-zero elements of $\mathrm{Hom}_G(\varphi, \varphi)$ are invertible by the first paragraph of the proof, it follows $\lambda I - T = 0$. Of course this is the same as saying $T = \lambda I$. □

Remark 4.1.7. It is not hard to deduce from Schur's lemma that if φ and ρ are equivalent irreducible representations, then $\dim \mathrm{Hom}_G(\varphi, \rho) = 1$.

We are now in a position to describe the irreducible representations of an abelian group.

Corollary 4.1.8. *Let G be an abelian group. Then any irreducible representation of G has degree one.*

Proof. Let $\varphi\colon G \to GL(V)$ be an irreducible representation. Fix for the moment $h \in G$. Then setting $T = \varphi_h$, we obtain, for all $g \in G$, that

$$T\varphi_g = \varphi_h\varphi_g = \varphi_{hg} = \varphi_{gh} = \varphi_g\varphi_h = \varphi_g T.$$

Consequently, Schur's lemma implies $\varphi_h = \lambda_h I$ for some scalar $\lambda_h \in \mathbb{C}$ (the subscript indicates the dependence on h). Let v be a non-zero vector in V and $k \in \mathbb{C}$. Then $\varphi_h(kv) = \lambda_h Ikv = \lambda_h kv \in \mathbb{C}v$. Thus $\mathbb{C}v$ is a G-invariant subspace, as h was arbitrary. We conclude that $V = \mathbb{C}v$ by irreducibility and so $\dim V = 1$. □

Let us present some applications of this result to linear algebra.

Corollary 4.1.9. *Let G be a finite abelian group and $\varphi\colon G \to GL_n(\mathbb{C})$ a representation. Then there is an invertible matrix T such that $T^{-1}\varphi_g T$ is diagonal for all $g \in G$ (T is independent of g).*

Proof. Since φ is completely reducible, we have that $\varphi \sim \varphi^{(1)} \oplus \cdots \oplus \varphi^{(m)}$ where $\varphi^{(1)}, \ldots, \varphi^{(m)}$ are irreducible. Since G is abelian, the degree of each $\varphi^{(i)}$ is 1 (and hence $n = m$). Consequently, $\varphi_g^{(i)} \in \mathbb{C}^*$ for all $g \in G$. Now if $T\colon \mathbb{C}^n \to \mathbb{C}^n$ gives the equivalence of φ with $\varphi^{(1)} \oplus \cdots \oplus \varphi^{(n)}$, then

$$T^{-1}\varphi_g T = \begin{bmatrix} \varphi_g^{(1)} & 0 & \cdots & 0 \\ 0 & \varphi_g^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \varphi_g^{(n)} \end{bmatrix}$$

is diagonal for all $g \in G$. □

As a corollary, we obtain the diagonalizability of matrices of finite order.

Corollary 4.1.10. *Let $A \in GL_m(\mathbb{C})$ be a matrix of finite order. Then A is diagonalizable. Moreover, if $A^n = I$, then the eigenvalues of A are n^{th}-roots of unity.*

Proof. Suppose $A^n = I$. Define a representation $\varphi\colon \mathbb{Z}_n \to GL_m(\mathbb{C})$ by setting $\varphi(\overline{k}) = A^k$. This is easily verified to give a well-defined representation since $A^n = I$. Thus there exists $T \in GL_n(\mathbb{C})$ such that $T^{-1}AT$ is diagonal by Corollary 4.1.9. Suppose

$$T^{-1}AT = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_m \end{bmatrix} = D.$$

Then

$$D^n = (T^{-1}AT)^n = T^{-1}A^nT = T^{-1}IT = I.$$

Therefore, we have

$$\begin{bmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_m^n \end{bmatrix} = D^n = I$$

and so $\lambda_i^n = 1$ for all i. This establishes that the eigenvalues of A are n^{th}-roots of unity. □

4.2 The orthogonality relations

From this point onwards, the group G shall always be assumed finite. Let $\varphi\colon G \to GL_n(\mathbb{C})$ be a representation. Then $\varphi_g = (\varphi_{ij}(g))$ where $\varphi_{ij}(g) \in \mathbb{C}$, $1 \leq i, j \leq n$. Thus there are n^2 functions $\varphi_{ij}\colon G \to \mathbb{C}$ associated to φ. What can be said about the functions φ_{ij} when φ is irreducible and unitary? It turns out that the functions of this sort form an orthogonal basis for $\mathbb{C}^G$.

Definition 4.2.1 (Group algebra)**.** Let G be a group and define

$$L(G) = \mathbb{C}^G = \{f \mid f\colon G \to \mathbb{C}\}.$$

Then $L(G)$ is an inner product space with addition and scalar multiplication given by

$$\begin{gathered} (f_1 + f_2)(g) = f_1(g) + f_2(g) \\ (cf)(g) = c \cdot f(g) \end{gathered}$$

and with the inner product defined by

$$\langle f_1, f_2 \rangle = \frac{1}{|G|} \sum_{g \in G} \overline{f_1(g)} f_2(g).$$

For reasons to become apparent later, $L(G)$ is called the *group algebra* of G.

One of our goals in this chapter is to prove the following important result. Recall that $U_n(\mathbb{C})$ is the group of $n \times n$ unitary matrices.

Theorem (Schur orthogonality relations)**.** *Suppose that* $\varphi\colon G \to U_n(\mathbb{C})$ *and* $\rho\colon G \to U_m(\mathbb{C})$ *are inequivalent irreducible unitary representations. Then:*

1. $\langle \rho_{k\ell}, \varphi_{ij} \rangle = 0$;

2. $\langle \varphi_{k\ell}, \varphi_{ij} \rangle = \begin{cases} 1/n & \text{if } i = k \text{ and } j = \ell \\ 0 & \text{else.} \end{cases}$

The proof requires a lot of preparation. We begin with our second usage of the "averaging trick."

Proposition 4.2.2. *Let $\varphi\colon G \to GL(V)$ and $\rho\colon G \to GL(W)$ be representations and suppose that $T\colon V \to W$ is a linear transformation. Then:*

(a) $T^\sharp = \dfrac{1}{|G|}\displaystyle\sum_{g\in G} \rho_{g^{-1}} T \varphi_g \in \mathrm{Hom}_G(\varphi, \rho)$

(b) If $T \in \mathrm{Hom}_G(\varphi, \rho)$, then $T^\sharp = T$.

(c) The map $P\colon \mathrm{Hom}(V, W) \to \mathrm{Hom}_G(\varphi, \rho)$ defined by $P(T) = T^\sharp$ is an onto linear map.

Proof. We verify (a) by a direct computation.

$$T^\sharp \varphi_h = \frac{1}{|G|}\sum_{g\in G} \rho_{g^{-1}} T \varphi_g \varphi_h = \frac{1}{|G|}\sum_{g\in G} \rho_{g^{-1}} T \varphi_{gh}. \tag{4.1}$$

The next step is to apply a change of variables $x = gh$. Since right multiplication by h is a permutation of G, as g varies over G, so does x. Noting that $g^{-1} = hx^{-1}$, we conclude that the right hand side of (4.1) is equal to

$$\frac{1}{|G|}\sum_{g\in G} \rho_{hx^{-1}} T \varphi_x = \frac{1}{|G|}\sum_{g\in G} \rho_h \rho_{x^{-1}} T \varphi_x = \rho_h \frac{1}{|G|}\sum_{x\in G} \rho_{x^{-1}} T \varphi_x = \rho_h T^\sharp.$$

This proves $T^\sharp \in \mathrm{Hom}_G(\varphi, \rho)$.

To prove (b), notice that if $T \in \mathrm{Hom}_G(\varphi, \rho)$, then

$$T^\sharp = \frac{1}{|G|}\sum_{g\in G} \rho_{g^{-1}} T \varphi_g = \frac{1}{|G|}\sum_{g\in G} \rho_{g^{-1}} \rho_g T = \frac{1}{|G|}\sum_{g\in G} T = \frac{1}{|G|}|G|T = T.$$

Finally, for (c) we establish linearity by checking

$$\begin{aligned} P(c_1T_1 + c_2T_2) &= (c_1T_1 + c_2T_2)^\sharp \\ &= \frac{1}{|G|}\sum_{g\in G} \rho_{g^{-1}} (c_1T_1 + c_2T_2) \varphi_g \\ &= c_1 \frac{1}{|G|}\sum_{g\in G} \rho_{g^{-1}} T_1 \varphi_g + c_2 \frac{1}{|G|}\sum_{g\in G} \rho_{g^{-1}} T_2 \varphi_g \\ &= c_1 T_1^\sharp + c_2 T_2^\sharp = c_1 P(T_1) + c_2 P(T_2). \end{aligned}$$

If $T \in \mathrm{Hom}_G(\varphi, \rho)$, then (b) implies $T = T^\sharp = P(T)$ and so P is onto. $\square$

The following variant of Schur's lemma will be the form in which we shall most commonly use it. It is based on the trivial observation that if I_n is the $n \times n$ identity matrix and $\lambda \in \mathbb{C}$, then $\mathrm{Tr}(\lambda I_n) = n\lambda$.

Proposition 4.2.3. *Let $\varphi\colon G \to GL(V)$, $\rho\colon G \to GL(W)$ be irreducible representations of G and let $T\colon V \to W$ be a linear map. Then:*

(a) If $\varphi \nsim \rho$, then $T^\sharp = 0$;

(b) If $\varphi = \rho$, then $T^\sharp = \dfrac{\mathrm{Tr}(T)}{\deg \varphi} I$.

Proof. Assume first $\varphi \nsim \rho$. Then $\mathrm{Hom}_G(\varphi, \rho) = 0$ by Schur's lemma and so $T^\sharp = 0$. Next suppose $\varphi = \rho$. By Schur's lemma, $T^\sharp = \lambda I$ some $\lambda \in \mathbb{C}$. Our goal is to solve for λ. As $T^\sharp\colon V \to V$, we have $\mathrm{Tr}(\lambda I) = \lambda \mathrm{Tr}(I) = \lambda \dim V = \lambda \deg \varphi$. It follows that $T^\sharp = \dfrac{\mathrm{Tr}(T^\sharp)}{\deg \varphi} I$.

On the other hand, we can also compute the trace directly from the definition of $T^\sharp$. Using $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, we obtain

$$\mathrm{Tr}(T^\sharp) = \frac{1}{|G|} \sum_{g \in G} \mathrm{Tr}(\varphi_{g^{-1}} T \varphi_g) = \frac{1}{|G|} \sum_{g \in G} \mathrm{Tr}(T) = \frac{|G|}{|G|} \mathrm{Tr}(T) = \mathrm{Tr}(T)$$

and so $T^\sharp = \dfrac{\mathrm{Tr}(T)}{\deg \varphi} I$, as required. $\square$

If $\varphi\colon G \to GL_n(\mathbb{C})$ and $\rho\colon G \to GL_m(\mathbb{C})$ are representations, then $\mathrm{Hom}(V, W) = M_{mn}(\mathbb{C})$ and $\mathrm{Hom}_G(\varphi, \rho)$ is a subspace of $M_{mn}(\mathbb{C})$. Hence the map P from Proposition 4.2.2 can be viewed as a linear transformation $P\colon M_{mn}(\mathbb{C}) \to M_{mn}(\mathbb{C})$. It would then be natural to compute the matrix of P with respect to the standard basis for $M_{mn}(\mathbb{C})$. It turns out that when φ and ρ are unitary representations, the matrix for P has a special form. Recall that the standard basis for $M_{mn}(\mathbb{C})$ consists of the matrices $E_{11}, E_{12}, \ldots, E_{mn}$ where E_{ij} is the $m \times n$-matrix with 1 in position ij and 0 elsewhere. One then has $(a_{ij}) = \sum_{ij} a_{ij} E_{ij}$.

The following lemma is a straightforward computation with the formula for matrix multiplication.

Lemma 4.2.4. *Let $A \in M_{rm}(\mathbb{C})$, $B \in M_{ns}(\mathbb{C})$ and $E_{ki} \in M_{mn}(\mathbb{C})$. Then the formula $(AE_{ki}B)_{\ell j} = a_{\ell k} b_{ij}$ holds where $A = (a_{ij})$ and $B = (b_{ij})$.*

Proof. By definition

$$(AE_{ki}B)_{\ell j} = \sum_{x,y} a_{\ell x}(E_{ki})_{xy}b_{yj}.$$

But all terms in this sum are 0, except when $x = k$, $y = i$, in which case one gets $a_{\ell k}b_{ij}$, as desired. □

Example 4.2.5. This example illustrates Lemma 4.2.4:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 0 & a_{11} \\ 0 & a_{21} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{21} & a_{11}b_{22} \\ a_{21}b_{21} & a_{21}b_{22} \end{bmatrix}.$$

Now we are prepared to compute the matrix of P with respect to the standard basis. We state the result in the form in which we shall use it.

Lemma 4.2.6. *Let $\varphi\colon G \to U_n(\mathbb{C})$ and $\rho\colon G \to U_m(\mathbb{C})$ be unitary representations. Let $A = E_{ki} \in M_{mn}(\mathbb{C})$. Then $A^\sharp_{\ell j} = \langle \rho_{k\ell}, \varphi_{ij} \rangle$.*

Proof. Since ρ is unitary, $\rho_{g^{-1}} = \rho_g^{-1} = \rho_g^*$. Thus $\rho_{\ell k}(g^{-1}) = \overline{\rho_{k\ell}(g)}$. Keeping this in mind, we compute

$$\begin{aligned} A^\sharp_{\ell j} &= \frac{1}{|G|} \sum_{g \in G} (\rho_{g^{-1}} E_{ki} \varphi_g)_{\ell j} \\ &= \frac{1}{|G|} \sum_{g \in G} \rho_{\ell k}(g^{-1}) \varphi_{ij}(g) \qquad \text{by Lemma 4.2.4} \\ &= \frac{1}{|G|} \sum_{g \in G} \overline{\rho_{k\ell}(g)} \varphi_{ij}(g) \\ &= \langle \rho_{k\ell}, \varphi_{ij} \rangle \end{aligned}$$

as required. □

Remark 4.2.7. Let $P\colon M_{mn}(\mathbb{C}) \to M_{mn}(\mathbb{C})$ be the linear transformation given by $P(T) = T^\sharp$ and let B be the matrix of P. Then B is an $mn \times mn$ matrix whose rows and columns are indexed by pairs $\ell j, ki$ where $1 \leq \ell, k \leq m$ and $1 \leq j, i \leq n$. The content of Lemma 4.2.6 is that the $\ell j, ki$ entry of B is the inner product $\langle \rho_{k\ell}, \varphi_{ij} \rangle$.

We can now prove the Schur orthogonality relations.

Theorem 4.2.8 (Schur orthogonality relations)**.** *Let $\varphi\colon G \to U_n(\mathbb{C})$ and $\rho\colon G \to U_m(\mathbb{C})$ be inequivalent irreducible unitary representations. Then:*

1. $\langle \rho_{k\ell}, \varphi_{ij} \rangle = 0$;

2. $\langle \varphi_{k\ell}, \varphi_{ij} \rangle = \begin{cases} 1/n & \text{if } i = k \text{ and } j = \ell \\ 0 & \text{else.} \end{cases}$

Proof. For 1, let $A = E_{ki} \in M_{mn}(\mathbb{C})$. Then $A^\sharp = 0$ by Proposition 4.2.3. On the other hand, $A^\sharp_{\ell j} = \langle \rho_{k\ell}, \varphi_{ij} \rangle$ by Lemma 4.2.6. This establishes 1.

Next, we apply Proposition 4.2.3 and Lemma 4.2.6 with $\varphi = \rho$. Let $A = E_{ki} \in M_n(\mathbb{C})$. Then

$$A^\sharp = \frac{\operatorname{Tr}(E_{ki})}{n} I$$

by Proposition 4.2.3. Lemma 4.2.6 shows that $A^\sharp_{\ell j} = \langle \varphi_{k\ell}, \varphi_{ij} \rangle$. First suppose that $j \neq \ell$. Then since $I_{\ell j} = 0$, it follows $0 = A^\sharp_{\ell j} = \langle \varphi_{k\ell}, \varphi_{ij} \rangle$. Next suppose that $i \neq k$. Then E_{ki} has only zeroes on the diagonal and so $\operatorname{Tr}(E_{ki}) = 0$. Thus we again have $0 = A^\sharp_{\ell j} = \langle \varphi_{k\ell}, \varphi_{ij} \rangle$. Finally, in the case where $\ell = j$ and $i = k$, E_{ki} has a single 1 on the diagonal and all other entries are 0. Thus $\operatorname{Tr}(E_{ki}) = 1$ and so $1/n = A^\sharp_{\ell j} = \langle \varphi_{k\ell}, \varphi_{ij} \rangle$. This proves the theorem. □

A simple renormalization establishes:

Corollary 4.2.9. *Let φ be an irreducible unitary representation of G of degree d. Then the d^2 functions $\{\sqrt{d}\varphi_{ij} \mid 1 \leq i, j \leq d\}$ form an orthonormal set.*

An important corollary of Theorem 4.2.8 is that there are only finitely many equivalence classes of irreducible representations of G. First recall every equivalence class contains a unitary representation. Next, because $\dim L(G) = |G|$, no linearly independent set of vectors from $L(G)$ can have more than $|G|$ elements. Theorem 4.2.8 says that the entries of inequivalent unitary representations of G form an orthogonal set of non-zero vectors in $L(G)$. It follows that G has at most $|G|$ equivalence classes of irreducible representations. In fact, if $\varphi^{(1)}, \ldots, \varphi^{(s)}$ are a complete set of representatives of the equivalence classes of irreducible representations of G and $d_i = \deg \varphi^{(i)}$, then the $d_1^2 + d_2^2 + \cdots + d_s^2$ functions $\{\sqrt{d_k}\varphi^{(k)}_{ij} \mid 1 \leq k \leq s, 1 \leq i, j \leq d_k\}$ form an orthonormal set of vectors in $L(G)$ and hence $s \leq d_1^2 + \cdots + d_s^2 \leq |G|$ (the first inequality holds since $d_i \geq 1$ all i). We summarize this discussion in the following proposition.

Proposition 4.2.10. *Let G be a finite group. Let $\varphi^{(1)}, \ldots, \varphi^{(s)}$ be a complete set of representatives of the equivalence classes of irreducible representations of G and set $d_i = \deg \varphi^{(i)}$. Then the functions*

$$\{\sqrt{d_k}\varphi_{ij}^{(k)} \mid 1 \leq k \leq s, 1 \leq i, j \leq d_k\}$$

form an orthonormal set in $L(G)$ and hence $s \leq d_1^2 + \cdots + d_s^2 \leq |G|$.

Later, we shall see that the second inequality in the proposition is in fact an equality; the first one is only an equality for abelian groups.

4.3 Characters and class functions

In this section, we finally prove the uniqueness of the decomposition of a representation into irreducible representations. The key ingredient is to associate to each representation φ a function $\chi_\varphi \colon G \to \mathbb{C}$ which encodes the entire representation.

Definition 4.3.1 (Character)**.** Let $\varphi \colon G \to GL(V)$ be a representation. The *character* $\chi_\varphi \colon G \to \mathbb{C}$ of φ is defined by setting $\chi_\varphi(g) = \mathrm{Tr}(\varphi_g)$. The character of an irreducible representation is called an *irreducible character*.

So if $\varphi \colon G \to GL_n(\mathbb{C})$ is a representation given by $\varphi_g = (\varphi_{ij}(g))$, then

$$\chi_\varphi(g) = \sum_{i=1}^{n} \varphi_{ii}(g).$$

In general, to compute the character one must choose a basis and so when talking about characters, we may assume without loss of generality that we are talking about matrix representations.

Remark 4.3.2. If $\varphi \colon G \to \mathbb{C}^*$ is a degree 1 representation, then $\chi_\varphi = \varphi$. From now on, we will not distinguish between a degree 1 representation and its character.

The first piece of information that we shall read off the character is the degree of the representation.

Proposition 4.3.3. *Let φ be a representation of G. Then $\chi_\varphi(1) = \deg \varphi$.*

Proof. Indeed, suppose that $\varphi \colon G \to GL(V)$ is a representation. Then $\mathrm{Tr}(\varphi_1) = \mathrm{Tr}(I) = \dim V = \deg \varphi$. □

A key property of the character is that it depends only on the equivalence class of the representation.

Proposition 4.3.4. *If φ and ρ are equivalent representations, then $\chi_\varphi = \chi_\rho$.*

Proof. Since the trace is computed by selecting a basis, we are able to assume that $\varphi, \rho \colon G \to GL_n(\mathbb{C})$. Then, since they are equivalent, there is an invertible matrix $T \in GL_n(\mathbb{C})$ such that $\varphi_g = T\rho_g T^{-1}$, for all $g \in G$. Recalling $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, we obtain

$$\chi_\varphi(g) = \mathrm{Tr}(\varphi_g) = \mathrm{Tr}(T\rho_g T^{-1}) = \mathrm{Tr}(T^{-1}T\rho_g) = \mathrm{Tr}(\rho_g) = \chi_\rho(g)$$

as required. □

The same proof illuminates another crucial property of characters: they are constant on conjugacy classes.

Proposition 4.3.5. *Let φ be a representation of G. Then, for all $g, h \in G$, the equality $\chi_\varphi(g) = \chi_\varphi(hgh^{-1})$ holds.*

Proof. Indeed, we compute

$$\begin{aligned}\chi_\varphi(hgh^{-1}) &= \mathrm{Tr}(\varphi_{hgh^{-1}}) = \mathrm{Tr}(\varphi_h\varphi_g\varphi_h^{-1}) \\ &= \mathrm{Tr}(\varphi_h^{-1}\varphi_h\varphi_g) = \mathrm{Tr}(\varphi_g) = \chi_\varphi(g)\end{aligned}$$

again using $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. □

Functions which are constant on conjugacy classes play an important role in representation theory and hence deserve a name of their own.

Definition 4.3.6 (Class function)**.** A function $f \colon G \to \mathbb{C}$ is called a *class function* if $f(g) = f(hgh^{-1})$ for all $g, h \in G$, or equivalently, if f is constant on conjugacy classes of G. The space of class functions is denoted $Z(L(G))$.

In particular, characters are class functions. The notation $Z(L(G))$ suggests that the class functions should be the center of some ring, and this will indeed be the case. If $f \colon G \to \mathbb{C}$ is a class function and C is a conjugacy class, $f(C)$ will denote the constant value that f takes on C.

Proposition 4.3.7. *$Z(L(G))$ is a subspace of $L(G)$.*

Proof. Let f_1, f_2 be class functions on G and let $c_1, c_2 \in \mathbb{C}$. Then

$$\begin{aligned}(c_1f_1 + c_2f_2)(hgh^{-1}) &= c_1f_1(hgh^{-1}) + c_2f_2(hgh^{-1}) \\ &= c_1f_1(g) + c_2f_2(g) = (c_1f_1 + c_2f_2)(g)\end{aligned}$$

showing that $c_1f_1 + c_2f_2$ is a class function. □

Next, let's compute the dimension of $Z(L(G))$. Let $Cl(G)$ be the set of conjugacy classes of G. Define, for $C \in Cl(G)$, the function $\delta_C \colon G \to \mathbb{C}$ by

$$\delta_C(g) = \begin{cases} 1 & g \in C \\ 0 & g \notin C. \end{cases}$$

Proposition 4.3.8. *The set* $B = \{\delta_C \mid C \in Cl(G)\}$ *is a basis for* $Z(L(G))$*. Consequently* $\dim Z(L(G)) = |Cl(G)|$.

Proof. Clearly δ_C is constant on conjugacy classes, and hence is a class function. Let us begin by showing that B spans $Z(L(G))$. If $f \in Z(L(G))$, then one easily verifies that

$$f = \sum_{C \in Cl(G)} f(C)\delta_C.$$

Indeed, if C' is the conjugacy class of g, then when you evaluate the right hand side at g you get $f(C')$. Since $g \in C'$, by definition $f(C') = f(g)$. To establish linear independence, we verify that B is an orthogonal set of non-zero vectors. For if $C, C' \in Cl(G)$, then

$$\frac{1}{|G|}\sum_{g \in G} \overline{\delta_C(g)}\delta_{C'}(g) = \begin{cases} \frac{|C|}{|G|} & C = C' \\ 0 & C \neq C'. \end{cases}$$

This completes the proof that B is a basis. Since $|B| = |Cl(G)|$, the calculation of the dimension follows. □

The next theorem is one of the fundamental results in group representation theory. It shows that the irreducible characters form an orthonormal set of class functions. This will be used to establish the uniqueness of the decomposition of a representation into irreducible constituents and to obtain a better bound on the number of equivalence classes of irreducible representations.

Theorem 4.3.9 (First orthogonality relations)**.** *Let φ, ρ be irreducible representations of G. Then*

$$\langle \chi_\varphi, \chi_\rho \rangle = \begin{cases} 1 & \varphi \sim \rho \\ 0 & \varphi \nsim \rho. \end{cases}$$

Thus the irreducible characters of G form an orthonormal set of class functions.

Proof. Thanks to Propositions 3.2.4 and 4.3.4, we may assume without loss of generality that $\varphi\colon G \to U_n(\mathbb{C})$ and $\rho\colon G \to U_m(\mathbb{C})$ are unitary. Next we compute

$$\begin{aligned}
\langle \chi_\varphi, \chi_\rho \rangle &= \frac{1}{|G|} \sum_{g\in G} \overline{\chi_\varphi(g)} \chi_\rho(g) \\
&= \frac{1}{|G|} \sum_{g\in G} \sum_{i=1}^{n} \overline{\varphi_{ii}(g)} \sum_{j=1}^{m} \rho_{jj}(G) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{1}{|G|} \sum_{g\in G} \overline{\varphi_{ii}(g)} \rho_{jj}(G) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} \langle \varphi_{ii}(g), \rho_{jj}(g) \rangle.
\end{aligned}$$

The Schur orthogonality relations (Theorem 4.2.8) yield $\langle \varphi_{ii}(g), \rho_{jj}(g) \rangle = 0$ if $\varphi \nsim \rho$ and so $\langle \chi_\varphi, \chi_\rho \rangle = 0$ if $\varphi \nsim \rho$. If $\varphi \sim \rho$, then we may assume $\varphi = \rho$ by Proposition 4.3.4. In this case, the Schur orthogonality relations tell us

$$\langle \varphi_{ii}, \varphi_{jj} \rangle = \begin{cases} 1/n & i = j \\ 0 & i \neq j \end{cases}$$

and so

$$\langle \chi_\varphi, \chi_\varphi \rangle = \sum_{i=1}^{n} \langle \varphi_{ii}, \varphi_{ii} \rangle = \sum_{i=1}^{n} \frac{1}{n} = 1$$

as required. □

Corollary 4.3.10. *There are at most $|Cl(G)|$ equivalence classes of irreducible representations of G.*

Proof. First note that Theorem 4.3.9 implies inequivalent irreducible representations have distinct characters and, moreover, the irreducible characters form an orthonormal set. Since $\dim Z(L(G)) = |Cl(G)|$ and orthonormal sets are linearly independent, the corollary follows. □

Let us introduce some notation. If V is a vector space, φ is a representation and $m > 0$, then we set

$$mV = V \oplus \overbrace{\cdots}^{\times m} \oplus V \text{ and } m\varphi = \varphi \oplus \overbrace{\cdots}^{\times m} \oplus \varphi.$$

Let $\varphi^{(1)}, \ldots, \varphi^{(s)}$ be a complete set of irreducible unitary representations of G, up to equivalence. Again, set $d_i = \deg \varphi^{(i)}$.

Definition 4.3.11 (Multiplicity)**.** If $\rho \sim m_1\varphi^{(1)} \oplus m_2\varphi^{(2)} \oplus \cdots \oplus m_s\varphi^{(s)}$, then m_i is called the *multiplicity* of $\varphi^{(i)}$ in ρ. If $m_i > 0$, then we say that $\varphi^{(i)}$ is an *irreducible constituent* of ρ.

It is not clear at the moment that the multiplicity is well defined because we have not yet established the uniqueness of the decomposition of a representation into irreducibles. To show that it is well defined, we come up with a way to compute m_i directly from the character of ρ. Since the character only depends on the equivalence class, it follows that the multiplicity of $\varphi^{(i)}$ will be the same no matter how we decompose ρ.

Remark 4.3.12. If $\rho \sim m_1\varphi^{(1)} \oplus m_2\varphi^{(2)} \oplus \cdots \oplus m_s\varphi^{(s)}$, then

$$\deg \rho = m_1d_1 + m_2d_2 + \cdots + m_sd_s.$$

Lemma 4.3.13. *Let $\varphi = \rho \oplus \psi$. Then $\chi_\varphi = \chi_\rho + \chi_\psi$.*

Proof. We may assume that $\rho\colon G \to GL_n(\mathbb{C})$ and $\psi\colon G \to GL_m(\mathbb{C})$. Then $\varphi\colon G \to GL_{n+m}(\mathbb{C})$ has block form

$$\varphi_g = \begin{bmatrix} \rho_g & 0 \\ 0 & \psi_g \end{bmatrix}.$$

Since the trace is the sum of the diagonal elements, it follows that

$$\chi_\varphi(g) = \mathrm{Tr}(\varphi_g) = \mathrm{Tr}(\rho_g) + \mathrm{Tr}(\psi_g) = \chi_\rho(g) + \chi_\psi(g).$$

We conclude that $\chi_\varphi = \chi_\rho + \chi_\psi$. □

The above lemma implies that each character is an integral linear combination of irreducible characters. We can then use the orthonormality of the irreducible characters to extract the coefficients.

Theorem 4.3.14. *Let $\varphi^{(1)}, \ldots, \varphi^{(s)}$ be a complete set of representatives of the equivalence classes of irreducible representations of G and let*

$$\rho \sim m_1\varphi^{(1)} \oplus m_2\varphi^{(2)} \oplus \cdots \oplus m_s\varphi^{(s)}.$$

Then $m_i = \langle \chi_{\varphi^{(i)}}, \chi_\rho \rangle$. Consequently, the decomposition of ρ into irreducible constituents is unique and ρ is determined up to equivalence by its character.

Proof. By the previous lemma, $\chi_\rho = m_1\chi_{\varphi^{(1)}} + \cdots + m_s\chi_{\varphi^{(s)}}$. By the first orthogonality relations

$$\langle \chi_{\varphi^{(i)}}, \chi_\rho \rangle = m_1\langle \chi_{\varphi^{(i)}}, \chi_{\varphi^{(1)}} \rangle + \cdots + m_s\langle \chi_{\varphi^{(i)}}, \chi_{\varphi^{(s)}} \rangle = m_i,$$

proving the first statement. Proposition 4.3.4 implies the second and third statements. □

Theorem 4.3.14 offers a convenient criterion to check whether a representation is irreducible.

Corollary 4.3.15. *A representation ρ is irreducible if and only if $\langle \chi_\rho, \chi_\rho \rangle = 1$.*

Proof. Suppose $\rho \sim m_1\varphi^{(1)} \oplus m_2\varphi^{(2)} \oplus \cdots \oplus m_s\varphi^{(s)}$. Using the orthonormality of the irreducible characters, we obtain $\langle \chi_\rho, \chi_\rho \rangle = m_1^2 + \cdots + m_s^2$. The m_i are non-negative integers, so $\langle \chi_\rho, \chi_\rho \rangle = 1$ if and only if there is an index j so that $m_j = 1$ and $m_i = 0$ for $i \neq j$. But this happens precisely if ρ is irreducible. □

Let's use Corollary 4.3.15 to show that the representation from Example 3.1.8 is irreducible.

Example 4.3.16. Let ρ be the representation of S_3 from Example 3.1.8. Since Id, $(1\ 2)$ and $(1\ 2\ 3)$ form a complete set of representatives of the conjugacy classes of S_3, we can compute the inner product $\langle \chi_\rho, \chi_\rho \rangle$ from the values of the character on these elements. Now $\chi_\rho(Id) = 2$, $\chi_\rho((1\ 2)) = 0$ and $\chi_\rho((1\ 2\ 3)) = -1$. Since there are 3 transpositions and 2 three-cycles, we have

$$\langle \chi_\rho, \chi_\rho \rangle = \frac{1}{6}\left(2^2 + 3 \cdot 0^2 + 2 \cdot (-1)^2\right) = 1$$

and so ρ is irreducible.

Let us try to find all the irreducible characters of S_3 and to decompose the standard representation (c.f. Example 3.1.4).

Example 4.3.17 (Characters of S_3)**.** We know that S_3 admits the trivial character $\chi_1 \colon S_3 \to \mathbb{C}^*$ given by $\chi_1(\sigma) = 1$ for all $\sigma \in S_3$ (recall we identify a degree one representation with its character). We also have the character χ_3 of the irreducible representation from Example 3.1.8. Since S_3 has 3 conjugacy classes, we might hope that there are 3 inequivalent irreducible representations of S_3. From Proposition 4.2.10, we know that if d is the degree of the missing representation, then $1^2 + d^2 + 2^2 \leq 6$ and so $d = 1$. In fact, we can define a second degree one representation by

$$\chi_2(\sigma) = \begin{cases} 1 & \sigma \text{ is even} \\ -1 & \sigma \text{ is odd} \end{cases}$$

Let us form a table encoding this information (such a table is called a *character table*). The rows of Table 4.1 correspond to the irreducible characters, whereas the columns correspond to the conjugacy classes.

	Id	$(1\ 2)$	$(1\ 2\ 3)$
χ_1	1	1	1
χ_2	1	-1	1
χ_3	2	0	-1

Table 4.1: Character table of S_3

The standard representation of S_3 from Example 3.1.4 is given by the matrices

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

Hence we have character values

	Id	$(1\ 2)$	$(1\ 2\ 3)$
χ_φ	3	1	0

Inspection of Table 4.1 shows that $\chi_\varphi = \chi_1 + \chi_3$ and hence $\varphi \sim \chi_1 \oplus \rho$, as was advertised in Example 3.1.8. Alternatively, one could use Theorem 4.3.14 to obtain this result. Indeed,

$$\begin{aligned} \langle \chi_1, \chi_\varphi \rangle &= \frac{1}{6}(3 + 3 \cdot 1 + 2 \cdot 0) = 1 \\ \langle \chi_2, \chi_\varphi \rangle &= \frac{1}{6}(3 + 3 \cdot (-1) + 2 \cdot 0) = 0 \\ \langle \chi_3, \chi_\varphi \rangle &= \frac{1}{6}(6 + 3 \cdot 0 + 2 \cdot 0) = 1. \end{aligned}$$

We will study the character table in detail later, in particular we shall show that the columns are always pairwise orthogonal, as is the case in Table 4.1.

4.4 The regular representation

Cayley's theorem asserts that G is isomorphic to a subgroup of S_n where $n = |G|$. The standard representation from Example 3.1.4 provides a representation $\varphi\colon S_n \to GL_n(\mathbb{C})$. The restriction of this representation to G will be called the regular representation of G, although we will construct it in a different way.

Let X be a finite set. We build synthetically a vector space with basis X by setting

$$\mathbb{C}X = \left\{ \sum_{x\in X} c_x x \mid c_x \in \mathbb{C} \right\}.$$

So $\mathbb{C}X$ consists of all formal linear combinations of elements of X. Two elements $\sum_{x\in X} a_x x$ and $\sum_{x\in X} b_x x$ are declared to be equal if and only if $a_x = b_x$ all $x \in X$. Addition is given by

$$\sum_{x\in X} a_x x + \sum_{x\in X} b_x x = \sum_{x\in X} (a_x + b_x)x;$$

scalar multiplication is defined similarly. We identify $x \in X$ with the linear combination $1 \cdot x$. Clearly X is a basis for $\mathbb{C}X$. An inner product can be defined on $\mathbb{C}X$ by setting

$$\left\langle \sum_{x\in X} a_x x, \sum_{x\in X} b_x x \right\rangle = \sum_{x\in X} \overline{a_x} b_x.$$

Definition 4.4.1 (Regular representation)**.** Let G be a finite group. The *regular representation* of G is the homomorphism $L\colon G \to GL(\mathbb{C}G)$ defined by

$$L_g \sum_{h\in G} c_h h = \sum_{h\in G} c_h gh = \sum_{x\in G} c_{g^{-1}x} x, \tag{4.2}$$

for $g \in G$ (where the last equality comes from the change of variables $x = gh$).

The L stands for "left." Notice that on a basis element $h \in G$, we have $L_g h = gh$, i.e., L_g acts on the basis via left multiplication by g. The formula in (4.2) is then the usual formula for a linear operator acting on a linear combination of basis vectors given the action on the basis. It follows that L_g is a linear map for all $g \in G$. The regular representation is never irreducible when G is non-trivial, but it has the positive feature that it contains all the irreducible representations of G as constituents. Let us first prove that it is a representation.

Proposition 4.4.2. *The regular representation is a unitary representation of G.*

Proof. We already pointed out the map L_g is linear for $g \in G$. Also if $g_1, g_2 \in G$ and $h \in G$ is a basis element of $\mathbb{C}G$, then

$$L_{g_1} L_{g_2} h = L_{g_1} g_2 h = g_1 g_2 h = L_{g_1 g_2} h$$

so L is a homomorphism. If we show that L_g is unitary, it will then follow L_g is invertible and that L is a unitary representation. Now by (4.2)

$$\left\langle L_g \sum_{h\in G} c_h h, L_g \sum_{h\in G} k_h h \right\rangle = \left\langle \sum_{x\in G} c_{g^{-1}x} x, \sum_{x\in G} k_{g^{-1}x} x \right\rangle = \sum_{x\in G} \overline{c_{g^{-1}x}} k_{g^{-1}x}. \tag{4.3}$$

Setting $y = g^{-1}x$ turns the right hand side of (4.3) into

$$\sum_{y\in G} \overline{c_y} k_y = \left\langle \sum_{y\in G} c_y y, \sum_{y\in G} k_y y \right\rangle$$

establishing that L_g is unitary. □

Let's next compute the character of L. It turns out to have a particularly simple form.

Proposition 4.4.3. *The character of the regular representation L is given by*

$$\chi_L(g) = \begin{cases} |G| & g = 1 \\ 0 & g \neq 1. \end{cases}$$

Proof. Let $G = \{g_1, \ldots, g_n\}$ where $n = |G|$. Then $L_g g_j = g g_j$. Thus if $[L_g]$ is the matrix of L_g with respect to the basis G with this ordering, then

$$\begin{aligned} [L_g]_{ij} &= \begin{cases} 1 & g_i = g g_j \\ 0 & \text{else} \end{cases} \\ &= \begin{cases} 1 & g = g_i g_j^{-1} \\ 0 & \text{else.} \end{cases} \end{aligned}$$

In particular,

$$[L_g]_{ii} = \begin{cases} 1 & g = 1 \\ 0 & \text{else} \end{cases}$$

from which we conclude

$$\chi_L(g) = \mathrm{Tr}(L_g) = \begin{cases} |G| & g = 1 \\ 0 & g \neq 1 \end{cases}$$

as required. □

We now decompose the regular representation L into irreducible constituents. Fix again a complete set $\{\varphi^{(1)}, \ldots, \varphi^{(s)}\}$ of inequivalent irreducible unitary representations of our finite group G and set $d_i = \deg \varphi^{(i)}$. For convenience, we put $\chi_i = \chi_{\varphi^{(i)}}$, for $i = 1, \ldots, s$.

Theorem 4.4.4. *Let L be the regular representation of G. Then the decomposition*

$$L \sim d_1\varphi^{(1)} \oplus d_2\varphi^{(2)} \oplus \cdots \oplus d_s\varphi^{(s)}$$

holds.

Proof. We compute

$$\begin{aligned} \langle \chi_i, \chi_L \rangle &= \frac{1}{|G|} \sum_{g \in G} \overline{\chi_i(g)} \chi_L(g) \\ &= \frac{1}{|G|} \overline{\chi_i(1)} |G| \\ &= \deg \varphi^{(i)} \\ &= d_i \end{aligned}$$

since $\chi_L(g) = 0$ for $g \neq 1$ and $\chi_L(1) = |G|$. This finishes the proof thanks to Theorem 4.3.14. □

With this theorem, we may complete the line of investigation initiated in this chapter.

Corollary 4.4.5. *The formula $|G| = d_1^2 + d_2^2 + \cdots + d_s^2$ holds.*

Proof. Since $\chi_L = d_1\chi_1 + d_2\chi_2 + \cdots + d_s\chi_s$ by Theorem 4.4.4, evaluating at 1 yields

$$|G| = \chi_L(1) = d_1\chi_1(1) + \cdots + d_s\chi_s(1) = d_1^2 + \cdots + d_s^2$$

as required. □

Consequently, we may infer that the matrix coefficients of irreducible unitary representations form an orthogonal basis for the space of all functions on G.

Theorem 4.4.6. *The set* $B = \{\sqrt{d_k}\varphi_{ij}^{(k)} \mid 1 \leq k \leq s, 1 \leq i, j \leq d_k\}$ *is an orthonormal basis for* $L(G)$.

Proof. We already know B is an orthonormal set by the orthogonality relations (Theorem 4.2.8). Since $|B| = d_1^2 + \cdots + d_s^2 = |G| = \dim L(G)$, it follows B is a basis. □

Next we show that $\chi_1, \ldots, \chi_s$ is an orthonormal basis for the space of class functions.

Theorem 4.4.7. *The set* $\chi_1, \ldots, \chi_s$ *is an orthonormal basis for* $Z(L(G))$.

Proof. The first orthogonality relations (Theorem 4.3.9) tell us that the irreducible characters form an orthonormal set of class functions. We must show that they span $Z(L(G))$. Let $f \in Z(L(G))$. By the previous theorem,

$$f = \sum_{i,j,k} c_{ij}^{(k)} \varphi_{ij}^{(k)}$$

for some $c_{ij}^{(k)} \in \mathbb{C}$ where $1 \leq k \leq s$ and $1 \leq i, j \leq d_k$. Since f is a class function, for any $x \in G$, we have

$$\begin{aligned}
f(x) &= \frac{1}{|G|} \sum_{g \in G} f(g^{-1}xg) \\
&= \frac{1}{|G|} \sum_{g \in G} \sum_{i,j,k} c_{ij}^{(k)} \varphi_{ij}^{(k)}(g^{-1}xg) \\
&= \sum_{i,j,k} c_{ij}^{(k)} \frac{1}{|G|} \sum_{g \in G} \varphi_{ij}^{(k)}(g^{-1}xg) \\
&= \sum_{i,j,k} c_{ij}^{(k)} \left[\frac{1}{|G|} \sum_{g \in G} \varphi_{g^{-1}}^{(k)} \varphi_x^{(k)} \varphi_g^{(k)} \right]_{ij} \\
&= \sum_{i,j,k} c_{ij}^{(k)} [(\varphi_x^{(k)})^{\sharp}]_{ij} \\
&= \sum_{i,j,k} c_{ij}^{(k)} \frac{\mathrm{Tr}(\varphi_x^{(k)})}{\deg \varphi^{(k)}} I_{ij} \\
&= \sum_{i,k} c_{ii}^{(k)} \frac{1}{d_k} \chi_k(x).
\end{aligned}$$

This establishes that

$$f = \sum_{i,k} c_{ii}^{(k)} \frac{1}{d_k} \chi_k$$

is in the span of $\chi_1, \ldots, \chi_s$, completing the proof that the irreducible characters form an orthonormal basis for $Z(L(G))$. □

Corollary 4.4.8. *The number of equivalence classes of irreducible representations of G is the number of conjugacy classes of G.*

Proof. The above theorem implies $s = \dim Z(L(G)) = |Cl(G)|$. □

Corollary 4.4.9. *A finite group G is abelian if and only if it has $|G|$ equivalence classes of irreducible representations.*

Proof. A finite group G is abelian if and only if $|G| = |Cl(G)|$. □

Example 4.4.10 (Irreducible representations of $\mathbb{Z}_n$)**.** Let $\omega = e^{2\pi i/n}$. Define $\chi_k \colon \mathbb{Z}_n \to \mathbb{C}^*$ by $\chi_k(\overline{m}) = \omega^{km}$ for $0 \leq k \leq n-1$. Then $\chi_0, \ldots, \chi_{n-1}$ are the distinct irreducible representations of $\mathbb{Z}_n$.

The representation theoretic information about a finite group G can be encoded in a matrix known as its character table.

Definition 4.4.11 (Character table)**.** Let G be a finite group with irreducible characters $\chi_1, \ldots, \chi_s$ and conjugacy classes $C_1, \ldots, C_s$. The *character table* of G is the $s \times s$ matrix X with $\mathsf{X}_{ij} = \chi_i(C_j)$. In other words, the rows of X are indexed by the characters of G, the columns by the conjugacy classes of G and the ij-entry is the value of the i^{th}-character on the j^{th}-conjugacy class.

The character table of S_3 is recorded in Table 4.1, while that of $\mathbb{Z}_4$ can be found in Table 4.2.

	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
χ_1	1	1	1	1
χ_2	1	-1	1	-1
χ_3	1	i	-1	$-i$
χ_4	1	$-i$	-1	i

Table 4.2: Character table of $\mathbb{Z}_4$

Notice that in both examples the columns are orthogonal with respect to the standard inner product. Let's prove that this is always the case.

If $g, h \in G$, then the inner product of the columns corresponding to their conjugacy classes is $\displaystyle\sum_{i=1}^{s} \overline{\chi_i(g)}\chi_i(h)$.

Recall that if C is a conjugacy class, then

$$\delta_C(g) = \begin{cases} 1 & g \in C \\ 0 & \text{else.} \end{cases}$$

The δ_C with $C \in Cl(G)$ form a basis for $Z(L(G))$, as do the irreducible characters. It is natural to express the δ_C in terms of the irreducible characters. This will yield the orthogonality of the columns of the character table.

Theorem 4.4.12 (Second orthogonality relations)**.** *Let C, C' be conjugacy classes of G and let $g \in C$ and $h \in C'$. Then*

$$\sum_{i=1}^{s} \overline{\chi_i(g)}\chi_i(h) = \begin{cases} |G|/|C| & C = C' \\ 0 & C \neq C'. \end{cases}$$

Consequently, the columns of the character table are orthogonal and hence the character table is invertible.

Proof. Using that $\delta_C = \sum_{i=1}^{s} \langle \chi_i, \delta_C \rangle \chi_i$, we compute

$$\begin{aligned} \delta_C(h) &= \sum_{i=1}^{s} \langle \chi_i, \delta_C \rangle \chi_i(h) \\ &= \sum_{i=1}^{s} \frac{1}{|G|} \sum_{x \in G} \overline{\chi_i(x)}\delta_C(x)\chi_i(h) \\ &= \sum_{i=1}^{s} \frac{1}{|G|} \sum_{x \in C} \overline{\chi_i(x)}\chi_i(h) \\ &= \frac{|C|}{|G|} \sum_{i=1}^{s} \overline{\chi_i(g)}\chi_i(h). \end{aligned}$$

Since the left hand side is 1 when $h \in C$ and 0 otherwise, we conclude

$$\sum_{i=1}^{s} \overline{\chi_i(g)}\chi_i(h) = \begin{cases} |G|/|C| & C = C' \\ 0 & C \neq C' \end{cases}$$

as was required.

It now follows that the columns of the character table form an orthogonal set of non-zero vectors and hence are linearly independent. This yields the invertibility of the character table. □

4.5 Representations of abelian groups

In this section, we compute the characters of an abelian group. Example 4.4.10 provides the characters of the group $\mathbb{Z}_n$. Since any finite abelian group is a direct product of cyclic groups, all we need to know is how to compute the characters of a direct product of abelian groups. Let us proceed to the task at hand!

Proposition 4.5.1. *Let G_1, G_2 be abelian groups. Let $\chi_1, \ldots, \chi_m$ and $\varphi_1, \ldots, \varphi_n$ be the irreducible representations of G_1, G_2, respectively. In particular, $m = |G_1|$ and $n = |G_2|$. Then the functions $\alpha_{ij} \colon G_1 \times G_2 \to \mathbb{C}^*$ with $1 \leq i \leq m$, $1 \leq j \leq n$ given by*

$$\alpha_{ij}(g_1, g_2) = \chi_i(g_1)\varphi_j(g_2)$$

form a complete set of irreducible representations of $G_1 \times G_2$.

Proof. First we check that the α_{ij} are homomorphisms. Indeed,

$$\begin{aligned}
\alpha_{ij}(g_1, g_2)\alpha_{ij}(g_1', g_2') &= \chi_i(g_1)\varphi_j(g_2)\chi_i(g_1')\varphi_j(g_2') \\
&= \chi_i(g_1)\chi_i(g_1')\varphi_j(g_2)\varphi_j(g_2') \\
&= \chi_i(g_1g_1')\varphi_j(g_2g_2') \\
&= \alpha_{ij}(g_1g_1', g_2g_2') \\
&= \alpha_{ij}((g_1, g_2)(g_1', g_2')).
\end{aligned}$$

Next we verify that $\alpha_{ij} = \alpha_{k\ell}$ implies $i = k$ and $j = \ell$. For if this is the case, then

$$\chi_i(g) = \alpha_{ij}(g, 1) = \alpha_{k\ell}(g, 1) = \chi_k(g)$$

and so $i = k$. Similarly, $j = \ell$. Since $G_1 \times G_2$ has $|G_1 \times G_2| = mn$ distinct irreducible representations, it follows that the α_{ij} with $1 \leq i \leq m$, $1 \leq j \leq n$ are all of them. $\square$

Example 4.5.2. Let's compute the character table of the Klein four group $\mathbb{Z}_2 \times \mathbb{Z}_2$. The character table of $\mathbb{Z}_2$ is given in Table 4.3 and so for $\mathbb{Z}_2 \times \mathbb{Z}_2$

	$\overline{0}$	$\overline{1}$
χ_1	1	1
χ_2	1	-1

Table 4.3: The character table of $\mathbb{Z}_2$

the character table is as in Table 4.4.

	$(\overline{0},\overline{0})$	$(\overline{0},\overline{1})$	$(\overline{1},\overline{0})$	$(\overline{1},\overline{1})$
α_{11}	1	1	1	1
α_{12}	1	-1	1	-1
α_{21}	1	1	-1	-1
α_{22}	1	-1	-1	1

Table 4.4: The character table of $\mathbb{Z}_2 \times \mathbb{Z}_2$

Exercises

Exercise 4.1. Let $\varphi\colon G \to GL(V)$ be an irreducible representation. Let

$$Z(G) = \{a \in G \mid ag = ga, \forall g \in G\}$$

be the center of G. Show that if $a \in Z(G)$, then $\varphi(a) = \lambda I$ some $\lambda \in \mathbb{C}^*$.

Exercise 4.2. Let $\operatorname{sgn}\colon S_n \to \mathbb{C}^*$ be the representation given by

$$\operatorname{sgn}(\sigma) = \begin{cases} 1 & \sigma \text{ is even} \\ -1 & \sigma \text{ is odd.} \end{cases}$$

Show that if χ is the character of an irreducible representation of S_n not equivalent to sgn, then

$$\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\chi(\sigma) = 0.$$

Exercise 4.3. Let $\varphi\colon G \to GL_n(\mathbb{C})$ and $\rho\colon G \to GL_m(\mathbb{C})$ be representations. Let $V = M_{mn}(\mathbb{C})$ be the vector space of $m \times n$-matrices over $\mathbb{C}$. Define $\tau\colon G \to GL(V)$ by $\tau_g(A) = \rho_g A \varphi_g^T$ where B^T is the transpose of a matrix B.

1. Show that τ is a representation of G.
2. Show that
$$\tau_g E_{k\ell} = \sum_{i,j} \rho_{ik}(g)\varphi_{j\ell}(g)E_{ij}.$$
3. Prove that $\chi_\tau(g) = \chi_\rho(g)\chi_\varphi(g)$. (Hint: you need to compute the coefficient of $E_{k\ell}$ in $\tau_g E_{k\ell}$ and add this up over all k, ℓ.)

Exercise 4.4. Let $\alpha\colon S_n \to GL_n(\mathbb{C})$ be the representation given by defining $\alpha_\sigma(e_i) = e_{\sigma(i)}$ on the standard basis $\{e_1 \ldots, e_n\}$ for $\mathbb{C}^n$.

1. Show that $\chi_\alpha(\sigma)$ is the number of fixed points of σ, that is, the number of elements $k \in \{1, \ldots, n\}$ such that $\sigma(k) = k$.

2. Show that if $n = 3$, then $\langle \chi_\alpha, \chi_\alpha \rangle = 2$ and hence α is not irreducible.

Exercise 4.5. Let χ be a non-trivial irreducible character of a finite group G. Show that

$$\sum_{g \in G} \chi(g) = 0.$$

Exercise 4.6. Let $\varphi\colon G \to H$ be a surjective homomorphism and let $\psi\colon H \to GL(V)$ be an irreducible representation. Prove that $\psi \circ \varphi$ is an irreducible representation of G.

Exercise 4.7. Let G_1 and G_2 be finite groups and let $G = G_1 \times G_2$. Suppose $\rho\colon G_1 \to GL_m(\mathbb{C})$ and $\varphi\colon G_2 \to GL_n(\mathbb{C})$ are representations. Let $V = M_{mn}(\mathbb{C})$ be the vector space of $m \times n$-matrices over $\mathbb{C}$. Define $\tau\colon G \to GL(V)$ by $\tau_{(g_1, g_2)}(A) = \rho_{g_1} A \varphi_{g_2}^T$ where B^T is the transpose of a matrix B.

1. Show that τ is a representation of G.

2. Prove that $\chi_\tau(g_1, g_2) = \chi_\rho(g_1)\chi_\varphi(g_2)$.

3. Show that if ρ and φ are irreducible, then τ is irreducible.

4. Prove that every irreducible representation of $G_1 \times G_2$ can be obtained in this way.

Exercise 4.8. Let $Q = \{\pm 1, \pm \hat{\imath}, \pm \hat{\jmath}, \pm \hat{k}\}$ be the group of quaternions. The key rules to know are that $\hat{\imath}^2 = \hat{\jmath}^2 = \hat{k}^2 = \hat{\imath}\hat{\jmath}\hat{k} = -1$.

1. Show that $\rho\colon Q \to GL_2(\mathbb{C})$ defined by

$$\rho(\pm 1) = \pm \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \rho(\pm \hat{\imath}) = \pm \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix},$$
$$\rho(\pm \hat{\jmath}) = \pm \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \rho(\pm \hat{k}) = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$$

is an irreducible representation of Q. Just verify that it is irreducible. You may assume that it is a representation (although you should check this on scrap paper for your own edification).

2. Find 4 inequivalent degree one representations of Q. Hint: $N = \{\pm 1\}$ is a normal subgroup of Q and $Q/N \cong \mathbb{Z}_2 \times \mathbb{Z}_2$. Use this to obtain the 4 inequivalent representations of degree 1.

3. Show that the conjugacy classes of Q are $\{1\}$, $\{-1\}$, $\{\pm\hat{\imath}\}$, $\{\pm\hat{\jmath}\}$, $\{\pm\hat{k}\}$.

4. Write down the character table for Q.

Exercise 4.9. Let G be a group and let G' be the commutator subgroup of G. That is, G' is the subgroup of G generated by all commutators $[g,h] = g^{-1}h^{-1}gh$ with $g,h \in G$. You may take for granted the following facts that are typically proved in a first course in group theory:

i. G' is a normal subgroup of G.

ii. G/G' is an abelian group.

iii. if N is a normal subgroup of G, then G/N is a abelian if and only if $G' \subseteq N$.

Let $\varphi\colon G \to G/G'$ be the canonical homomorphism given by $\varphi(g) = gG'$. Prove that every degree one representation $\rho\colon G \to \mathbb{C}^*$ is of the form $\psi \circ \varphi$ where $\psi\colon G/G' \to \mathbb{C}^*$ is a degree one representation of the abelian group G/G'.

Exercise 4.10. Show that if G is a finite group and g is a non-trivial element of G, then there is an irreducible representation φ with $\varphi(g) \neq I$.
Hint: Let $L\colon G \to GL(\mathbb{C}G)$ be the regular representation. Show that $L_g \neq I$. Use the decomposition of L into irreducible representations to show that $\varphi_g \neq I$ for some irreducible.

Chapter 5

Fourier Analysis on Finite Groups

In this chapter we introduce an algebraic structure on $L(G)$ coming from the convolution product. The Fourier transform then permits us to analyze this structure more clearly in terms of known rings. In particular, we prove Wedderburn's theorem for group algebras over the complex numbers. Due to its applications in signal and image processing, Fourier analysis is one of the most important aspects of mathematics. There are entire books dedicated to Fourier analysis on finite groups. Unfortunately, we merely scratch the surface of this rich theory in this text. In particular, the only application that we give is to computing the eigenvalues of the adjacency matrix of a Cayley graph of an abelian group.

5.1 Periodic functions on cyclic groups

We begin with the classical case of periodic functions on the integers.

Definition 5.1.1 (Periodic function)**.** A function $f\colon \mathbb{Z} \to \mathbb{C}$ is said to be *periodic* with period n if $f(x) = f(x+n)$ for all $x \in \mathbb{Z}$.

Notice that if n is a period for f, then so is any multiple of n. It is easy to see that periodic functions with period n are in bijection with elements of $L(\mathbb{Z}_n)$, that is, functions $f\colon \mathbb{Z}_n \to \mathbb{C}$. Indeed, the definition of a periodic function says precisely that f is constant on residue classes modulo n. Now the irreducible characters form a basis for $L(\mathbb{Z}_n)$ and are given in Example 4.4.10. It follows that if $f\colon \mathbb{Z}_n \to \mathbb{C}$ is a function, then

$$f = \langle \chi_0, f\rangle \chi_0 + \cdots + \langle \chi_{n-1}, f\rangle \chi_{n-1}. \tag{5.1}$$

The Fourier transform encodes this information as a function.

Definition 5.1.2 (Fourier transform)**.** Let $f\colon \mathbb{Z}_n \to \mathbb{C}$. Define the *Fourier transform* $\widehat{f}\colon \mathbb{Z}_n \to \mathbb{C}$ of f by

$$\widehat{f}(\overline{m}) = n\langle \chi_m, f\rangle = \sum_{k=0}^{n-1} e^{-2\pi imk} f(\overline{k})$$

It is immediate that the Fourier transform is a linear transformation $T\colon L(\mathbb{Z}_n) \to L(\mathbb{Z}_n)$ by the linearity of inner products in the second variable. We can rewrite (5.1) as:

Proposition 5.1.3 (Fourier inversion)**.** *The Fourier transform is invertible. More precisely,* $f = \dfrac{1}{n}\displaystyle\sum_{k=0}^{n-1} \widehat{f}(\overline{k})\chi_k$.

The Fourier transform on cyclic groups is used in signal and image processing. The idea is that the values of $\widehat{f}$ correspond to the wavelengths associated to the wave function f. One sets to zero all sufficiently small values of $\widehat{f}$, thereby compressing the wave. To recover something close enough to the original wave, as far as our eyes and ears are concerned, one applies Fourier inversion.

5.2 The convolution product

We now introduce the convolution product on $L(G)$, thereby explaining the terminology group algebra for $L(G)$.

Definition 5.2.1 (Convolution)**.** Let G be a finite group and $a, b \in L(G)$. Then the *convolution* $a * b\colon G \to \mathbb{C}$ is defined by

$$a * b(x) = \sum_{y \in G} a(xy^{-1})b(y). \tag{5.2}$$

Our eventual goal is to show that convolution gives $L(G)$ the structure of a ring. Before that, let us motivate the definition of convolution. To each element $g \in G$, we have associated the delta function δ_g. What could be more natural than to try and assign a multiplication $*$ to $L(G)$ so that $\delta_g * \delta_h = \delta_{gh}$? Let's show that convolution has this property. Indeed

$$\delta_g * \delta_h(x) = \sum_{y \in G} \delta_g(xy^{-1})\delta_h(y)$$

and the only non-zero term is when $y = h$ and $g = xy^{-1} = xh^{-1}$, i.e., $x = gh$. In this case, one gets 1, so we have proved:

Proposition 5.2.2. *For $g, h \in G$, $\delta_g * \delta_h = \delta_{gh}$.* □

Now if $a, b \in L(G)$, then

$$a = \sum_{g \in G} a(g)\delta_g, \ b = \sum_{g \in G} b(g)\delta_g$$

so if $L(G)$ were really a ring, then the distributive law would yield

$$a * b = \sum_{g,h \in G} a(g)b(h)\delta_g * \delta_h = \sum_{g,h \in G} a(g)b(h)\delta_{gh}.$$

Applying the change of variables $x = gh$, $y = h$ then gives us

$$a * b = \sum_{x \in G} \left(\sum_{y \in G} a(xy^{-1})b(y) \right) \delta_x$$

which is equivalent to the formula (5.2). Another motivation for the definition of convolution comes from statistics; see Exercise 5.8.

Theorem 5.2.3. *The set $L(G)$ is a ring with addition taken pointwise and convolution as multiplication. Moreover, δ_1 is a multiplicative identity.*

Proof. We will only verify that δ_1 is the identity and the associativity of convolution. The remaining verifications that $L(G)$ is a ring are straightforward and will be left to the reader. Let $a \in L(G)$. Then

$$a * \delta_1(x) = \sum_{y \in G} a(xy^{-1})\delta_1(y^{-1}) = a(x)$$

since $\delta_1(y^{-1}) = 0$ except when $y = 1$. Similarly, $\delta_1 * a = a$. This proves δ_1 is the identity.

For associativity, let $a, b, c \in L(G)$. Then

$$[(a * b) * c](x) = \sum_{y \in G} [a * b](xy^{-1})c(y) = \sum_{y \in G} \sum_{z \in G} a(xy^{-1}z^{-1})b(z)c(y). \quad (5.3)$$

We make the change of variables $u = zy$ (and so $y^{-1}z^{-1} = u^{-1}$, $z = uy^{-1}$). The right hand side of (5.3) then becomes

$$\begin{aligned}\sum_{y\in G}\sum_{u\in G} a(xu^{-1})b(uy^{-1})c(y) &= \sum_{u\in G} a(xu^{-1})\sum_{y\in G} b(uy^{-1})c(y)\\ &= \sum_{u\in G} a(xu^{-1})[b*c](u)\\ &= [a*(b*c)](x)\end{aligned}$$

completing the proof. □

It is now high time to justify the notation $Z(L(G))$ for the space of class functions on G. Recall that the center $Z(R)$ of a ring R consists of all elements $a \in R$ such that $ab = ba$ all $b \in R$. For instance, the scalar matrices form the center of $M_n(\mathbb{C})$.

Proposition 5.2.4. *$Z(L(G))$ is the center of $L(G)$. That is, $f\colon G \to \mathbb{C}$ is a class function if and only if $a * f = f * a$ for all $a \in L(G)$.*

Proof. Suppose first that f is a class function and let $a \in L(G)$. Then

$$a * f(x) = \sum_{y\in G} a(xy^{-1})f(y) = \sum_{y\in G} a(xy^{-1})f(xyx^{-1}) \tag{5.4}$$

since f is a class function. Setting $z = xy^{-1}$ turns the right hand side of (5.4) into

$$\sum_{z\in G} a(z)f(xz^{-1}) = \sum_{z\in G} f(xz^{-1})a(z) = f * a(x)$$

and hence $a * f = f * a$.

For the other direction, let f be in the center of $L(G)$.

Claim. $f(gh) = f(hg)$ for all $g, h \in G$.

Proof of claim. Observe that

$$\begin{aligned}f(gh) &= \sum_{y\in G} f(gy^{-1})\delta_{h^{-1}}(y) = f * \delta_{h^{-1}}(g)\\ &= \delta_{h^{-1}} * f(g) = \sum_{y\in G} \delta_{h^{-1}}(gy^{-1})f(y) = f(hg)\end{aligned}$$

since $\delta_{h^{-1}}(gy^{-1})$ is non-zero if and only if $gy^{-1} = h^{-1}$, that is, $y = hg$. □

To complete the proof, we note that by the claim $f(ghg^{-1}) = f(hg^{-1}g) = f(h)$, establishing that f is a class function. □

5.3 Fourier analysis on finite abelian groups

In this section, we consider the case of abelian groups as the situation is much simpler and frequently is sufficient for applications to signal processing and number theory. In number theory, the groups of interest are usually $\mathbb{Z}_n$ and $\mathbb{Z}_n^*$.

Let $G = \{g_1, \ldots, g_n\}$ be a finite abelian group. Then class functions on G are the same thing as functions, that is $L(G) = Z(L(G))$. Therefore $L(G)$ is a commutative ring. Let's try to identify it (up to isomorphism) with a known ring. We know that G has $n = |G|$ irreducible characters $\chi_1, \ldots, \chi_n$ and that they form an orthonormal basis for $L(G)$. The secret to analyzing the ring structure on $L(G)$ is the Fourier transform.

Definition 5.3.1 (Fourier transform)**.** Let $f\colon G \to \mathbb{C}$ be a complex-valued function on G. Then the *Fourier transform* $\widehat{f}\colon G \to \mathbb{C}$ is defined by

$$\widehat{f}(g_i) = n\langle \chi_i, f\rangle = \sum_{g\in G} \overline{\chi_i(g)} f(g).$$

The complex numbers $n\langle \chi_i, f\rangle$ are often called the *Fourier coefficients* of f.

Notice that the definition of the Fourier transform depends on an ordering of both G and the characters. For the case of $G = \mathbb{Z}_n$ there are natural orderings for these, namely the ones used in Section 5.1.

Example 5.3.2. If χ_j is an irreducible character of G, then

$$\widehat{\chi_j}(g_i) = n\langle \chi_i, \chi_j\rangle = \begin{cases} n & i = j \\ 0 & \text{else} \end{cases}$$

by the orthogonality relations and so $\widehat{\chi_j} = n\delta_{g_j}$.

Theorem 5.3.3 (Fourier inversion)**.** *If $f \in L(G)$, then*

$$f = \frac{1}{n}\sum_{i=1}^{n} \widehat{f}(g_i)\chi_i.$$

Proof. The proof is a straightforward computation:

$$f = \sum_{i=1}^{n} \langle \chi_i, f\rangle \chi_i = \frac{1}{n}\sum_{i=1}^{n} n\langle \chi_i, f\rangle \chi_i = \frac{1}{n}\sum_{i=1}^{n} \widehat{f}(g_i)\chi_i$$

as required. □

Next we observe that the Fourier transform is a linear operator on $L(G)$.

Proposition 5.3.4. *The map* $T\colon L(G) \to L(G)$ *given by* $Tf = \widehat{f}$ *belongs to* $GL(L(G))$.

Proof. By definition $T(c_1f_1 + c_2f_2) = \widehat{c_1f_1 + c_2f_2}$. Now

$$\begin{aligned}\widehat{c_1f_1 + c_2f_2}(g_i) &= n\langle \chi_i, c_1f_1 + c_2f_2\rangle \\ &= c_1n\langle \chi_i, f_1\rangle + c_2n\langle \chi_i, f_2\rangle \\ &= c_1\widehat{f_1}(g_i) + c_2\widehat{f_2}(g_i)\end{aligned}$$

and so $\widehat{c_1f_1 + c_2f_2} = c_1\widehat{f_1} + c_2\widehat{f_2}$, establishing that T is linear. Theorem 5.3.3 immediately implies T is injective and hence T is invertible. □

There are two ways to make $L(G)$ into a ring: one way is to use convolution; the other is to use pointwise multiplication: $(f \cdot g)(x) = f(x)g(x)$. The reader should observe that δ_1 is the identity for convolution and that the constant map to 1 is the identity for the pointwise product. The next theorem shows that the Fourier transform gives an isomorphism between these two ring structures, that is, it sends convolution to pointwise multiplication.

Theorem 5.3.5. *The Fourier transform satisfies*

$$\widehat{a * b} = \widehat{a} \cdot \widehat{b}.$$

Consequently, the linear map $T\colon L(G) \to L(G)$ *given by* $Tf = \widehat{f}$ *provides a ring isomorphism between* $(L(G), +, *)$ *and* $(L(G), +, \cdot)$.

Proof. We know by Proposition 5.3.4 that T is an isomorphism of vector spaces. Therefore, to show that it is a ring isomorphism it suffices to show $T(a * b) = Ta \cdot Tb$, that is $\widehat{a * b} = \widehat{a} \cdot \widehat{b}$. Let us endeavor to do this.

$$\begin{aligned}\widehat{a * b}(g_i) &= n\langle \chi_i, a * b\rangle \\ &= n \cdot \frac{1}{n}\sum_{x\in G}\overline{\chi_i(x)}(a * b)(x) \\ &= \sum_{x\in G}\overline{\chi_i(x)}\sum_{y\in G}a(xy^{-1})b(y) \\ &= \sum_{y\in G}b(y)\sum_{x\in G}\overline{\chi_i(x)}a(xy^{-1}).\end{aligned}$$

Changing variables, we put $z = xy^{-1}$ (and so $x = zy$). Then we obtain

$$\begin{aligned}\widehat{a*b}(g_i) &= \sum_{y\in G} b(y) \sum_{z\in G} \overline{\chi_i(zy)}a(z) \\ &= \sum_{y\in G} \overline{\chi_i(y)}b(y) \sum_{z\in G} \overline{\chi_i(z)}a(z) \\ &= \sum_{z\in G} \overline{\chi_i(z)}a(z) \sum_{y\in G} \overline{\chi_i(y)}b(y) \\ &= n\langle \chi_i, a\rangle \cdot n\langle \chi_i, b\rangle \\ &= \widehat{a}(g_i)\widehat{b}(g_i)\end{aligned}$$

and so $\widehat{a*b} = \widehat{a} \cdot \widehat{b}$, as was required. □

Let us summarize what we have for the classical case of periodic functions on $\mathbb{Z}$.

Example 5.3.6 (Periodic functions on $\mathbb{Z}$)**.** Let $f, g\colon \mathbb{Z} \to \mathbb{C}$ have period n. Their convolution is defined by

$$f * g(m) = \sum_{k=0}^{n-1} f(m-k)g(k).$$

The Fourier transform is then

$$\widehat{f}(m) = \sum_{k=0}^{n-1} e^{-2\pi imk/n} f(k).$$

The Fourier inversion theorem says that

$$f(m) = \frac{1}{n}\sum_{k=0}^{n-1} e^{2\pi imk/n} \widehat{f}(k).$$

The multiplication formula says that $\widehat{f*g} = \widehat{f} \cdot \widehat{g}$. In practice it is more efficient to compute $\widehat{f} \cdot \widehat{g}$ and then apply Fourier inversion to obtain $f * g$ than to compute $f * g$ directly thanks to the existence of the fast Fourier transform.

The original Fourier transform was invented by Fourier in the continuous context in the early 1800s to study the heat equation. For absolutely integrable complex-valued functions $f, g\colon \mathbb{R} \to \mathbb{C}$, their convolution is defined by

$$f * g(x) = \int_{-\infty}^{\infty} f(x-y)g(y)dy.$$

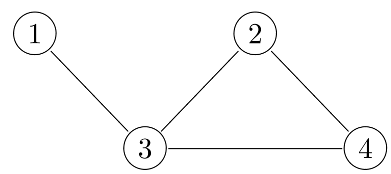


Figure 5.1: A graph

The Fourier transform of f is

$$\widehat{f}(x) = \int_{-\infty}^{\infty} e^{-2\pi ixt} f(t)dt.$$

Fourier inversion says that

$$f(x) = \int_{-\infty}^{\infty} e^{2\pi ixt} \widehat{f}(t)dt.$$

Once again the multiplication rule $\widehat{f * g} = \widehat{f} \cdot \widehat{g}$ holds.

5.4 An application to graph theory

A *graph* Γ consists of a set V of *vertices* and a set E of unordered pairs of elements of V, called *edges*. One often views graphs pictorially by selecting a point for each vertex and drawing a line segment between two vertices that form an edge.

For instance, if Γ has vertex set $V = \{1, 2, 3, 4\}$ and edge set $E = \{\{1, 3\}, \{2, 3\}, \{2, 4\}, \{3, 4\}\}$, then the picture is as in Figure 5.1.

We shall only consider finite graphs in this section. One can usefully encode a graph by its adjacency matrix.

Definition 5.4.1 (Adjacency matrix)**.** Let Γ be a graph with vertex set $V = \{v_1, \ldots, v_n\}$ and edge set E. Then the adjacency matrix $A = (a_{ij})$ is given by

$$a_{ij} = \begin{cases} 1 & \{v_i, v_j\} \in E \\ 0 & \text{else.} \end{cases}$$

Example 5.4.2. For the graph in Figure 5.1, the adjacency matrix is

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

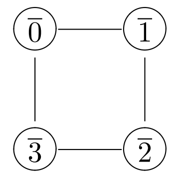


Figure 5.2: The Cayley Graph of $\mathbb{Z}_4$ with respect to $\{\pm\overline{1}\}$

Notice that the adjacency matrix is always symmetric and hence diagonalizable with real eigenvalues by the spectral theorem for matrices. The set of eigenvalues of A is called the *spectrum* of the graph. One can obtain important information from the eigenvalues, such as the number of spanning trees. Also one can verify that A^n_{ij} is the number of paths of length n from v_i to v_j. For a diagonalizable matrix, knowing the eigenvalues already gives a lot of information about powers of the matrix. There is a whole area of graph theory, called spectral graph theory, dedicated to studying graphs via their eigenvalues. The adjacency matrix is also closely related to the study of random walks on the graph.

A natural source of graphs, known as Cayley graphs, comes from group theory. Representation theory affords us a means to analyze the eigenvalues of Cayley graphs, at least for abelian groups.

Definition 5.4.3 (Cayley graph)**.** Let G be a finite group. By a *symmetric subset* of G, we mean a subset $S \subseteq G$ such that:

- $1 \notin S$;
- $s \in S$ implies $s^{-1} \in S$.

If S is a symmetric subset of G, then the *Cayley graph* of G with respect to S is the graph with vertex set G and with an edge $\{g, h\}$ connecting g and h if $gh^{-1} \in S$, or equivalently $hg^{-1} \in S$.

Remark 5.4.4. In this definition S can be empty, in which case the Cayley graph has no edges. One can verify that the Cayley graph is connected (any two vertices can be connected by a path) if and only if S generates G.

Example 5.4.5. Let $G = \mathbb{Z}_4$ and $S = \{\pm\overline{1}\}$. Then the Cayley graph of G with respect to S is drawn in Figure 5.2. The adjacency matrix of this

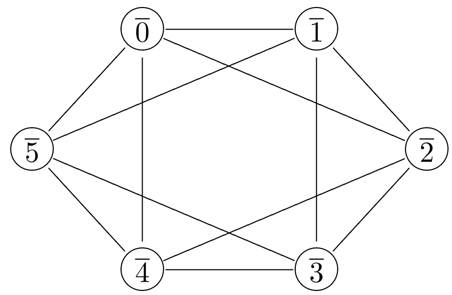


Figure 5.3: The Cayley graph of $\mathbb{Z}_6$ with respect to $\{\pm\overline{1}, \pm\overline{2}\}$

Cayley graph is given by

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

Example 5.4.6. In this example we take $G = \mathbb{Z}_6$ and $S = \{\pm\overline{1}, \pm\overline{2}\}$. The resulting Cayley graph can be found in Figure 5.3. The adjacency matrix of this graph is

$$\begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

The graphs we have been considering are Cayley graphs of cyclic groups. Such graphs have a special name.

Definition 5.4.7 (Circulant)**.** A Cayley graph of $\mathbb{Z}_n$ is a called a *circulant graph* (on n vertices). The adjacency matrix of a circulant graph is called a *circulant matrix.*

Our goal is to describe the eigenvalues of the Cayley graph of an abelian group. First we need a lemma.

Lemma 5.4.8. *Let $G = \{g_1, \ldots, g_n\}$ be an abelian group with irreducible characters $\chi_1, \ldots, \chi_n$ and let $a \in L(G)$. Define the convolution operator $F\colon L(G) \to L(G)$ by $F(b) = a * b$. Then χ_j is an eigenvector of F with eigenvalue $\widehat{a}(g_j)$ for all $1 \leq j \leq n$. Consequently, F is a diagonalizable operator.*

Proof. Using the distributivity of convolution over addition, it is easy to verify that F is linear. Next observe that

$$\widehat{a * \chi_j} = \widehat{a} \cdot \widehat{\chi_j} = \widehat{a} \cdot n\delta_{g_j}$$

where the last equality uses Example 5.3.2. Clearly, one has that

$$(\widehat{a} \cdot n\delta_{g_j})(g_i) = \begin{cases} \widehat{a}(g_j)n & i = j \\ 0 & \text{else} \end{cases}$$

and so $\widehat{a} \cdot n\delta_{g_j} = \widehat{a}(g_j)n\delta_{g_j}$. Applying the Fourier inversion theorem to $\widehat{a * \chi_j} = \widehat{a}(g_j)n\delta_{g_j}$ and using that $\widehat{\chi_j} = n\delta_{g_j}$, we obtain $a * \chi_j = \widehat{a}(g_j)\chi_j$. In other words, $F(\chi_j) = \widehat{a}(g_j)\chi_j$ and so χ_j is an eigenvector of F with eigenvalue $\widehat{a}(g_j)$.

Since $\chi_1, \ldots, \chi_n$ form an orthonormal basis of eigenvectors for F, it follows that F is diagonalizable. □

Lemma 5.4.8 is the key ingredient to computing the eigenvalues of the adjacency matrix of a Cayley graph of an abelian group. It only remains to realize the adjacency matrix as the matrix of a convolution operator.

Theorem 5.4.9. *Let $G = \{g_1, \ldots, g_n\}$ be an abelian group and $S \subseteq G$ a symmetric set. Let $\chi_1, \ldots, \chi_n$ be the irreducible characters of G and let A be the adjacency matrix of the Cayley graph of G with respect to S (using this ordering for the elements of G). Then:*

1. *The eigenvalues of the adjacency matrix A are the real numbers*

$$\lambda_i = \sum_{s \in S} \chi_i(s)$$

where $1 \leq i \leq n$;

2. *The corresponding orthonormal basis of eigenvectors is given by the vectors $\{v_1, \ldots, v_n\}$ where $v_i = (\chi_i(g_1), \ldots, \chi_i(g_n))^T$.*

Proof. Let $G = \{g_1, \ldots, g_n\}$ and let $\delta_S = \sum_{s \in S} \delta_s$ be the characteristic (or indicator) function of S; so

$$\delta_S(x) = \begin{cases} 1 & x \in S \\ 0 & \text{else.} \end{cases}$$

Let $F \colon L(G) \to L(G)$ be the convolution operator

$$F(b) = \delta_S * b.$$

Lemma 5.4.8 implies that the irreducible characters χ_i are eigenvectors of F and that the corresponding eigenvalue is

$$\widehat{\delta_S}(g_i) = n\langle \chi_i, \delta_S\rangle = \sum_{x\in G} \overline{\chi_i(x)}\delta_S(x) = \sum_{x\in S} \overline{\chi_i(x)} = \sum_{s\in S} \chi_i(s) = \lambda_i$$

where the penultimate equality is obtained by putting $s = x^{-1}$ and using that degree one representations are unitary, whence $\chi_i(x^{-1}) = \overline{\chi_i(x)}$, and that S is symmetric.

It follows that if B is the basis $\{\delta_{g_1}, \ldots, \delta_{g_n}\}$ for $L(G)$, then the matrix $[F]_B$ of F with respect to this basis has eigenvalues $\lambda_1, \ldots, \lambda_n$ and eigenvectors $v_1, \ldots, v_n$ (where the orthonormality of the v_i follows from the orthonormality of the χ_i). Therefore, it remains to prove that $A = [F]_B$.

To this end we compute

$$F(\delta_{g_i}) = \delta_S * \delta_{g_j} = \sum_{s\in S} \delta_s * \delta_{g_j} = \sum_{s\in S} \delta_{sg_j}$$

by Proposition 5.2.2. Recalling that $([F]_B)_{ij}$ is the coefficient of δ_{g_i} in $F(\delta_{g_j})$, we conclude that

$$\begin{aligned} ([F]_B)_{ij} &= \begin{cases} 1 & g_i = sg_j \text{ for some } s \in S \\ 0 & \text{else} \end{cases} \\ &= \begin{cases} 1 & g_i g_j^{-1} \in S \\ 0 & \text{else} \end{cases} \\ &= A_{ij} \end{aligned}$$

as required.

Finally, to verify that λ_i is real, we just observe that if $s \in S$, then either $s = s^{-1}$, and so $\chi_i(s) = \chi_i(s^{-1}) = \overline{\chi_i(s)}$ is real, or $s \neq s^{-1} \in S$ and $\chi(s) + \chi(s^{-1}) = \chi(s) + \overline{\chi(s)}$ is real. □

Specializing to the case of circulant matrices, we obtain:

Corollary 5.4.10. *Let A be a circulant matrix of degree n, say it is the adjacency matrix of the Cayley graph of $\mathbb{Z}_n$ with respect to the symmetric set S. Then the eigenvalues of A are*

$$\lambda_k = \sum_{s\in S} e^{2\pi iks/n}$$

where $k = 0, \ldots, n-1$ and a corresponding basis of orthonormal eigenvectors is given by $v_0, \ldots, v_{n-1}$ where $v_k = (1, e^{2\pi ik2/n}, \ldots, e^{2\pi ik(n-1)/n})^T$.

Example 5.4.11. Let A be the adjacency matrix of the circulant graph in Example 5.4.6. Then the eigenvalues of A are $\lambda_1, \ldots, \lambda_6$ where

$$\lambda_k = e^{\pi ik/3} + e^{-\pi ik/3} + e^{2\pi ik/3} + e^{-2\pi ik/3} = 2\cos \pi k/3 + 2\cos 2\pi k/3$$

for $k = 1, \ldots, 6$.

Remark 5.4.12. This approach can be generalized to non-abelian groups provided the symmetric set S is closed under conjugation. For more on the relationship between graph theory and representation theory, as well as the related subject of random walks on graphs, see [1, 3, 4].

5.5 Fourier analysis on non-abelian groups

For a non-abelian group G, we have $L(G) \neq Z(L(G))$ and so $L(G)$ is a non-commutative ring. Therefore, we cannot find a Fourier transform that turns convolution into pointwise multiplication (as pointwise multiplication is commutative). Instead, we try to replace pointwise multiplication by matrix multiplication. To achieve this, let us first recast the abelian case in a different form.

Suppose $G = \{g_1, \ldots, g_n\}$ is a finite abelian group with irreducible characters $\chi_1, \ldots, \chi_n$. Then to each function $f\colon G \to \mathbb{C}$, we can associate its vector of Fourier coefficients. That is, we define $T\colon L(G) \to \mathbb{C}^n$ by

$$Tf = (n\langle \chi_1, f\rangle, n\langle \chi_2, f\rangle, \ldots, n\langle \chi_n, f\rangle) = (\widehat{f}(g_1), \widehat{f}(g_2), \ldots, \widehat{f}(g_n)).$$

The map T is injective by the Fourier inversion theorem since we can recover $\widehat{f}$, and hence f, from Tf. It is also linear (this is essentially a reformulation of Proposition 5.3.4) and hence a vector space isomorphism since $\dim L(G) = n$. Now $\mathbb{C}^n = \mathbb{C} \times \cdots \times \mathbb{C}$ has the structure of a direct product of rings where multiplication is taken coordinate-wise:

$$(a_1, \ldots, a_n)(b_1, \ldots, b_n) = (a_1b_1, \ldots, a_nb_n).$$

The map T is in fact a ring isomorphism since

$$\begin{aligned} T(a * b) &= (\widehat{a * b}(g_1), \ldots, \widehat{a * b}(g_n)) = (\widehat{a}(g_1)\widehat{b}(g_1), \ldots, \widehat{a}(g_n)\widehat{b}(g_n)) \\ &= (\widehat{a}(g_1), \ldots, \widehat{a}(g_n))(\widehat{b}(g_1), \ldots, \widehat{b}(g_n)) = Ta \cdot Tb \end{aligned}$$

Consequently, we have reinterpreted Theorem 5.3.5 in the following way.

Theorem 5.5.1. *Let G be a finite abelian group of order n. Then $L(G) \cong \mathbb{C}^n$.*

One might guess that this reflects the fact that all irreducible representations of an abelian group have degree one and that for non-abelian groups, we must replace $\mathbb{C}$ by matrix rings over $\mathbb{C}$. This is indeed the case. So without further ado, let G be a finite group of order n with complete set $\varphi^{(1)}, \ldots, \varphi^{(s)}$ of unitary representatives of the equivalence classes of irreducible representations of G. As usual, we put $d_k = \deg \varphi^{(k)}$. The matrix coefficients are the functions $\varphi_{ij}^{(k)} \colon G \to \mathbb{C}$ given by $\varphi_g^{(k)} = (\varphi_{ij}^{(k)}(g))$. Theorem 4.4.6 tells us that the functions $\sqrt{d_k}\varphi_{ij}^{(k)}$ form an orthonormal basis for $L(G)$.

Definition 5.5.2 (Fourier transform)**.** Define

$$T \colon L(G) \to M_{d_1}(\mathbb{C}) \times \cdots \times M_{d_s}(\mathbb{C})$$

by $Tf = (\widehat{f}(\varphi^{(1)}), \ldots, \widehat{f}(\varphi^{(s)}))$ where

$$\widehat{f}(\varphi^{(k)})_{ij} = n\langle \varphi_{ij}^{(k)}, f\rangle = \sum_{g\in G} \overline{\varphi_{ij}^{(k)}(g)} f(g). \tag{5.5}$$

We call Tf the *Fourier transform* of f.

Notice that (5.5) can be written more succinctly in the form

$$\widehat{f}(\varphi^{(k)}) = \sum_{g\in G} \overline{\varphi_g^{(k)}} f(g)$$

which is the form that we shall most frequently use[1]. Let us begin with the Fourier inversion theorem.

Theorem 5.5.3 (Fourier inversion)**.** *Let $f \colon G \to \mathbb{C}$ be a complex-valued function on G. Then*

$$f = \frac{1}{n}\sum_{i,j,k} d_k \widehat{f}(\varphi^{(k)})_{ij}\varphi_{ij}^{(k)}.$$

Proof. We compute

$$\begin{aligned} f &= \sum_{i,j,k} \langle \sqrt{d_k}\varphi_{ij}^{(k)}, f\rangle \sqrt{d_k}\varphi_{ij}^{(k)} = \frac{1}{n}\sum_{i,j,k} d_k n \langle \varphi_{ij}^{(k)}, f\rangle \varphi_{ij}^{(k)} \\ &= \frac{1}{n}\sum_{i,j,k} d_k \widehat{f}(\varphi^{(k)})_{ij}\varphi_{ij}^{(k)} \end{aligned}$$

as required. □

[1]Some authors define $\widehat{f}(\varphi^{(k)}) = \sum_{g\in G} \varphi_g^{(k)} f(g)$.

Next we show that T is a vector space isomorphism.

Proposition 5.5.4. *The map* $T\colon L(G) \to M_{d_1}(\mathbb{C}) \times \cdots \times M_{d_s}(\mathbb{C})$ *is a vector space isomorphism.*

Proof. To show that T is linear it suffices to prove

$$\left(\widehat{c_1 f_1 + c_2 f_2}\right)(\varphi^{(k)}) = c_1 \widehat{f_1}(\varphi^{(k)}) + c_2 \widehat{f_2}(\varphi^{(k)})$$

for $1 \leq k \leq s$. Indeed,

$$\begin{aligned}\left(\widehat{c_1 f_1 + c_2 f_2}\right)(\varphi^{(k)}) &= \sum_{g \in G} \overline{\varphi_g^{(k)}}(c_1 f_1 + c_2 f_2)(g) \\ &= c_1 \sum_{g \in G} \overline{\varphi_g^{(k)}} f_1(g) + c_2 \sum_{g \in G} \overline{\varphi_g^{(k)}} f_2(g) \\ &= c_1 \widehat{f_1}(\varphi^{(k)}) + c_2 \widehat{f_2}(\varphi^{(k)})\end{aligned}$$

as was to be proved.

The Fourier inversion theorem implies that T is injective. Since

$$\dim L(G) = |G| = d_1^2 + \cdots + d_s^2 = \dim M_{d_1}(\mathbb{C}) \times \cdots \times M_{d_s}(\mathbb{C})$$

it follows that T is an isomorphism. $\square$

All the preparation has now been completed to show that the Fourier transform is a ring isomorphism. This leads us to a special case of a more general theorem of Wedderburn that is often used as the starting point for studying the representation theory of finite groups.

Theorem 5.5.5 (Wedderburn)**.** *The Fourier transform*

$$T\colon L(G) \to M_{d_1}(\mathbb{C}) \times \cdots \times M_{d_s}(\mathbb{C})$$

is an isomorphism of rings.

Proof. Proposition 5.5.4 asserts that T is an isomorphism of vector spaces. Therefore, to show that it is a ring isomorphism it suffices to show $T(a*b) = Ta \cdot Tb$. In turn, by the definition of multiplication in a direct product, to do this it suffices to establish $\widehat{a*b}(\varphi^{(k)}) = \widehat{a}(\varphi^{(k)}) \cdot \widehat{b}(\varphi^{(k)})$ for $1 \leq k \leq s$.

The computation is analogous to the abelian case:

$$\begin{aligned}\widehat{a*b}(\varphi^{(k)}) &= \sum_{x\in G} \overline{\varphi_x^{(k)}}(a*b)(x)\\ &= \sum_{x\in G} \overline{\varphi_x^{(k)}} \sum_{y\in G} a(xy^{-1})b(y)\\ &= \sum_{y\in G} b(y) \sum_{x\in G} \overline{\varphi_x^{(k)}} a(xy^{-1}).\end{aligned}$$

Setting $z = xy^{-1}$ (and so $x = zy$) yields

$$\begin{aligned}\widehat{a*b}(\varphi^{(k)}) &= \sum_{y\in G} b(y) \sum_{z\in G} \overline{\varphi_{zy}^{(k)}} a(z)\\ &= \sum_{y\in G} b(y) \sum_{z\in G} \overline{\varphi_z^{(k)}} \cdot \overline{\varphi_y^{(k)}} a(z)\\ &= \sum_{z\in G} \overline{\varphi_z^{(k)}} a(z) \sum_{y\in G} \overline{\varphi_y^{(k)}} b(y)\\ &= \widehat{a}(\varphi^{(k)}) \cdot \widehat{b}(\varphi^{(k)})\end{aligned}$$

This concludes the proof that T is a ring isomorphism. □

For non-abelian groups, it is still true that computing $Ta \cdot Tb$ and inverting T can sometimes be faster than computing $a * b$ directly.

Remark 5.5.6. Note that

$$\widehat{\delta_g}(\varphi^{(k)}) = \sum_{x\in G} \overline{\varphi_x^{(k)}} \delta_g(x) = \overline{\varphi_g^{(k)}}.$$

Since the conjugate of an irreducible representation is easily verified to be irreducible, it follows that $T\delta_g$ is a vector whose entries consist of the images of g under all the irreducible representations of G, in some order.

The next example gives some indication how the representation theory of S_n can be used to analyze voting.

Example 5.5.7 (Diaconis)**.** Suppose that in an election each voter has to rank n candidates on a ballot. Let us call the candidates $\{1, \ldots, n\}$. Then to each ballot we can correspond a permutation $\sigma \in S_n$. For example if the ballot ranks the candidates in the order 312, then the corresponding permutation is

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

An election then corresponds to a function $f\colon S_n \to \mathbb{N}$ where $f(\sigma)$ is the number of people whose ballot corresponds to the permutation σ. Using the fast Fourier transform for the symmetric group, Diaconis was able analyze various elections. As with signal processing, one can discard Fourier coefficients of small magnitude to compress data. Also for S_n, the Fourier coefficients $n!\langle \varphi_{ij}^{(k)}, f\rangle$ have nice interpretations. For instance, an appropriate coefficient measures how many people ranked candidate m first amongst all candidates. See [4, 5].

Exercises

Exercise 5.1. Let $f\colon \mathbb{Z}_3 \to \mathbb{C}$ be give by $f(\overline{k}) = \sin(2\pi k/3)$. Compute the Fourier transform $\widehat{f}$ of f.

Exercise 5.2. Draw the Cayley graph of $\mathbb{Z}_6$ with respect to the set $S = \{\pm 2, \pm 3\}$ and compute the eigenvalues of the adjacency matrix.

Exercise 5.3. Let $G = \{g_1, \ldots, g_n\}$ be an abelian group with irreducible characters $\chi_1, \ldots, \chi_n$. Let $a, b \in L(G)$. Prove the Plancherel formula

$$\langle a, b\rangle = \frac{1}{n}\langle \widehat{a}, \widehat{b}\rangle.$$

Exercise 5.4. Prove Lemma 5.4.8 directly from the definition of convolution.

Exercise 5.5. Prove that $Z(M_n(\mathbb{C})) = \{\lambda I \mid \lambda \in \mathbb{C}\}$.

Exercise 5.6. Let G be a finite group of order n and let $\varphi^{(1)}, \ldots, \varphi^{(s)}$ be a complete set of representatives of the equivalence classes of irreducible representations of G. Let χ_i be the character of $\varphi^{(i)}$ and let $e_i = \dfrac{d_i}{n}\chi_i$ where d_i is the degree of $\varphi^{(i)}$.

1. Show that if $f \in Z(L(G))$, then
$$\widehat{f}(\varphi^{(k)}) = \frac{n}{d_k}\langle \chi_k, f\rangle I.$$
2. Deduce that
$$\widehat{e_i}(\varphi^{(k)}) = \begin{cases} I & i = k \\ 0 & \text{else.}\end{cases}$$
3. Deduce that
$$e_i * e_j = \begin{cases} e_i & i = j \\ 0 & \text{else.}\end{cases}$$

4. Deduce that $e_1 + \cdots + e_s$ is the identity δ_1 of $L(G)$.

Exercise 5.7. Let G be a finite group of order n and let $\varphi^{(1)}, \ldots, \varphi^{(s)}$ be a complete set of representatives of the equivalence classes of irreducible representations of G. Let χ_i be the character of $\varphi^{(i)}$ and d_i be the degree of $\varphi^{(i)}$. Suppose $a \in Z(L(G))$ and define a linear operator $F\colon L(G) \to L(G)$ by $F(b) = a * b$.

1. Fix $1 \leq k \leq s$. Show that $\varphi_{ij}^{(k)}$ is an eigenvector of F with eigenvalue $\frac{n}{d_k}\langle \chi_k, a\rangle$. Hint: show that

$$\widehat{\varphi_{ij}^{(m)}}(\varphi^{(k)}) = \begin{cases} \frac{n}{d_k} E_{ij} & m = k \\ 0 & \text{else.} \end{cases}$$

   Now compute $\widehat{a * \varphi_{ij}^{(k)}}$ using Exercise 5.6(1) and apply the Fourier inversion theorem.

2. Conclude that F is a diagonalizable operator.

3. Let $S \subseteq G$ be a symmetric set and assume further that $gSg^{-1} = S$ for all $g \in G$. Show that the eigenvalues of the adjacency matrix A of the Cayley graph of G with respect to S are $\lambda_1, \ldots, \lambda_s$ where

$$\lambda_k = \frac{1}{d_k} \sum_{s \in S} \chi_k(s)$$

   and that λ_k has multiplicity d_k^2.

4. Compute the eigenvalues of the Cayley graph of S_3 with respect to $S = \{(1\ 2), (1\ 3), (2\ 3)\}$.

Exercise 5.8. The following exercise is for readers familiar with probability and statistics. Let G be a finite group and suppose that X, Y are random variables taking values in G with distributions μ, ν respectively, that is,

$$\mathrm{Prob}[X = g] = \mu(g) \quad \text{and} \quad \mathrm{Prob}[Y = g] = \nu(g)$$

for $g \in G$. Show that if X and Y are independent, then the random variable XY has distribution the convolution $\mu * \nu$. Thus the Fourier transform is useful for studying products of group-valued random variables [4].

Chapter 6

Burnside's Theorem

In this chapter, we look at one of the first major applications of representation theory: Burnside's pq-theorem. This theorem states that a non-abelian group of order $p^a q^b$ can never be simple; recall that a group is simple if it contains no non-trivial proper normal subgroups. To prove this we shall need to take a brief excursion into number theory.

6.1 A little number theory

A complex number is called an *algebraic number* if it is the root of a polynomial with integer coefficients. Numbers that are not algebraic are called *transcendental*. For instance $\frac{1}{2}$ is algebraic, being a root of the polynomial $2z-1$, and so is $\sqrt{2}$, as it is a root of z^2-2. A standard course in rings and fields shows that the set $\overline{\mathbb{Q}}$ of algebraic numbers is a field. A fairly straightforward counting argument shows that $\overline{\mathbb{Q}}$ is countable, while $\mathbb{C}$ is uncountable. Thus most numbers are not algebraic, but it is very hard to prove that a given number is transcendental. For example e and π are transcendental, but this is highly non-trivial to prove. Number theory is concerned with integers and so for our purposes we are interested in a special type of algebraic number called an algebraic integer.

Definition 6.1.1 (Algebraic integer)**.** A complex number α is said to be an *algebraic integer* if it is a root of a monic polynomial with integer coefficients. That is to say, there is a polynomial $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$ with $a_0, \ldots, a_{n-1} \in \mathbb{Z}$ and $p(\alpha) = 0$.

The fact that the leading coefficient is 1 is crucial to the definition. Notice that if α is an algebraic integer, then so is $-\alpha$ since if $p(z)$ is a monic polynomial with integer coefficients such that $p(\alpha) = 0$, then either $p(-z)$ or $-p(-z)$ is a monic polynomial and $-\alpha$ is a root of both these polynomials.

Example 6.1.2 (n^{th}-roots)**.** Let m be an integer. Then $z^n - m$ is a monic polynomial with integer coefficients, so any n^{th}-root of m is an algebraic integer. Thus, for example, $\sqrt{2}$ is an algebraic integer, as is $e^{2\pi i/n}$. In fact any n^{th}-root of unity is an algebraic integer.

Example 6.1.3 (Eigenvalues of integer matrices)**.** Let $A = (a_{ij})$ with the $a_{ij} \in \mathbb{Z}$ be an $n \times n$ integer matrix. Then the characteristic polynomial $p_A(z) = \det(zI - A)$ is a monic polynomial with integer coefficients. Thus each eigenvalue of A is an algebraic integer.

A rational number like $1/2$ is a root of a non-monic integral polynomial $2z - 1$. One would guess then that rational numbers cannot be algebraic integers unless they are integers. This is indeed the case, as follows from the "Rational Roots Test" from high school.

Proposition 6.1.4. *A rational number r is an algebraic integer if and only if it is an integer.*

Proof. Write $r = m/n$ with $m, n \in \mathbb{Z}$, $n > 0$ and $\gcd(m, n) = 1$. Suppose r is a root of $z^k + a_{k-1}z^{k-1} + \cdots + a_0$. Then

$$0 = \left(\frac{m}{n}\right)^k + a_{k-1}\left(\frac{m}{n}\right)^{k-1} + \cdots + a_0$$

and so clearing denominators (by multiplying by n^k) yields

$$0 = m^k + a_{k-1}m^{k-1}n + \cdots + a_1mn^{k-1} + a_0n^k.$$

In other words,

$$m^k = -n(a_{k-1}m^{k-1} + \cdots + a_1mn^{k-1} + a_0n^{k-1})$$

and so $n \mid m^k$. As $\gcd(m, n) = 1$, we conclude $n = 1$. Thus $r = m \in \mathbb{Z}$. □

A general strategy to show that an integer d divides an integer n is to show that n/d is an algebraic integer. Proposition 6.1.4 then implies $d \mid n$. First we need to learn more about algebraic integers. Namely, we want to show that they form a subring $\mathbb{A}$ of $\mathbb{C}$. To do this we need the following lemma.

Lemma 6.1.5. *An element $y \in \mathbb{C}$ is an algebraic integer if and only if there exist $y_1, \ldots, y_t \in \mathbb{C}$, not all zero, such that*

$$yy_i = \sum_{j=1}^{t} a_{ij} y_j$$

with the $a_{ij} \in \mathbb{Z}$ for all $1 \le i \le t$ (i.e., yy_i is an integral linear combination of the y_i for all i).

Proof. Suppose first that y is an algebraic integer. Let y be a root of

$$p(z) = z^n + a_{n-1} z^{n-1} + \cdots + a_0$$

and take $y_i = y^{i-1}$ for $1 \le i \le n$. Then, for $1 \le i \le n-2$, we have $yy_i = yy^{i-1} = y^i = y_{i+1}$ and $yy_{n-1} = y^n = -a_0 - \cdots - a_{n-1}y^{n-1}$.

Conversely, if $y_1, \ldots, y_t$ are as in the statement of the lemma, let $A = (a_{ij})$ and

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_t \end{bmatrix} \in \mathbb{C}^t.$$

Then

$$[AY]_i = \sum_{j=1}^{t} a_{ij} y_j = yy_i = y[Y]_i$$

and so $AY = yY$. Since $Y \neq 0$ by assumption, it follows that y is an eigenvalue of the $t \times t$ integer matrix A and hence is an algebraic integer by Example 6.1.3. ☐

Corollary 6.1.6. *The set $\mathbb{A}$ of algebraic integers is a subring of $\mathbb{C}$. In particular, the sum and product of algebraic integers is algebraic.*

Proof. We already observed that $\mathbb{A}$ is closed under taking negatives. Let $y, y' \in \mathbb{A}$. Choose $y_1, y_2, \ldots, y_t \in \mathbb{C}$ not all 0 and $y'_1, \ldots, y'_s \in \mathbb{C}$ not all 0 such that

$$yy_i = \sum_{j=1}^{t} a_{ij} y_j, \ y'y'_k = \sum_{j=1}^{s} b_{kj} y'_j$$

as guaranteed by Lemma 6.1.5. Then

$$(y + y')y_i y'_k = yy_i y'_k + y'y'_k y_i = \sum_{j=1}^{t} a_{ij} y_j y'_k + \sum_{j=1}^{s} b_{kj} y'_j y_i$$

is an integral linear combination of the $y_j y'_\ell$, establishing that $y + y' \in \mathbb{A}$ by Lemma 6.1.5. Similarly, $yy'y_i y'_k = yy_i y' y'_k$ is an integral linear combination of the $y_j y'_\ell$ and so $yy' \in \mathbb{A}$. □

We shall also need that the complex conjugate of an algebraic integer is an algebraic integer. Indeed, if $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$ is a polynomial with integer coefficients and α is a root of $p(z)$, then

$$p(\overline{\alpha}) = \overline{\alpha}^n + a_{n-1}\overline{\alpha}^{n-1} + \cdots + a_0 = \overline{\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_0} = \overline{p(\alpha)} = 0.$$

6.2 The dimension theorem

The relevance of algebraic integers to group representation theorem becomes apparent with the following corollary to Corollary 6.1.6.

Corollary 6.2.1. *Let χ be a character of a finite group G. Then $\chi(g)$ is an algebraic integer all $g \in G$.*

Proof. Let $\varphi\colon G \to GL_m(\mathbb{C})$ be a representation with character χ. Let n be the order of G. Then $g^n = 1$ and so $\varphi_g^n = I$. Corollary 4.1.10 then implies that φ_g is diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_m$ that are n^{th}-roots of unity. In particular, the eigenvalues of φ_g are algebraic integers. Since

$$\chi(g) = \mathrm{Tr}(\varphi_g) = \lambda_1 + \cdots + \lambda_m$$

and algebraic integers form a ring, we conclude that $\chi(g)$ is an algebraic integer. □

Remark 6.2.2. Notice that the proof of Corollary 6.2.1 shows that $\chi_\varphi(g)$ is a sum of m n^{th}-roots of unity. We shall use this fact later.

Our next goal is to show that the degree of an irreducible representation divides the order of the group. To do this we need to conjure up some more algebraic integers.

Theorem 6.2.3. *Let φ be an irreducible representation of a finite group G of degree d. Let $g \in G$ and let h be the size of the conjugacy class of g. Then $\frac{h}{d}\chi_\varphi(g)$ is an algebraic integer.*

Proof. Let $C_1, \ldots, C_s$ be the conjugacy classes of G. Set $h_i = |C_i|$ and let χ_i be the value of χ_φ on the class C_i. We want to show that $\frac{h_i}{d}\chi_i$ is an algebraic integer for each i. Consider the operator

$$T_i = \sum_{x \in C_i} \varphi_x.$$

Claim. $T_i = \dfrac{h_i}{d}\chi_i \cdot I$.

Proof of claim. We first show that $\varphi_g T_i \varphi_{g^{-1}} = T_i$ for all $g \in G$. Indeed,

$$\varphi_g T_i \varphi_{g^{-1}} = \sum_{x \in C_i} \varphi_g \varphi_x \varphi_{g^{-1}} = \sum_{x \in C_i} \varphi_{gxg^{-1}} = \sum_{y \in C_i} \varphi_y = T_i$$

since C_i is closed under conjugation and conjugation by g is a permutation. By Schur's lemma, $T_i = \lambda I$ some $\lambda \in \mathbb{C}$. Then since I is the identity operator on a d-dimensional vector space

$$d\lambda = \mathrm{Tr}(\lambda I) = \mathrm{Tr}(T_i) = \sum_{x \in C_i} \mathrm{Tr}(\varphi_x) = \sum_{x \in C_i} \chi_\varphi(x) = \sum_{x \in C_i} \chi_i = |C_i|\chi_i = h_i\chi_i$$

and so $\lambda = \dfrac{h_i}{d}\chi_i$, establishing the claim. □

We need yet another claim, which says the T_i "behave" like algebraic integers.

Claim. $T_i T_j = \sum_{k=1}^{s} a_{ijk} T_k$ some $a_{ijk} \in \mathbb{Z}$.

Proof of claim. Routine calculation shows

$$T_i T_j = \sum_{x \in C_i} \varphi_x \cdot \sum_{y \in C_j} \varphi_y = \sum_{x \in C_i, y \in C_j} \varphi_{xy} = \sum_{g \in G} a_{ijg} \varphi_g$$

where $a_{ijg} \in \mathbb{Z}$ is the number of ways to write $g = xy$ with $x \in C_i$ and $y \in C_j$. We claim that a_{ijg} depends only on the conjugacy class of g. Suppose that this is indeed the case and let a_{ijk} be the value of a_{ijg} with $g \in C_k$. Then

$$\sum_{g \in G} a_{ijg}\varphi_g = \sum_{k=1}^{s} \sum_{g \in C_k} a_{ijg}\varphi_g = \sum_{k=1}^{s} a_{ijk} \sum_{g \in C_k} \varphi_g = \sum_{k=1}^{s} a_{ijk} T_k$$

proving the claim.

So let's check that a_{ijg} depends only on the conjugacy class of g. Let

$$X_g = \{(x, y) \in C_i \times C_j \mid xy = g\};$$

so $a_{ijg} = |X_g|$. Let g' be conjugate to g. We show that $|X_g| = |X_{g'}|$. Suppose that $g' = kgk^{-1}$ and define a bijection $\psi\colon X_g \to X_g'$ by

$$\psi(x, y) = (kxk^{-1}, kyk^{-1}).$$

Notice that $kxk^{-1} \in C_i$, $kyk^{-1} \in C_j$ and $kxk^{-1}kyk^{-1} = kxyk^{-1} = kgk^{-1} = g'$, and so $\psi(x, y) \in X_g'$. Evidently, ψ has inverse $\tau\colon X_{g'} \to X_g$ given by $\tau(x', y') = (k^{-1}x'k, k^{-1}y'k)$ so ψ is a bijection and hence $|X_g| = |X_{g'}|$. □

We now complete the proof of the theorem. Substituting the formula for the T_i from the first claim into the formula from the second claim yields

$$\left(\frac{h_i}{d}\chi_i\right) \cdot \left(\frac{h_j}{d}\chi_j\right) = \sum_{k=1}^{s} a_{ijk}\left(\frac{h_k}{d}\chi_k\right)$$

and so $\dfrac{h_i}{d}\chi_i$ is an algebraic integer by Lemma 6.1.5. □

Theorem 6.2.4 (Dimension theorem)**.** *Let φ be an irreducible representation of G of degree d. Then d divides $|G|$.*

Proof. The first orthogonality relations (Theorem 4.3.9) provide

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

and so

$$\frac{|G|}{d} = \sum_{g\in G}\overline{\chi_\varphi(g)}\frac{\chi_\varphi(g)}{d}. \tag{6.1}$$

Let $C_1, \ldots, C_s$ be the conjugacy classes of G and let χ_i be the value of χ_φ on C_i. Let $h_i = |C_i|$. Then from (6.1) we obtain

$$\frac{|G|}{d} = \sum_{i=1}^{s}\sum_{g\in C_i}\overline{\chi_\varphi(g)}\frac{\chi_\varphi(g)}{d} = \sum_{i=1}^{s}\sum_{g\in C_i}\overline{\chi_i}\left(\frac{1}{d}\chi_i\right) = \sum_{i=1}^{s}\overline{\chi_i}\left(\frac{h_i}{d}\chi_i\right). \tag{6.2}$$

But $\dfrac{h_i}{d}\chi_i$ is an algebraic integer by Theorem 6.2.3, while $\overline{\chi_i}$ is an algebraic integer by Corollary 6.2.1 and the closure of algebraic integers under complex conjugation. Since the algebraic integers form a ring, it follows from (6.2) that $|G|/d$ is an algebraic integer and hence an integer by Proposition 6.1.4. Therefore, d divides $|G|$. □

The following corollaries are usually proved using facts about p-groups and Sylow's theorems.

Corollary 6.2.5. *Let p be a prime and $|G| = p^2$. Then G is a abelian.*

Proof. Let $d_1, \ldots, d_s$ be the degrees of the irreducible representations of G. Then d_i can be 1, p or p^2. Since the trivial representation has degree 1 and

$$p^2 = |G| = d_1^2 + \cdots + d_s^2$$

it follows that all $d_i = 1$ and hence G is abelian. □

Recall that the *commutator subgroup* G' of a group G is the subgroup generated by all elements of the form $g^{-1}h^{-1}gh$ with $g, h \in G$. It is a normal subgroup and has the properties that G/G' is abelian and if N is any normal subgroup with G/N abelian, then $G' \subseteq N$.

Lemma 6.2.6. *Let G be a finite group. Then the number of degree one representations of G divides $|G|$. More precisely, if G' is the commutator subgroup of G, then there is a bijection between degree one representations of G and irreducible representations of the abelian group G/G'. Hence G has $|G/G'| = [G : G']$ degree one representations.*

Proof. Let $\varphi\colon G \to G/G'$ be the canonical projection. If $\psi\colon G/G' \to \mathbb{C}^*$ is an irreducible representation, then $\psi\varphi\colon G \to \mathbb{C}^*$ is a degree one representation. We now show that every degree one representation of G is obtained in this way. Let $\rho\colon G \to \mathbb{C}^*$ be a degree one representation. Then $\operatorname{Im}\rho \cong G/\ker\rho$ is abelian. Therefore $G' \subseteq \ker\rho$. Define $\psi\colon G/G' \to \mathbb{C}^*$ by $\psi(gG') = \rho(g)$. This is well defined since if $gG' = hG'$, then $h^{-1}g \in G' \subseteq \ker\rho$ and so $\rho(h^{-1}g) = 1$. Thus $\rho(h) = \rho(g)$. Clearly $\psi(gG'hG') = \psi(ghG') = \rho(gh) = \rho(g)\rho(h) = \psi(gG')\psi(hG')$ and so ψ is a homomorphism. By construction $\rho = \psi\varphi$, completing the proof. □

Corollary 6.2.7. *Let p, q be primes with $p < q$ and $q \not\equiv 1 \bmod p$. Then any group G of order pq is abelian.*

Proof. Let $d_1, \ldots, d_s$ be the degrees of the irreducible representations of G. Since d_i divides $|G|$, $p < q$ and

$$pq = |G| = d_1^2 + \cdots + d_s^2$$

it follows that $d_i = 1, p$ all i. Let n be the number of degree p representations of G and let m be the number of degree 1 representations of G. Then $pq = m + np^2$. Since m divides $|G|$ by Lemma 6.2.6, $m \geq 1$ (there is at least the trivial representation) and $p \mid m$, we must have $m = p$ or $m = pq$. If $m = p$, then $q = 1 + np$ contradicting that $q \not\equiv 1 \bmod p$. Therefore, $m = pq$ and so all the irreducible representations of G have degree one. Thus G is abelian. □

6.3 Burnside's theorem

Let G be a group of order n and suppose that $\varphi\colon G \to GL_d(\mathbb{C})$ is a representation. Then $\chi_\varphi(g)$ is a sum of d n^{th}-roots of unity, as was noted in Remark 6.2.2. This explains the relevance of our next lemma.

Lemma 6.3.1. *Let $\lambda_1, \ldots, \lambda_d$ be n^{th}-roots of unity. Then*

$$|\lambda_1 + \cdots + \lambda_d| \leq d$$

and equality holds if and only if $\lambda_1 = \lambda_2 = \cdots = \lambda_d$.

Proof. If $v, w \in \mathbb{R}^2$ are vectors, then

$$\|v + w\|^2 = \|v\|^2 + 2\langle v, w\rangle + \|w\|^2 = \|v\|^2 + 2\|v\| \cdot \|w\| \cos\theta + \|w\|^2$$

where θ is the angle between v and w. Since $\cos\theta \leq 1$ with equality if and only if $\theta = 0$, it follows that $\|v + w\| \leq \|v\| + \|w\|$ with equality if and only if $v = \lambda w$ or $w = \lambda v$ some $\lambda \geq 0$.

Induction then yields $|\lambda_1 + \cdots + \lambda_d| \leq |\lambda_1| + \cdots + |\lambda_d|$ with equality if and only if the λ_i are non-negative scalar multiples of some complex number z. But $|\lambda_1| = \cdots = |\lambda_d| = 1$, so they can only be non-negative multiples of the same complex number if they are the equal. This completes the proof. □

Let $\omega = e^{2\pi i/n}$. Denote by $\mathbb{Q}[\omega]$ the smallest subfield of $\mathbb{C}$ containing ω. This is the smallest subfield F of $\mathbb{C}$ so that $z^n - 1 = (z - \alpha_1) \cdots (z - \alpha_n)$ with $\alpha_1, \ldots, \alpha_n \in F$, i.e., the splitting field of $z^n - 1$. Fields of the form $\mathbb{Q}[\omega]$ are called *cyclotomic fields*. Let ϕ be the Euler ϕ-function; so $\phi(n)$ is the number of positive integers less than n that are relatively prime to it. The following is usually proved in a course on rings and fields.

Lemma 6.3.2. *The field $\mathbb{Q}[\omega]$ has dimension $\phi(n)$ as a $\mathbb{Q}$-vector space.*

Actually, all we really require is that the dimension is finite, which follows since ω is an algebraic number. We shall also need a little bit of Galois theory. Let $\Gamma = \mathrm{Gal}(\mathbb{Q}[w] : \mathbb{Q})$. That is Γ is the group of all field automorphisms $\sigma\colon \mathbb{Q}[\omega] \to \mathbb{Q}[\omega]$ such that $\sigma(r) = r$ all $r \in \mathbb{Q}$ (actually this last condition is automatic). It follows from the fundamental theorem of Galois theory that $|\Gamma| = \phi(n)$ since $\dim \mathbb{Q}[\omega] = \phi(n)$ as a $\mathbb{Q}$-vector space and $\mathbb{Q}[\omega]$ is the splitting field of the polynomial $z^n - 1$. In fact, one can prove that $\Gamma \cong \mathbb{Z}_n^*$, although we will not use this; for us the important thing is that Γ is finite.

A crucial fact is that if $p(z)$ is a polynomial with rational coefficients, then Γ permutes the roots of p in $\mathbb{Q}[\omega]$.

Lemma 6.3.3. *Let $p(z)$ be a polynomial with rational coefficients and suppose that $\alpha \in \mathbb{Q}[\omega]$ is a root of p. Then $\sigma(\alpha)$ is also a root of p all $\sigma \in \Gamma$.*

Proof. Suppose $p(z) = a_k z^k + a_{k-1} z^{k-1} + \cdots + a_0$ with the $a_i \in \mathbb{Q}$. Then

$$\begin{aligned} p(\sigma(\alpha)) &= a_k \sigma(\alpha)^k + a_{k-1}\sigma(\alpha)^{k-1} + \cdots + a_0 \\ &= \sigma(a_k \alpha^k + a_{k-1}\alpha^{k-1} + \cdots + a_0) \\ &= \sigma(0) \\ &= 0 \end{aligned}$$

since $\sigma(a_i) = a_i$ for all i. □

Corollary 6.3.4. *Let α be an n^{th}-root of unity. Then $\sigma(\alpha)$ is also an n^{th}-root of unity for all $\sigma \in \Gamma$.*

Proof. Apply Lemma 6.3.3 to the polynomial $z^n - 1$. □

Remark 6.3.5. The proof that $\Gamma \cong \mathbb{Z}_n^*$ follows fairly easily from Corollary 6.3.4; we sketch it here. Since Γ permutes the roots of $z^n - 1$, it acts by automorphisms on the cyclic group $C_n = \{\omega^k \mid 0 \leq k \leq n-1\}$ of order n. As the automorphism group of a cyclic group of order n is isomorphic to $\mathbb{Z}_n^*$, this determines a homomorphism $\tau \colon \Gamma \to \mathbb{Z}_n^*$ by $\tau(\sigma) = \sigma|_{C_n}$. Since $\mathbb{Q}[\omega]$ is generated over $\mathbb{Q}$ by ω, each element of Γ is determined by what it does to ω and hence τ is injective. Since $|\Gamma| = \phi(n) = |\mathbb{Z}_n^*|$, it must be that τ is an isomorphism.

Corollary 6.3.6. *Let $\alpha \in \mathbb{Q}[\omega]$ be an algebraic integer and suppose $\sigma \in \Gamma$. Then $\sigma(\alpha)$ is an algebraic integer.*

Proof. If α is a root of the monic polynomial p with integer coefficients, then so is $\sigma(\alpha)$ by Lemma 6.3.3. □

Another consequence of the fundamental theorem of Galois theory that we shall require is:

Theorem 6.3.7. *Let $\alpha \in \mathbb{Q}[\omega]$. Then $\sigma(\alpha) = \alpha$ all $\sigma \in \Gamma$ if and only if $\alpha \in \mathbb{Q}$.*

The following corollary is a thinly disguised version of the averaging trick.

Corollary 6.3.8. *Let $\alpha \in \mathbb{Q}[\omega]$. Then $\prod_{\sigma\in\Gamma} \sigma(\alpha) \in \mathbb{Q}$.*

Proof. Let $\tau \in \Gamma$. Then we have

$$\tau\left(\prod_{\sigma\in\Gamma} \sigma(\alpha)\right) = \prod_{\sigma\in\Gamma} \tau\sigma(\alpha) = \prod_{\rho\in\Gamma} \rho(\alpha)$$

where the last equality is obtained by setting $\rho = \tau\sigma$. Theorem 6.3.7 now yields the desired conclusion. □

The next theorem is of a somewhat technical nature, but is crucial to proving Burnside's theorem.

Theorem 6.3.9. *Let G be a group of order n and let C be a conjugacy class of G. Let $\varphi\colon G \to GL_d(\mathbb{C})$ be an irreducible representation and assume $h = |C|$ is relatively prime to d. Then either*

1. *$\varphi_g = \lambda I$ some $\lambda \in \mathbb{C}^*$ for all $g \in C$; or*
2. *$\chi_\varphi(g) = 0$ all $g \in C$.*

Proof. Set $\chi = \chi_\varphi$. First note that if $\varphi_g = \lambda I$ for some $g \in C$, then $\varphi_x = \lambda I$ for all $x \in C$ since conjugating a scalar matrix does not change it. Also since χ is a class function, if it vanishes on any element of C, it must vanish on all elements of C. Therefore it suffices to show that if $\varphi_g \neq \lambda I$ for some $g \in C$, then $\chi_\varphi(g) = 0$.

By Theorem 6.2.3 we know that $\frac{h}{d}\chi(g)$ is an algebraic integer; also $\chi(g)$ is an algebraic integer by Corollary 6.2.1. Since $\gcd(d, h) = 1$, we can find integers k, j so that $kh + jd = 1$. Let

$$\alpha = k\left(\frac{h}{d}\chi(g)\right) + j\chi(g) = \frac{kh + jd}{d}\chi(g) = \frac{\chi(g)}{d}.$$

Then α is an algebraic integer. By Corollary 4.1.10, φ_g is diagonalizable and its eigenvalues $\lambda_1, \ldots, \lambda_d$ are n^{th}-roots of unity. Since φ_g is diagonalizable but not a scalar matrix, its eigenvalues are not all the same. Applying Lemma 6.3.1 to $\chi(g) = \lambda_1 + \cdots + \lambda_d$ yields $|\chi(g)| < d$, and so

$$|\alpha| = \left|\frac{\chi(g)}{d}\right| < 1.$$

Also note that $\alpha \in \mathbb{Q}[\omega]$. Let $\sigma \in \Gamma$. Lemma 6.3.6 implies that $\sigma(\alpha)$ is an algebraic integer. Corollary 6.3.4 tells us that

$$\sigma(\chi(g)) = \sigma(\lambda_1) + \cdots + \sigma(\lambda_d)$$

is again a sum of d n^{th}-roots of unity, not all equal. Hence, another application of Lemma 6.3.1 yields

$$|\sigma(\alpha)| = \left|\frac{\sigma(\chi(g))}{d}\right| < 1.$$

Putting all this together, we obtain that $q = \prod_{\sigma\in\Gamma}\sigma(\alpha)$ is an algebraic integer with

$$|q| = \left|\prod_{\sigma\in\Gamma}\sigma(\alpha)\right| = \prod_{\sigma\in\Gamma}|\sigma(\alpha)| < 1.$$

But Corollary 6.3.8 tells us that $q \in \mathbb{Q}$. Therefore, $q \in \mathbb{Z}$ by Proposition 6.1.4. Since $|q| < 1$, we may conclude that $q = 0$ and hence $\sigma(\alpha) = 0$ for some $\sigma \in \Gamma$. But since σ is an automorphism, this implies $\alpha = 0$. We conclude $\chi(g) = 0$, as was to be proved. □

We are just one lemma away from proving Burnside's theorem.

Lemma 6.3.10. *Let G be a finite non-abelian group. Suppose that there is a conjugacy class $C \neq \{1\}$ of G such that $|C| = p^t$ with p prime, $t \geq 0$. Then G is not simple.*

Proof. Assume that G is simple and let $\varphi^{(1)}, \ldots, \varphi^{(s)}$ be a complete set of representatives of the equivalence classes of irreducible representations of G. Let $\chi_1, \ldots, \chi_s$ be their respective characters and $d_1, \ldots, d_s$ their degrees. We may take $\varphi^{(1)}$ to be the trivial representation. Since G is simple, $\ker\varphi^{(k)} = \{1\}$ for $k > 1$ (since $\ker p^{(k)} = G$ implies $\varphi^{(k)}$ is the trivial representation). Therefore, $\varphi^{(k)}$ is injective for $k > 1$ and so, since G is non-abelian and $\mathbb{C}^*$ is abelian, it follows that $d_k > 1$ for $k > 1$. Also, since G is simple, non-abelian $Z(G) = \{1\}$ and so $t > 0$.

Let $g \in C$ and $k > 1$. Let Z_k be the set of all elements of G such that $\varphi^{(k)}_g$ is a scalar matrix. Let $H = \{\lambda I_{d_k} \mid \lambda \in \mathbb{C}^*\}$; then H is a subgroup of $GL_{d_k}(\mathbb{C})$ contained in the center, and hence normal (actually it is the center). As Z_k is the inverse image of H under $\varphi^{(k)}$, we conclude that Z_k is a normal subgroup of G. Since $d_k > 1$, we cannot have $Z_k = G$. Thus $Z_k = \{1\}$ by simplicity of G. Suppose for the moment that $p \nmid d_k$; then $\chi_k(g) = 0$ by Theorem 6.3.9.

Let L be the regular representation of G. Recall $L \sim d_1\varphi^{(1)} \oplus \cdots \oplus d_s\varphi^{(s)}$. Since $g \neq 1$, Proposition 4.4.3 yields

$$\begin{aligned}
0 = \chi_L(g) &= d_1\chi_1(g) + \cdots + d_s\chi_s(g)\\
&= 1 + \sum_{k=2}^{s} d_k\chi_k(g)\\
&= 1 + \sum_{p\mid d_k} d_k\chi_k(g)\\
&= 1 + pz
\end{aligned}$$

where z is an algebraic integer. Hence $1/p = -z$ is an algebraic integer, and thus an integer by Proposition 6.1.4. This contradiction establishes the lemma. □

We are now ready to prove the deepest theorem in this text.

Theorem 6.3.11 (Burnside)**.** *Let G be a group of order $p^a q^b$ with p, q primes. Then G is not simple unless it is cyclic of prime order.*

Proof. Since an abelian group is simple if and only if it is cyclic of prime order, we may assume that G is non-abelian. Since groups of prime power order have non-trivial centers, if a or b is zero, then we are done. Suppose next that $a, b \geq 1$. By Sylow's theorem, G has a subgroup H of order q^b. Let $1 \neq g \in Z(H)$ and let $N_G(g) = \{x \in G \mid xg = gx\}$ be the normalizer of g in G. Then $H \subseteq N_G(g)$ as $g \in Z(H)$. Thus

$$p^a = [G : H] = [G : N_G(g)][N_G(g) : H]$$

and so $[G : N_G(g)] = p^t$ for some $t \geq 0$. But $[G : N_G(g)]$ is the size of the conjugacy class of g. The previous lemma now implies that G is not simple. □

Remark 6.3.12. Burnside's theorem is often stated in the equivalent form that all groups of order $p^a q^b$, with p, q primes, are solvable.

Exercises

Exercise 6.1. Let G be a non-abelian group of order 39.

1. Determine the degrees of the irreducible representations of G and how many irreducible representations G has of each degree.
2. Determine the number of conjugacy classes of G.

Exercise 6.2. Prove that if there is a non-solvable group of order $p^a q^b$ with p, q primes, then there is a simple non-abelian group of order $p^{a'} q^{b'}$.

Exercise 6.3. Show that if $\varphi\colon G \to GL_d(\mathbb{C})$ is a representation with character χ, then $g \in \ker \varphi$ if and only if $\chi(g) = d$. Hint: Use Corollary 4.1.10 and Lemma 6.3.1.

Chapter 7

Group Actions and Permutation Representations

In this chapter we link representation theory with the theory of group actions and permutation groups. Once again, we are only able to provide a brief glimpse of these connections; see [1] for more. In this chapter all groups are assumed to be finite and all actions of groups are taken to be on finite sets.

7.1 Group actions

Let us begin by recalling the definition of a group action. If X is a set, then S_X will denote the symmetric group on X. We shall tacitly assume $|X| \geq 2$, as the case $|X| = 1$ is uninteresting.

Definition 7.1.1 (Group action)**.** An *action* of a group G on a set X is a homomorphism $\sigma\colon G \to S_X$. We often write σ_g for $\sigma(g)$. The cardinality of X is called the *degree* of the action.

Example 7.1.2 (Regular action)**.** Let G be a group and define $\lambda\colon G \to S_G$ by $\lambda_g(x) = gx$. Then λ is called the *regular action* of G on G.

A subset $Y \subseteq X$ is called G-*invariant* if $\sigma_g(y) \in Y$ for all $y \in Y$, $g \in G$. One can always partition X into a disjoint union of minimal G-invariant subsets called orbits.

Definition 7.1.3 (Orbit)**.** Let $\sigma\colon G \to S_X$ be a group action. The *orbit* of $x \in X$ under G is the set $G \cdot x = \{\sigma_g(x) \mid g \in G\}$.

Clearly the orbits are G-invariant. A standard course in group theory proves that distinct orbits are disjoint and the union of all the orbits is X, that is, the orbits form a partition of X. Of particular importance is the case where there is just one orbit.

Definition 7.1.4 (Transitive)**.** A group action $\sigma\colon G \to S_X$ is *transitive* if, for all $x, y \in X$, there exists $g \in G$ such that $\sigma_g(x) = y$. Equivalently, the action is transitive if there is one orbit of G on X.

Example 7.1.5 (Coset action)**.** If G is a group and H a subgroup, then there is an action $\sigma\colon G \to S_{G/H}$ given by $\sigma_g(xH) = gxH$. This action is transitive.

An even stronger property than transitivity is that of 2-transitivity.

Definition 7.1.6 (2-transitive)**.** An action $\sigma\colon G \to S_X$ of G on X is *2-transitive* if given any two pairs of distinct elements $x, y \in X$ and $x', y' \in X$, there exists $g \in G$ such that $\sigma_g(x) = x'$ and $\sigma_g(y) = y'$.

Example 7.1.7 (Symmetric groups)**.** For $n \geq 2$, the action of S_n on $\{1, \ldots, n\}$ is 2-transitive. Indeed, let $i \neq j$ and $k \neq \ell$ be pairs of elements of X. Let $X = \{1, \ldots, n\} \setminus \{i, j\}$ and $Y = \{1, \ldots, n\} \setminus \{k, \ell\}$. Then $|X| = n - 2 = |Y|$, so we can choose a bijection $\alpha\colon X \to Y$. Define $\tau \in S_n$ by

$$\tau(m) = \begin{cases} k & m = i \\ \ell & m = j \\ \alpha(m) & \text{else.} \end{cases}$$

Then $\tau(i) = k$ and $\tau(j) = \ell$. This establishes that S_n is 2-transitive.

Let's put this notion into a more general context.

Definition 7.1.8 (Orbital)**.** Let $\sigma\colon G \to S_X$ be a transitive group action. Define $\sigma^2\colon G \to S_{X\times X}$ by

$$\sigma^2_g(x_1, x_2) = (\sigma_g(x_1), \sigma_g(x_2)).$$

An orbit of σ^2 is termed an *orbital* of σ. The number of orbitals of is called the *rank* of σ.

Let $\Delta = \{(x, x) \mid x \in X\}$. As $\sigma^2_g(x, x) = (\sigma_g(x), \sigma_g(x))$, it follows from the transitivity of G on X that Δ is an orbital. It is called the *diagonal* or *trivial orbital*.

Remark 7.1.9. Orbitals are closely related to graph theory. If G acts transitively on X, then any non-trivial orbital can be viewed as the edge set of a graph with vertex set X (by symmetrizing). The group G acts on the resulting graph as a vertex-transitive group of automorphisms.

Proposition 7.1.10. *Let $\sigma\colon G \to S_X$ be a group action (with $X \geq 2$). Then σ is 2-transitive if and only if σ is transitive and* $\mathrm{rank}(\sigma) = 2$.

Proof. First we observe that transitivity is necessary for 2-transitivity since if G is 2-transitive on X and $x, y \in X$, then we may choose $x' \neq x$ and $y' \neq y$. By 2-transitivity there exists $g \in G$ with $\sigma_g(x) = y$ and $\sigma_g(x') = y'$. This shows that σ is transitive. Next observe that

$$(X \times X) \setminus \Delta = \{(x, y) \mid x \neq y\}$$

and so the complement of Δ is an orbital if and only for any two pairs $x \neq y$ and $x' \neq y'$ of distinct elements there exists $g \in G$ with $\sigma_g(x) = x'$ and $\sigma_g(y) = y'$, that is, σ is 2-transitive. □

Consequently the rank of S_n is 2. Let $\sigma\colon G \to S_X$ be a group action. Then, for $g \in G$, we define

$$\mathrm{Fix}(g) = \{x \in X \mid \sigma_g(x) = x\}$$

to be the set of *fixed points* of g. Let $\mathrm{Fix}^2(g)$ be the set of fixed points of g on $X \times X$. The notation is unambiguous because of the following proposition.

Proposition 7.1.11. *Let $\sigma\colon G \to S_X$ be a group action. Then the equality*

$$\mathrm{Fix}^2(g) = \mathrm{Fix}(g) \times \mathrm{Fix}(g)$$

holds. Hence $|\mathrm{Fix}^2(g)| = |\mathrm{Fix}(g)|^2$.

Proof. Let $(x, y) \in X \times X$. Then $\sigma_g^2(x, y) = (\sigma_g(x), \sigma_g(y))$ and so $(x, y) = \sigma_g^2(x, y)$ if and only if $\sigma_g(x) = x$ and $\sigma_g(y) = y$. We conclude $\mathrm{Fix}^2(g) = \mathrm{Fix}(g) \times \mathrm{Fix}(g)$. □

7.2 Permutation representations

Given a permutation representation $\sigma\colon G \to S_n$, we may compose it with the standard representation $\alpha\colon S_n \to GL_n(\mathbb{C})$ to obtain a representation of G. Let us formalize this.

Definition 7.2.1 (Permutation representation)**.** Let $\sigma\colon G \to S_X$ be a group action. Define a representation $\widetilde{\sigma}\colon G \to GL(\mathbb{C}X)$ by setting

$$\widetilde{\sigma}_g\left(\sum_{x\in X} c_x x\right) = \sum_{x\in X} c_x\sigma_g(x) = \sum_{y\in X} c_{\sigma_{g^{-1}}(y)} y.$$

One calls $\widetilde{\sigma}$ the *permutation representation* associated to σ.

Remark 7.2.2. Notice that $\widetilde{\sigma}_g$ is the linear extension of the map defined on the basis X of $\mathbb{C}X$ by sending x to $\sigma_g(x)$. Also observe that the degree of the representation $\widetilde{\sigma}$ is the same as the degree of the group action σ.

Example 7.2.3 (Regular representation)**.** Let $\lambda\colon G \to S_G$ be the regular action. Then one has $\widetilde{\lambda} = L$, the regular representation.

The following proposition is proved exactly as in the case of the regular representation, so we omit the proof.

Proposition 7.2.4. *Let $\sigma\colon G \to S_X$ be a group action. Then the permutation representation $\widetilde{\sigma}\colon G \to GL(\mathbb{C}X)$ is a unitary representation of G.*

Next we compute the character of $\widetilde{\sigma}$.

Proposition 7.2.5. *Let $\sigma\colon G \to S_X$ be a group action. Then*

$$\chi_{\widetilde{\sigma}}(g) = |\mathrm{Fix}(g)|.$$

Proof. Let $X = \{x_1, \ldots, x_n\}$ and let $[\widetilde{\sigma}_g]$ be the matrix of $\widetilde{\sigma}$ with respect to this basis. Then $\widetilde{\sigma}_g(x_j) = \sigma_g(x)$ so

$$[\widetilde{\sigma}_g]_{ij} = \begin{cases} 1 & x_i = \sigma_g(x_j) \\ 0 & \text{else.} \end{cases}$$

In particular,

$$\begin{aligned} [\widetilde{\sigma}_g]_{ii} &= \begin{cases} 1 & x_i = \sigma_g(x_i) \\ 0 & \text{else} \end{cases} \\ &= \begin{cases} 1 & x_i \in \mathrm{Fix}(g) \\ 0 & \text{else} \end{cases} \end{aligned}$$

and so $\chi_{\widetilde{\sigma}}(g) = \mathrm{Tr}([\widetilde{\sigma}_g]) = |\mathrm{Fix}(g)|$. □

Like the regular representation, permutation representations are never irreducible (if $|X| > 1$). To understand better how it decomposes, we first consider the trivial component.

Definition 7.2.6 (Fixed subspace)**.** Let $\varphi\colon G \to GL(V)$ be a representation. Then

$$V^G = \{v \in V \mid \varphi_g(v) = v \text{ all } g \in G\}$$

is the fixed subspace of G.

One easily verifies that V^G is a G-invariant subspace and the subrepresentation $\varphi|_{V^G}$ is equivalent to $\dim V^G$ copies of the trivial representation. Let us prove that V^G is the direct sum of all the copies of the trivial representation in φ.

Proposition 7.2.7. *Let $\varphi\colon G \to GL(V)$ be a representation and let χ_1 be the trivial character of G. Then $\langle \chi_1, \chi_\varphi \rangle = \dim V^G$.*

Proof. Write $V = m_1V_1 \oplus \cdots \oplus m_sV_s$ where $V_1, \ldots, V_s$ are irreducible G-invariant subspaces whose associated subrepresentations range over the distinct equivalence classes of irreducible representations of G (we allow $m_i = 0$). Without loss of generality, we may assume that V_1 is equivalent to the trivial representation. Let $\varphi^{(i)}$ be the restriction of φ to V_i. Now if $v \in V$, then $v = v_1 + \cdots + v_s$ with the $v_i \in m_iV_i$ and

$$\varphi_g v = (m_1\varphi^{(1)})_g v_1 + \cdots + (m_s\varphi^{(s)})_g v_s = v_1 + (m_2\varphi^{(2)})_g v_2 + \cdots + (m_s\varphi^{(s)})_g v_s$$

and so $g \in V^G$ if and only if $v_i \in m_iV_i^G$ for all $2 \leq i \leq s$. In other words,

$$V^G = m_1V_1 \oplus m_2V_2^G \oplus \cdots \oplus m_sV_s^G.$$

Let $i \geq 2$. Since V_i is irreducible and not equivalent to the trivial representation and V_i^G is G-invariant, it follows $V_i^G = 0$. Thus $V^G = m_1V_1$ and so the multiplicity of the trivial representation in φ is $\dim V^G$, as required. □

Now we compute $\mathbb{C}X^G$ when we have a permutation representation.

Proposition 7.2.8. *Let $\sigma\colon G \to S_X$ be a group action. Let $\mathcal{O}_1, \ldots, \mathcal{O}_m$ be the orbits of G on X and define $v_i = \sum_{x \in \mathcal{O}_i} x$. Then $v_1, \ldots, v_m$ is a basis for $\mathbb{C}X^G$ and hence $\dim \mathbb{C}X^G$ is the number of orbits of G on X.*

Proof. First observe that

$$\widetilde{\sigma}_g v_i = \sum_{x \in \mathcal{O}_i} \sigma_g(x) = \sum_{y \in \mathcal{O}_i} y = v_i$$

as is seen by setting $y = \sigma_g(x)$ and using that σ_g permutes the orbit $\mathcal{O}_i$. Thus $v_1, \ldots, v_m \in \mathbb{C}X^G$. Since the orbits are disjoint, we have

$$\langle v_i, v_j \rangle = \begin{cases} |\mathcal{O}_i| & i = j \\ 0 & i \neq j \end{cases}$$

and so $\{v_1, \ldots, v_s\}$ is an orthogonal set of non-zero vectors and hence linearly independent. It remain to prove that this set spans $\mathbb{C}X^G$.

Suppose $v = \sum_{x \in X} c_x x \in \mathbb{C}X^G$. We show that if $z \in G \cdot y$, then $c_y = c_z$. Indeed, let $z = \sigma_g(y)$. Then we have

$$\sum_{x \in X} c_x x = v = \widetilde{\sigma}_g v = \sum_{x \in X} c_x \sigma_g(x) \tag{7.1}$$

and so the coefficient of z in the left hand side of (7.1) is c_z while the coefficient of z in the right hand side is c_y since $z = \sigma_g(y)$. Thus $c_z = c_y$. It follows that there are complex numbers $c_1, \ldots, c_m$ such that $c_x = c_i$ all $x \in \mathcal{O}_i$. Then

$$v = \sum_{x \in X} c_x x = \sum_{i=1}^{m} \sum_{x \in \mathcal{O}_i} c_x x = \sum_{i=1}^{m} c_i \sum_{x \in \mathcal{O}_i} x = \sum_{i=1}^{m} c_i v_i$$

and so $v_1, \ldots, v_m$ span $\mathbb{C}X^G$, completing the proof. □

Since G always has at least one orbit on X, the above result shows that the trivial representation appears as a constituent in $\widetilde{\sigma}$ and so if $|X| > 1$, then $\widetilde{\sigma}$ is not irreducible. As a corollary to the above proposition we prove a useful result known as Burnside's lemma, although it seems to have been known earlier to Cauchy and Frobenius. It has many applications in combinatorics to counting problems. The lemma says that the number of orbits of G on X is the average number of fixed points.

Corollary 7.2.9 (Burnside's lemma)**.** *Let $\sigma \colon G \to S_X$ be a group action and let m be the number of orbits of G on X. Then*

$$m = \frac{1}{|G|} \sum_{g \in G} |\mathrm{Fix}(g)|.$$

Proof. Let χ_1 be the trivial character of G. By Propositions 7.2.5, 7.2.7 and 7.2.8 we have

$$m = \langle \chi_1, \chi_{\widetilde{\sigma}} \rangle = \frac{1}{|G|} \sum_{g \in G} \overline{\chi_1(g)} \chi_{\widetilde{\sigma}}(g) = \frac{1}{|G|} \sum_{g \in G} |\mathrm{Fix}(g)|$$

as required. □

As a corollary, we obtain two formulas for the rank of σ.

Corollary 7.2.10. *Let $\sigma\colon G \to S_X$ be a transitive group action. Then the equalities*

$$\operatorname{rank}(\sigma) = \frac{1}{|G|}\sum_{g\in G} |\mathrm{Fix}(g)|^2 = \langle \chi_{\widetilde{\sigma}}, \chi_{\widetilde{\sigma}}\rangle$$

hold.

Proof. Since $\operatorname{rank}(\sigma)$ is the number of orbits of σ^2 on $X \times X$ and the number of fixed points of g on $X \times X$ is $|\mathrm{Fix}(g)|^2$ (Proposition 7.1.11), the first equality is a consequence of Burnside's lemma. For the second we compute

$$\langle \chi_{\widetilde{\sigma}}, \chi_{\widetilde{\sigma}}\rangle = \frac{1}{|G|}\sum_{g\in G} \overline{|\mathrm{Fix}(g)|}|\mathrm{Fix}(g)| = \frac{1}{|G|}\sum_{g\in G} |\mathrm{Fix}(g)|^2$$

completing the proof. □

Assume now that $\sigma\colon G \to S_X$ is a transitive action. Let $v_0 = \sum_{x\in X} x$. Then $\mathbb{C}X^G = \mathbb{C}v_0$ by Proposition 7.2.8. Since $\widetilde{\sigma}$ is a unitary representation, $V_0 = \mathbb{C}v_0^{\perp}$ is a G-invariant subspace (c.f. the proof of Proposition 3.2.3). Usually $\mathbb{C}v_0$ is called the *trace* of σ and V_0 is called the *augmentation* of σ. Let $\widetilde{\sigma}'$ be the restriction of $\widetilde{\sigma}$ to V_0; since $\mathbb{C}X = V_0 \oplus \mathbb{C}v_0$, it follows that $\chi_{\widetilde{\sigma}} = \chi_{\widetilde{\sigma}'} + \chi_1$ where χ_1 is the trivial character. We now characterize when the augmentation representation $\widetilde{\sigma}'$ is irreducible.

Theorem 7.2.11. *Let $\sigma\colon G \to S_X$ be a transitive group action. Then $\widetilde{\sigma}'$ is irreducible if and only if G is 2-transitive on X.*

Proof. This is a simple calculation using Corollary 7.2.10 and the fact that G is 2-transitive on X if and only if $\operatorname{rank}(\sigma) = 2$ (Proposition 7.1.10). Indeed, if χ_1 is the trivial character of G, then

$$\begin{aligned}\langle \chi_{\widetilde{\sigma}'}, \chi_{\widetilde{\sigma}'}\rangle &= \langle \chi_{\widetilde{\sigma}} - \chi_1, \chi_{\widetilde{\sigma}} - \chi_1\rangle \\ &= \langle \chi_{\widetilde{\sigma}}, \chi_{\widetilde{\sigma}}\rangle - \langle \chi_{\widetilde{\sigma}}, \chi_1\rangle - \langle \chi_1, \chi_{\widetilde{\sigma}}\rangle + \langle \chi_1, \chi_1\rangle.\end{aligned} \tag{7.2}$$

Now by Proposition 7.2.8 $\langle \chi_1, \chi_{\widetilde{\sigma}}\rangle = 1$, since G is transitive, and hence $\langle \chi_{\widetilde{\sigma}}, \chi_1\rangle = 1$. Also $\langle \chi_1, \chi_1\rangle = 1$. Thus (7.2) becomes, in light of Corollary 7.2.10,

$$\langle \chi_{\widetilde{\sigma}'}, \chi_{\widetilde{\sigma}'}\rangle = \operatorname{rank}(\sigma) - 1$$

and so $\chi_{\widetilde{\sigma}'}$ is an irreducible character if and only if $\operatorname{rank}(\sigma) = 2$, that is, if and only if G is 2-transitive on X. □

Remark 7.2.12. The decomposition of the standard representation of S_3 in Example 4.3.17 corresponds precisely to the decomposition into the direct sum of the augmentation and the trace.

With Theorem 7.2.11 in hand, we may now compute the character table of S_4.

Example 7.2.13 (Character table of S_4)**.** First of all S_4 has 5 conjugacy classes, represented by $Id, (1\ 2), (1\ 2\ 3), (1\ 2\ 3\ 4), (1\ 2)(3\ 4)$. Let χ_1 be the trivial character and χ_2 the character of the sign homomorphism. Since S_4 acts 2-transitively on $\{1, \ldots, 4\}$, Theorem 7.1.10 implies that the augmentation representation is irreducible. Let χ_4 be the character of this representation; it is the character of the standard representation minus the trivial character so $\chi_4(g) = |\mathrm{Fix}(g)| - 1$. Let $\chi_5 = \chi_2 \cdot \chi_4$. That is if τ is the representation associated to χ_4, then we can define a new representation $\tau^{\chi_2} \colon S_4 \to GL_3(\mathbb{C})$ by $\tau_g^{\chi_2} = \chi_2(g)\tau_g$. It is easily verified that $\chi_{\tau^{\chi_2}}(g) = \chi_2(g)\chi_4(g)$ and τ^{χ_2} is irreducible. This gives us four of the five irreducible representations. How do we get the fifth? Let d be the degree of the missing representation. Then

$$24 = |S_4| = 1^2 + 1^2 + d^2 + 3^2 + 3^2 = 20 + d^2$$

and so $d = 2$. Let χ_3 be the character of the missing irreducible representation and let L be the regular representation of S_4. Then

$$\chi_L = \chi_1 + \chi_2 + 2\chi_3 + 3\chi_4 + 3\chi_5$$

so for $Id \neq g \in S_4$, we have

$$\chi_3(g) = \frac{1}{2}\left(-\chi_1(g) - \chi_2(g) - 3\chi_4(g) - 3\chi_5(g)\right).$$

In this way we are able to produce the character table of S_4 in Table 7.1.

	Id	$(1\ 2)$	$(1\ 2\ 3)$	$(1\ 2\ 3\ 4)$	$(1\ 2)(3\ 4)$
χ_1	1	1	1	1	1
χ_2	1	-1	1	-1	1
χ_3	2	0	-1	0	2
χ_4	3	1	0	-1	-1
χ_5	3	-1	0	1	-1

Table 7.1: Character table of S_4

The reader should try to produce a representation with character χ_3. As a hint, observe that $K = \{Id, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\}$ is a normal subgroup of S_4 and that $S_4/K \cong S_3$. Construct an irreducible representation by composing the surjective map $S_4 \to S_3$ with the degree 2 irreducible representation of S_3 coming from the augmentation representation for S_3.

Exercises

Exercise 7.1. Show that if $\sigma\colon G \to S_X$ is a group action, then the orbits of G on X partition X.

Exercise 7.2. Let $\sigma\colon G \to S_X$ be a transitive group action. If $x \in X$, let

$$G_x = \{g \in G \mid \sigma_g(x) = x\}. \tag{7.3}$$

You may take for granted that G_x is a subgroup of G (called the *stabilizer* of x). Prove that the following are equivalent:

1. G_x is transitive on $X \setminus \{x\}$ for *some* $x \in X$;
2. G_x is transitive on $X \setminus \{x\}$ for *all* $x \in X$;
3. G acts 2-transitively on X.

Exercise 7.3. Compute the character table of A_4. Hints:

1. Let $K = \{Id, (12)(34), (13)(24), (14)(23)\}$. Then K is a normal subgroup of A_4 and $A_4/K \cong \mathbb{Z}_3$. Use this to construct 3 degree one representations of A_4.
2. Show that A_4 acts 2-transitively on $\{1, 2, 3, 4\}$.
3. Conclude that A_4 has 4 conjugacy classes and find them.
4. Produce the character table.

Exercise 7.4. Let G be a group. Define a representation $\lambda\colon G \to GL(L(G))$ by $\lambda_g(f)(h) = f(g^{-1}h)$.

1. Verify that λ is a representation.
2. Prove that λ is equivalent to the regular representation.

3. Let K be a subgroup of G. Let $L(G/K)$ be the subspace of $L(G)$ consisting of functions $f\colon G \to \mathbb{C}$ that are *right K-invariant*, that is, $f(gk) = f(g)$ for all $k \in K$. Show that $L(G/K)$ is a G-invariant subspace of $L(G)$ and that the restriction of λ to $L(G/K)$ is equivalent to the permutation representation $\mathbb{C}(G/K)$.

Exercise 7.5. Two group actions $\sigma\colon G \to S_X$ and $\tau\colon G \to S_Y$ are isomorphic if there is a bijection $\psi\colon X \to Y$ such that $\psi\sigma_g = \tau_g\psi$ for all $g \in G$.

1. Show that if $\tau\colon G \to S_X$ is a transitive group action, $x \in X$ and G_x is the stabilizer of x (c.f. (7.3)), then τ is isomorphic to the coset action $\sigma\colon G \to S_{G/G_x}$.

2. Show that if σ and τ are isomorphic group actions, then the corresponding permutation representations are equivalent.

Exercise 7.6. Suppose that G is a finite group of order n with s conjugacy classes. Suppose that one chooses a pair $(g, h) \in G \times G$ uniformly at random. Prove that the probability g and h commute is s/n. Hint: Apply Burnside's lemma to the action of G on itself by conjugation.

Chapter 8

Induced Representations

If $\psi\colon G \to H$ is a group homomorphism, then from any representation $\varphi\colon H \to GL(V)$ we can obtain a representation $\rho\colon G \to GL(V)$ by composition: set $\rho = \varphi \circ \psi$. If ψ is onto and φ is irreducible, one can verify that ρ will also be irreducible. Lemma 6.2.6 shows that every degree one representation of G is obtained in this way by taking $\psi\colon G \to G/G'$. As G/G' is abelian, in principle, we know how to compute all its irreducible representations. Now we would like to consider the dual situation: suppose H is a subgroup of G; how can we construct a representation of G from a representation of H? There is a method to do this, due to Frobenius, via a procedure called induction. This is particularly useful when applied to abelian subgroups since we know how to construct all representations of an abelian group.

8.1 Induced characters and Frobenius reciprocity

We use the notation $H \leq G$ to indicate that H is a subgroup of G. Our goal is to first define the induced character on G associated to a character on H. This induced character will be a class function; we'll worry later about showing that it is actually the character of a representation. If $f\colon G \to \mathbb{C}$ is a function, then we can restrict f to H to obtain a map $\mathrm{Res}^G_H f\colon H \to \mathbb{C}$ called *restriction*. So $\mathrm{Res}^G_H f(h) = f(h)$ for $h \in H$.

Proposition 8.1.1. *Let* $H \leq G$. *Then* $\mathrm{Res}^G_H\colon Z(L(G)) \to Z(L(H))$ *is a linear map.*

Proof. First we need to verify that if $f\colon G \to \mathbb{C}$ is a class function, then so is $\mathrm{Res}^G_H f$. Indeed, if $x, h \in H$, then $\mathrm{Res}^G_H f(xhx^{-1}) = f(xhx^{-1}) = f(h) =$

$\mathrm{Res}^G_H f(h)$ since f is a class function. Linearity is immediate:

$$\mathrm{Res}^G_H(c_1f_1 + c_2f_2)(h) = c_1f_1(h) + c_2f_2(h) = c_1 \mathrm{Res}^G_H f_1(h) + c_2 \mathrm{Res}^G_H f_2(h).$$

This completes the proof. □

Our goal now is to construct a linear map $Z(L(H)) \to Z(L(G))$ going the other way. First we need a piece of notation. If $H \leq G$ and $f\colon H \to \mathbb{C}$ is a function, let us define $\dot{f}\colon G \to \mathbb{C}$ by

$$\dot{f}(x) = \begin{cases} f(x) & x \in H \\ 0 & x \notin H. \end{cases}$$

The reader should verify that the assignment $f \mapsto \dot{f}$ is a linear map from $L(H)$ to $L(G)$. Let us now define a map $\mathrm{Ind}^G_H\colon Z(L(H)) \to Z(L(G))$, called *induction*, by the formula

$$\mathrm{Ind}^G_H f(g) = \frac{1}{|H|}\sum_{x\in G} \dot{f}(x^{-1}gx).$$

In the case χ is a character of H, one calls $\mathrm{Ind}^G_H \chi$ the *induced character* on G.

Proposition 8.1.2. *Let $H \leq G$. Then the map*

$$\mathrm{Ind}^G_H\colon Z(L(H)) \to Z(L(G))$$

is linear.

Proof. First we verify that $\mathrm{Ind}^G_H f$ is really a class function. Let $y, g \in G$, then

$$\mathrm{Ind}^G_H f(y^{-1}gy) = \frac{1}{|H|}\sum_{x\in G} \dot{f}(x^{-1}y^{-1}gyx) = \frac{1}{|H|}\sum_{z\in G} \dot{f}(z^{-1}gz) = \mathrm{Ind}^G_H f(g)$$

where the penultimate equality follows by setting $z = yx$. Next we check linearity. Indeed, we compute

$$\begin{aligned}\mathrm{Ind}^G_H(c_1f_1 + c_2f_2)(g) &= \frac{1}{|H|}\sum_{x\in G} \overbrace{c_1f_2 + c_2f_2}^{\cdot}(x^{-1}gx) \\ &= c_1\frac{1}{|H|}\sum_{x\in G} \dot{f}_1(x^{-1}gx) + c_2\frac{1}{|H|}\sum_{x\in G}\dot{f}_2(x^{-1}gx) \\ &= c_1 \mathrm{Ind}^G_H f_1(g) + c_2 \mathrm{Ind}^G_H f_2(g)\end{aligned}$$

establishing the linearity of the induction map. □

The following theorem, known as Frobenius reciprocity, asserts that the linear maps Res^G_H and Ind^G_H are adjoint. What it says is in practice is that if χ is an irreducible character of G and θ is an irreducible character of H, then the multiplicity of χ in the induced character $\mathrm{Ind}^G_H \theta$ is exactly the same as the multiplicity of θ in $\mathrm{Res}^G_H \chi$.

Theorem 8.1.3 (Frobenius reciprocity)**.** *Suppose that H is a subgroup of G and let a be a class function on G and b be a class function on H. Then the formula*

$$\langle \mathrm{Res}^G_H a, b\rangle = \langle a, \mathrm{Ind}^G_H b\rangle$$

holds.

Proof. We begin by computing

$$\begin{aligned}\langle a, \mathrm{Ind}^G_H b\rangle &= \frac{1}{|G|}\sum_{g\in G}\overline{a(g)}\,\mathrm{Ind}^G_H b(g) &(8.1)\\ &= \frac{1}{|G|}\sum_{g\in G}\overline{a(g)}\frac{1}{|H|}\sum_{x\in G}\dot{b}(x^{-1}gx) &(8.2)\\ &= \frac{1}{|G|}\frac{1}{|H|}\sum_{x\in G}\sum_{g\in G}\overline{a(g)}\dot{b}(x^{-1}gx). &(8.3)\end{aligned}$$

Now in order for $\dot{b}(x^{-1}gx)$ not to be 0, we need $x^{-1}gx \in H$, that is, we need $g = xhx^{-1}$ with $h \in H$. This allows us to re-index the sum in (8.3) as

$$\begin{aligned}\frac{1}{|G|}\frac{1}{|H|}\sum_{x\in G}\sum_{h\in H}\overline{a(xhx^{-1})}b(h) &= \frac{1}{|G|}\frac{1}{|H|}\sum_{x\in G}\sum_{h\in H}\overline{a(h)}b(h)\\ &= \frac{1}{|G|}\sum_{x\in G}\langle \mathrm{Res}^G_H a, b\rangle\\ &= \langle \mathrm{Res}^G_H a, b\rangle\end{aligned}$$

where the first equality uses that a is a class function on G. □

The following formula for induction in terms of coset representatives is often extremely useful, especially for computational purposes.

Proposition 8.1.4. *Let G be a group and H a subgroup of G. Let $t_1, \ldots, t_m$ be a complete set of representatives of the left cosets of H in G. Then the formula*

$$\mathrm{Ind}^G_H f(g) = \sum_{i=1}^{m} \dot{f}(t_i^{-1}gt_i)$$

holds for any class function f on H.

Proof. Using that G is the disjoint union $t_1H \cup \cdots \cup t_mH$, we obtain

$$\operatorname{Ind}_H^G f(g) = \frac{1}{|H|} \sum_{x \in G} \dot{f}(x^{-1}gx) = \frac{1}{|H|} \sum_{i=1}^{m} \sum_{h \in H} \dot{f}(h^{-1}t_i^{-1}gt_ih). \tag{8.4}$$

Now if $h \in H$, then $h^{-1}t_i^{-1}gt_ih \in H$ if and only if $t_i^{-1}gt_i \in H$. Since f is a class function on H, it follows that $\dot{f}(h^{-1}t_i^{-1}gt_ih) = \dot{f}(t_i^{-1}gt_i)$ and so the right hand side of (8.4) equals

$$\frac{1}{|H|} \sum_{i=1}^{m} \sum_{h \in H} \dot{f}(t_i^{-1}gt_i) = \frac{1}{|H|} \sum_{h \in H} \sum_{i=1}^{m} \dot{f}(t_i^{-1}gt_i) = \sum_{i=1}^{m} \dot{f}(t_i^{-1}gt_i)$$

completing the proof. □

8.2 Induced representations

If $\varphi\colon G \to GL(V)$ is a representation of G and $H \leq G$, then we can restrict φ to H to obtain a representation $\operatorname{Res}_H^G \varphi\colon H \to GL(V)$. Since, for $h \in H$,

$$\chi_{\operatorname{Res}_H^G \varphi}(h) = \operatorname{Tr}(\operatorname{Res}_H^G \varphi(h)) = \operatorname{Tr}(\varphi(h)) = \chi_\varphi(h) = \operatorname{Res}_H^G \chi_\varphi(h)$$

it follows that $\chi_{\operatorname{Res}_H^G \varphi} = \operatorname{Res}_H^G \chi_\varphi$. Thus the restriction map sends characters to characters. In this section, we show that induction also sends characters to characters, but the construction in this case is much more complicated. Let's look at some examples to see why this might indeed be the case.

Example 8.2.1 (Regular representation)**.** Let χ_1 be the trivial character of the trivial subgroup $\{1\}$ of G. Then

$$\operatorname{Ind}_{\{1\}}^G \chi_1(g) = \sum_{x \in G} \dot{\chi}_1(x^{-1}gx),$$

but $x^{-1}gx \in \{1\}$ if and only if $g = 1$. Thus

$$\operatorname{Ind}_{\{1\}}^G \chi_1(g) = \begin{cases} |G| & g = 1 \\ 0 & g \neq 1, \end{cases}$$

i.e., $\operatorname{Ind}_{\{1\}}^G \chi_1$ is the character of the regular representation of G.

This example can be generalized.

Example 8.2.2 (Permutation representations)**.** Let $H \leq G$ and consider the associated group action $\sigma\colon G \to S_{G/H}$ given by $\sigma_g(xH) = gxH$. Notice that $xH \in \mathrm{Fix}(g)$ if and only if $gxH = xH$, that is, $x^{-1}gx \in H$. Now there are $|H|$ elements x giving the coset xH so $|\mathrm{Fix}(g)|$ is $1/|H|$ times the number of $x \in G$ such that $x^{-1}gx \in H$. Let χ_1 be the trivial character of H. Then

$$\dot{\chi}_1(x^{-1}gx) = \begin{cases} 1 & x^{-1}gx \in H \\ 0 & x^{-1}gx \notin H \end{cases}$$

and so we deduce

$$\chi_{\widetilde{\sigma}}(g) = |\mathrm{Fix}(g)| = \frac{1}{|H|}\sum_{x\in G}\dot{\chi}_1(x^{-1}gx) = \mathrm{Ind}_H^G\,\chi_1(g)$$

showing that $\mathrm{Ind}_H^G\,\chi_1$ is the character of the permutation representation $\widetilde{\sigma}$.

Fix now a group G and a subgroup H. Let $m = [G : H]$ be the index of H in G. Choose a complete set of representatives $t_1, \ldots, t_m$ of the left cosets of H in G. Without loss of generality we may always take $t_1 = 1$. Suppose $\varphi\colon H \to GL_d(\mathbb{C})$ is a representation of H. Let us introduce a dot notation in this context by setting

$$\dot{\varphi}_x = \begin{cases} \varphi_x & x \in H \\ 0 & x \notin H \end{cases}$$

where 0 is the $d \times d$ zero matrix. We now may define a representation $\mathrm{Ind}_H^G\,\varphi\colon G \to GL_{md}(\mathbb{C})$, called the *induced representation*, as follows. First, for ease of notation, we write φ^G for $\mathrm{Ind}_H^G\,\varphi$. Then, for $g \in G$, we construct φ^G_g as an $m \times m$ block matrix with $d \times d$ blocks by setting $[\varphi^G_g]_{ij} = \dot{\varphi}_{t_i^{-1}gt_j}$ for $1 \leq i, j \leq m$. In matrix form we have

$$\varphi^G_g = \begin{bmatrix} \dot{\varphi}_{t_1^{-1}gt_1} & \dot{\varphi}_{t_1^{-1}gt_2} & \cdots & \dot{\varphi}_{t_1^{-1}gt_m} \\ \dot{\varphi}_{t_2^{-1}gt_1} & \dot{\varphi}_{t_2^{-1}gt_2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \dot{\varphi}_{t_{m-1}^{-1}gt_m} \\ \dot{\varphi}_{t_m^{-1}gt_1} & \cdots & \dot{\varphi}_{t_m^{-1}gt_{m-1}} & \dot{\varphi}_{t_m^{-1}gt_m} \end{bmatrix}.$$

Before proving that $\mathrm{Ind}_H^G\,\varphi$ is indeed a representation, let's look at some examples.

Example 8.2.3 (Dihedral groups)**.** Let $G = D_n$, the dihedral group of order $2n$. If r is a rotation by $2\pi/n$ and s is a reflection, then $D_n = \{r^m, sr^m \mid 0 \leq m \leq n-1\}$. Let $H = \langle r \rangle$; so H is a cyclic subgroup of order n and index 2. For $0 \leq k \leq n-1$, let $\chi_k \colon H \to \mathbb{C}^*$ be the representation given by $\chi_k(r^m) = e^{2\pi i k m/n}$. Let's compute the induced representation $\varphi^{(k)} = \operatorname{Ind}_H^G \chi_k$. We choose coset representatives $t_1 = 1$ and $t_2 = s$. Then

$$\begin{aligned} t_1^{-1} r^m t_1 &= r^m & t_1^{-1} s r^m t_1 &= s r^m \\ t_1^{-1} r^m t_2 &= r^m s = s r^{-m} & t_1^{-1} s r^m t_2 &= s r^m s = r^{-m} \\ t_2^{-1} r^m t_1 &= s r^m & t_2^{-1} s r^m t_1 &= r^m \\ t_2^{-1} r^m t_2 &= r^{-m} & t_2^{-1} s r^m t_2 &= r^m s = s r^{-m} \end{aligned}$$

and so we obtain

$$\begin{aligned} \varphi^{(k)}_{r^m} &= \begin{bmatrix} \dot{\chi}_k(r^m) & \dot{\chi}_k(sr^{-m}) \\ \dot{\chi}_k(sr^m) & \dot{\chi}_k(r^{-m}) \end{bmatrix} = \begin{bmatrix} e^{2\pi i k m/n} & 0 \\ 0 & e^{-2\pi i k m/n} \end{bmatrix} \\ \varphi^{(k)}_{sr^m} &= \begin{bmatrix} \dot{\chi}_k(sr^m) & \dot{\chi}_k(r^{-m}) \\ \dot{\chi}_k(r^m) & \dot{\chi}_k(sr^{-m}) \end{bmatrix} = \begin{bmatrix} 0 & e^{-2\pi i k m/n} \\ e^{2\pi i k m/n} & 0 \end{bmatrix}. \end{aligned}$$

In particular, $\operatorname{Ind}_H^G \chi_k(r^m) = 2\cos(2\pi k m/n)$ and $\operatorname{Ind}_H^G \chi_k(sr^m) = 0$. It is easy to verify that $\varphi^{(k)}$ is irreducible for $1 \leq k < \frac{n}{2}$ and that this range of values gives inequivalent irreducible representations. Note that $\operatorname{Ind}_H^G \chi_k = \operatorname{Ind}_H^G \chi_{n-k}$, so there is no need to consider $k > \frac{n}{2}$. One can show that the $\varphi^{(k)}$ cover all the equivalence classes of irreducible representations of D_n except for the degree one representations. If n is odd there are two degree one characters while if n is even there are four degree one representations.

Example 8.2.4 (Quaternions)**.** Let $Q = \{\pm 1, \pm\hat{\imath}, \pm\hat{\jmath}, \pm\hat{k}\}$ be the group of quaternions. Here -1 is central and the rules $\hat{\imath}^2 = \hat{\jmath}^2 = \hat{k}^2 = \hat{\imath}\hat{\jmath}\hat{k} = -1$ are valid. One can verify that $Q' = \{\pm 1\}$ and that $Q/Q' \cong \mathbb{Z}_2 \times Z_2$. Thus Q has four degree one representations. Since each representation has degree dividing 8 and the sum of the squares of the degrees is 8, there can be only one remaining irreducible representation and it must have degree 2. Let's construct it as an induced representation. Let $H = \langle \hat{\imath} \rangle$. Then $|H| = 4$ and so $[Q : H] = 2$. Consider the representation $\varphi \colon H \to \mathbb{C}^*$ given by $\varphi(\hat{\imath}^k) = i^k$. Let $t_1 = 1$ and $t_2 = \hat{\jmath}$. Then one can compute

$$\begin{aligned} \varphi^Q_{\pm 1} &= \pm \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, & \varphi^Q_{\pm\hat{\imath}} &= \pm \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \\ \varphi^Q_{\pm\hat{\jmath}} &= \pm \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, & \varphi^Q_{\pm\hat{k}} &= \pm \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}. \end{aligned}$$

It is easy to see that φ^Q is irreducible since $\varphi^Q_{\hat{\imath}}$ and $\varphi^Q_{\hat{\jmath}}$ have no common eigenvector. The character table of Q appears in Table 8.1.

	1	-1	$\hat{\imath}$	$\hat{\jmath}$	$\hat{k}$
χ_1	1	1	1	1	1
χ_2	1	1	1	-1	-1
χ_3	1	1	-1	1	-1
χ_4	1	1	-1	-1	1
χ_5	2	-2	0	0	0

Table 8.1: Character table of the quaternions

We are now ready to prove that $\operatorname{Ind}_H^G \varphi$ is a representation with character $\operatorname{Ind}_H^G \chi_\varphi$.

Theorem 8.2.5. *Let H be a subgroup of G of index m and suppose that $\varphi\colon H \to GL_d(\mathbb{C})$ is a representation of H. Then $\operatorname{Ind}_H^G \varphi\colon G \to GL_{md}(\mathbb{C})$ is a representation and $\chi_{\operatorname{Ind}_H^G \varphi} = \operatorname{Ind}_H^G \chi_\varphi$. In particular, Ind_H^G maps characters to characters.*

Proof. Let $t_1, \ldots, t_m$ be a complete set of representatives for the cosets of H in G. Set $\varphi^G = \operatorname{Ind}_H^G \varphi$. We begin by showing that φ^G is a representation. Let $x, y \in G$. Then we have

$$[\varphi^G_x \varphi^G_y]_{ij} = \sum_{k=1}^{m} [\varphi^G_x]_{ik} [\varphi^G_y]_{kj} = \sum_{k=1}^{m} \dot{\varphi}_{t_i^{-1}xt_k} \dot{\varphi}_{t_k^{-1}yt_j}. \tag{8.5}$$

The only way $\dot{\varphi}_{t_k^{-1}yt_j} \neq 0$ is if $t_k^{-1}yt_j \in H$, or equivalently $t_kH = yt_jH$. So if t_ℓ is the representative of the coset yt_jH, then the right hand side of (8.5) becomes $\dot{\varphi}_{t_i^{-1}xt_\ell}\varphi_{t_\ell^{-1}yt_j}$. This in turn is non-zero if and only if $t_i^{-1}xt_\ell \in H$, that is, $t_iH = xt_\ell H = xyt_jH$ or equivalently $t_i^{-1}xyt_j \in H$. If this is the case, then the right hand side of (8.5) equals

$$\varphi_{t_i^{-1}xt_\ell}\varphi_{t_\ell^{-1}yt_j} = \varphi_{t_i^{-1}xyt_j}$$

and hence $[\varphi^G_x\varphi^G_y]_{ij} = \dot{\varphi}_{t_i^{-1}xyt_j} = [\varphi^G_{xy}]_{ij}$, establishing that φ^G is a homomorphism from G to $M_{md}(\mathbb{C})$. Next observe that $[\varphi^G_1]_{ij} = \dot{\varphi}_{t_i^{-1}t_j}$, but $t_i^{-1}t_j \in H$ implies $t_iH = t_jH$, which in turn implies $t_i = t_j$. Thus

$$[\varphi^G_1]_{ij} = \begin{cases} \varphi_1 = I & i = j \\ 0 & i \neq j \end{cases}$$

and so $\varphi_1^G = I$. Therefore, if $g \in G$ then $\varphi_g^G \varphi_{g^{-1}}^G = \varphi_{gg^{-1}}^G = \varphi_1^G = I$ establishing that $(\varphi_g^G)^{-1} = \varphi_{g^{-1}}^G$ and therefore φ^G is a representation. Let's compute its character.

Applying Proposition 8.1.4 we obtain

$$\chi_{\varphi^G}(g) = \mathrm{Tr}(\varphi_g^G) = \sum_{i=1}^{m} \mathrm{Tr}(\dot{\varphi}_{t_i^{-1}gt_i}) = \sum_{i=1}^{m} \dot{\chi}_\varphi(t_i^{-1}gt_i) = \mathrm{Ind}_H^G \chi_\varphi$$

as required. □

8.3 Mackey's irreducibility criterion

There is no guarantee that if χ is an irreducible character of H, then $\mathrm{Ind}_H^G \chi$ will be an irreducible character of G. For instance, $L = \mathrm{Ind}_{\{1\}}^G \chi_1$ is not irreducible. On the other hand, sometimes induced characters are irreducible, as we saw with the dihedral groups and the quaternions. There is a criterion, due to Mackey, describing when an induced character is irreducible. This is the subject of this section. By Frobenius reciprocity,

$$\langle \mathrm{Ind}_H^G \chi_\varphi, \mathrm{Ind}_H^G \chi_\varphi \rangle = \langle \mathrm{Res}_H^G \mathrm{Ind}_H^G \chi_\varphi, \chi_\varphi \rangle$$

and so our problem amounts to understanding $\mathrm{Res}_H^G \mathrm{Ind}_H^G \chi_\varphi$.

Definition 8.3.1 (Disjoint representations)**.** Two representations φ and ρ of G are said to be *disjoint* if they have no common irreducible constituents.

Proposition 8.3.2. *Representations φ and ρ of G are disjoint if and only if χ_φ and χ_ρ are orthogonal.*

Proof. Let $\varphi^{(1)}, \ldots, \varphi^{(s)}$ be a complete set of representatives of the equivalence classes of irreducible representations of G. Then

$$\varphi \sim m_1\varphi^{(1)} + \cdots + m_s\varphi^{(s)}$$
$$\rho \sim n_1\varphi^{(1)} + \cdots + n_s\varphi^{(s)}$$

for certain non-negative integers m_i, n_i. From the orthonormality of irreducible characters, we obtain

$$\langle \chi_\varphi, \chi_\rho \rangle = m_1n_1 + \cdots + m_sn_s. \tag{8.6}$$

Clearly the right hand side of (8.6) is 0 if and only if $m_in_i = 0$ all i, if and only if φ and ρ are disjoint. □

To understand $\operatorname{Res}^G_H \operatorname{Ind}^G_H f$, it turns out not much more difficult to analyze $\operatorname{Res}^G_H \operatorname{Ind}^G_K f$ where H, K are two subgroups of G. To perform this analysis we need the notion of a double coset.

Definition 8.3.3 (Double coset)**.** Let H, K be subgroups of a group G. Then define a group action $\sigma\colon H \times K \to G$ by $\sigma_{(h,k)}(g) = hgk^{-1}$. The orbit of g under $H \times K$ is then the set

$$HgK = \{hgk \mid h \in H, k \in K\}$$

and is called a *double coset* of g. We write $H\backslash G/K$ for the set of double cosets of H and K in G.

Notice that the double cosets are disjoint and have union G. Also, if H is a normal subgroup of G, then $H\backslash G/H = G/H$.

Example 8.3.4. Let $G = GL_2(\mathbb{C})$ and B be the group of invertible 2×2 upper triangular matrices over $\mathbb{C}$. Then $B\backslash G/B = \left\{B, B\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}B\right\}$.

The following theorem of Mackey explains how induction and restriction of class functions from different subgroups interact.

Theorem 8.3.5 (Mackey)**.** *Let $H, K \leq G$ and let S be a complete set of double coset representatives for $H\backslash G/K$. Then, for $f \in Z(L(K))$,*

$$\operatorname{Res}^G_H \operatorname{Ind}^G_K f = \sum_{s \in S} \operatorname{Ind}^H_{H \cap sKs^{-1}} \operatorname{Res}^{sKs^{-1}}_{H \cap sKs^{-1}} f^s$$

where $f^s \in Z(L(sKs^{-1}))$ is given by $f^s(x) = f(s^{-1}xs)$.

Proof. For this proof it is important to construct the correct set T of left coset representatives for K in G. Choose, for each $s \in S$, a complete set V_s of representatives of the left cosets of $H \cap sKs^{-1}$ in H. Then $H = \bigcup_{v \in V_S} v(H \cap sKs^{-1})$ and the union is disjoint. Now

$$HsK = HsKs^{-1}s = \bigcup_{v \in V_s} v(H \cap sKs^{-1})sKs^{-1}s = \bigcup_{v \in V_s} vsK$$

and moreover this union is disjoint. Indeed, if $vsK = v'sK$ with $v, v' \in V_s$, then $s^{-1}v^{-1}v's \in K$ and so $v^{-1}v' \in sKs^{-1}$. But also $v, v' \in H$ so $v^{-1}v' \in H \cap sKs^{-1}$ and hence $v(H \cap sKs^{-1}) = v'(H \cap sKs^{-1})$ and so $v = v'$ by definition of V_s.

Let $T_s = \{vs \mid v \in V_s\}$ and let $T = \bigcup_{s\in S} T_s$. This latter union is disjoint since if $vs = v's'$ for $v \in V_s$ and $v' \in V_{s'}$, then $HsK = Hs'K$ and so $s = s'$, as S is a complete set of double coset representatives, and therefore $v = v'$. Putting this together we have

$$G = \bigcup_{s\in S} HsK = \bigcup_{s\in S}\bigcup_{v\in V_s} vsK = \bigcup_{s\in S}\bigcup_{t\in T_s} tK = \bigcup_{t\in T} tK$$

and all these unions are disjoint. Therefore, T is a complete set of representatives for the left cosets of K in G.

Using Proposition 8.1.4, for $h \in H$, we compute

$$\begin{aligned}
\operatorname{Ind}_K^G f(h) &= \sum_{t\in T} \dot{f}(t^{-1}ht) \\
&= \sum_{s\in S}\sum_{t\in T_s} \dot{f}(t^{-1}ht) \\
&= \sum_{s\in S}\sum_{v\in V_s} \dot{f}(s^{-1}v^{-1}hvs) \\
&= \sum_{s\in S}\sum_{\substack{v\in V_s,\\ v^{-1}hv\in sKs^{-1}}} f^s(v^{-1}hv) \\
&= \sum_{s\in S}\sum_{\substack{v\in V_s,\\ v^{-1}hv\in H\cap sKs^{-1}}} \operatorname{Res}^{sKs^{-1}}_{H\cap sKs^{-1}} f^s(v^{-1}hv) \\
&= \sum_{s\in S} \operatorname{Ind}^H_{H\cap sKs^{-1}} \operatorname{Res}^{sKs^{-1}}_{H\cap sKs^{-1}} f^s
\end{aligned}$$

again by an application of Proposition 8.1.4. This completes the proof. □

From Theorem 8.3.5, we can obtain Mackey's irreducibility criterion.

Theorem 8.3.6 (Mackey's irreducibility criterion). *Let H be a subgroup of G and let $\varphi\colon H \to GL_d(\mathbb{C})$ be a representation. Then $\operatorname{Ind}_H^G \varphi$ is irreducible if and only if:*

1. *φ is irreducible;*

2. *the representations $\operatorname{Res}^H_{H\cap sHs^{-1}} \varphi$ and $\operatorname{Res}^{sHs^{-1}}_{H\cap sHs^{-1}} \varphi^s$ are disjoint*

for all $s \notin H$, where $\varphi^s(x) = \varphi(s^{-1}xs)$ for $x \in sHs^{-1}$.

Proof. Let χ be the character of φ. Let S be a complete set of double coset representatives of $H\backslash G/H$. Assume without loss of generality that $1 \in S$. Then, for $s = 1$, notice that $H \cap sHs^{-1} = H$, $\varphi^s = \varphi$. Let $S^\sharp = S \setminus \{1\}$. Theorem 8.3.5 then yields

$$\mathrm{Res}^G_H \mathrm{Ind}^G_H \chi = \chi + \sum_{s\in S^\sharp} \mathrm{Ind}^H_{H\cap sHs^{-1}} \mathrm{Res}^{sHs^{-1}}_{H\cap sHs^{-1}} \chi^s.$$

Applying Frobenius reciprocity twice, we obtain

$$\begin{aligned}\langle \mathrm{Ind}^G_H \chi, \mathrm{Ind}^G_H \chi\rangle &= \langle \mathrm{Res}^G_H \mathrm{Ind}^G_H \chi, \chi\rangle \\ &= \langle \chi, \chi\rangle + \sum_{s\in S^\sharp} \langle \mathrm{Ind}^H_{H\cap sHs^{-1}} \mathrm{Res}^{sHs^{-1}}_{H\cap sHs^{-1}} \chi^s, \chi\rangle \\ &= \langle \chi, \chi\rangle + \sum_{s\in S^\sharp} \langle \mathrm{Res}^{sHs^{-1}}_{H\cap sHs^{-1}} \chi^s, \mathrm{Res}^H_{H\cap sHs^{-1}} \chi\rangle.\end{aligned}$$

Since $\langle \chi, \chi\rangle \geq 1$ and all terms in the sum are non-negative, we see that the inner product $\langle \mathrm{Ind}^G_H \chi, \mathrm{Ind}^G_H \chi\rangle$ is 1 if and only if $\langle \chi, \chi\rangle = 1$ and

$$\langle \mathrm{Res}^{sHs^{-1}}_{H\cap sHs^{-1}} \chi^s, \mathrm{Res}^H_{H\cap sHs^{-1}} \chi\rangle = 0$$

all $s \in S^\sharp$. Thus $\mathrm{Ind}^G_H \varphi$ is irreducible if and only if φ is irreducible and the representations $\mathrm{Res}^{sHs^{-1}}_{H\cap sHs^{-1}} \varphi^s$ and $\mathrm{Res}^H_{H\cap sHs^{-1}} \varphi$ are disjoint for all $s \in S^\sharp$. Now any $s \notin H$ can be an element of $S^\sharp$ for an appropriately chosen set S of double coset representatives, from which the theorem follows. □

Remark 8.3.7. The proof shows that one need only check that 2 holds for all $s \notin H$ from a given set of double coset representatives, which is often easier to deal with in practice.

Mackey's criterion is most readily applicable for normal subgroups. If $H \lhd G$ is a normal subgroup, then $H\backslash G/H = G/H$ and $H \cap sHs^{-1} = H$. So Mackey's criterion in this case boils down to checking that $\varphi\colon H \to GL_d(\mathbb{C})$ is irreducible and that $\varphi^s\colon H \to GL_d(\mathbb{C})$ does not have φ as an irreducible constituent for $s \notin H$. Actually, one can easily verify that φ^s is irreducible if and only if φ is irreducible, so basically one just has to check that φ and φ^s are inequivalent irreducible representations when $s \notin H$. In fact, one just needs to check this as s ranges over a complete set of coset representatives for G/H.

Example 8.3.8. Let p be a prime and let

$$G = \left\{ \begin{bmatrix} \overline{a} & \overline{b} \\ \overline{0} & \overline{1} \end{bmatrix} \mid \overline{a} \in \mathbb{Z}_p^*, \overline{b} \in \mathbb{Z}_p \right\},$$
$$H = \left\{ \begin{bmatrix} \overline{1} & \overline{b} \\ \overline{0} & \overline{1} \end{bmatrix} \mid \overline{b} \in \mathbb{Z}_p \right\}.$$

Then $H \cong \mathbb{Z}_p$, $H \lhd G$ and $G/H \cong \mathbb{Z}_p^*$ (consider the projection to the upper left corner). A complete set of coset representatives is

$$S = \left\{ \begin{bmatrix} \overline{a} & \overline{0} \\ \overline{0} & \overline{1} \end{bmatrix} \mid \overline{a} \in \mathbb{Z}_p^* \right\}.$$

Let $\varphi\colon H \to \mathbb{C}^*$ be given by

$$\varphi\left(\begin{bmatrix} \overline{1} & \overline{b} \\ \overline{0} & \overline{1} \end{bmatrix} \right) = e^{2\pi i b/p}.$$

Then if $s = \begin{bmatrix} \overline{a}^{-1} & \overline{0} \\ \overline{0} & \overline{1} \end{bmatrix}$ with $\overline{a} \neq 1$, we have

$$\varphi^s\left(\begin{bmatrix} \overline{1} & \overline{b} \\ \overline{0} & \overline{1} \end{bmatrix} \right) = \varphi\left(\begin{bmatrix} \overline{1} & \overline{a}\overline{b} \\ \overline{0} & \overline{1} \end{bmatrix} \right) = e^{2\pi i ab/p}$$

and so φ, φ^s are inequivalent irreducible representations of H. Mackey's criterion now implies that $\operatorname{Ind}_H^G \varphi$ is an irreducible representation of G of degree $[G : H] = p - 1$. Notice that

$$p - 1 + (p-1)^2 = (p-1)[1 + p - 1] = (p-1)p = |G|.$$

Since one can lift the $p - 1$ degree one representations of $G/H \cong \mathbb{Z}_p^*$ to G, the above computation implies that $\operatorname{Ind}_H^G \varphi$ and the $p - 1$ degree one representations are all the irreducible representations of G.

Exercises

Exercise 8.1. Prove that if $G = GL_2(\mathbb{C})$ and $B = \left\{ \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \mid ac \neq 0 \right\}$, then $B \backslash G / B = \left\{ B, B \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} B \right\}$. Prove that G/B is infinite.

Exercise 8.2. Let G be a group with a cyclic normal subgroup $H = \langle a \rangle$ of order k. Suppose that $N_G(a) = H$, that is, $sa = as$ implies $s \in H$. Show that if $\chi\colon H \to \mathbb{C}^*$ is the character given by $\chi(a^m) = e^{2\pi i m/k}$, then $\mathrm{Ind}_H^G \chi$ is an irreducible character of G.

Exercise 8.3.

1. Construct the character table for the dihedral group D_4 of order 8. Suppose that s is the reflection over the x-axis and r is rotation by $\pi/2$. Hint: Observe that $Z = \{1, r^2\}$ is in the center of D_4 (actually, it is the center) and $D_4/Z \cong \mathbb{Z}_2 \times \mathbb{Z}_2$. Use this to get the degree one characters. Get the degree two character as an induced character from a representation of the subgroup $\langle r \rangle \cong \mathbb{Z}_4$.

2. Is the action of D_4 on the vertices of the square two-transitive?

Exercise 8.4. Compute the character table for the group in Example 8.3.8 of the text when $p = 5$.

Exercise 8.5. Let N be a normal subgroup of a group G and suppose that $\varphi\colon N \to GL_d(\mathbb{C})$ is a representation. For $s \in G$, define $\varphi^s\colon N \to GL_d(\mathbb{C})$ by $\varphi^s(n) = \varphi(s^{-1}ns)$. Prove that φ is irreducible if and only if φ^s is irreducible.

Exercise 8.6. Show that if G is a non-abelian group and $\varphi\colon Z(G) \to GL_d(\mathbb{C})$ is an irreducible representation of the center of G, then $\mathrm{Ind}_{Z(G)}^G \varphi$ is not irreducible.

Exercise 8.7. A representation is called *faithful* if it is one-to-one.

1. Let H be a subgroup of G and suppose $\varphi\colon H \to GL_d(\mathbb{C})$ is a faithful representation. Show that $\varphi^G = \mathrm{Ind}_H^G \varphi$ is a faithful representation of G.

2. Show that every representation of a simple group which is not a direct sum of copies of the trivial representation is faithful.

Exercise 8.8. Let G be a group and let H be a subgroup. Let $\sigma\colon G \to S_{G/H}$ be the group action given by $\sigma_g(xH) = gxH$.

1. Show that σ is transitive.

2. Show that H is the stabilizer of the coset H.

3. Recall that if 1_H is the trivial character of H, then $\mathrm{Ind}_H^G 1_H$ is the character $\chi_{\widetilde{\sigma}}$ of the permutation representation $\widetilde{\sigma}\colon G \to GL(\mathbb{C}(G/H))$. Use Frobenius reciprocity to show that the rank of σ is the number of orbits of H on G/H.

4. Conclude that G is two-transitive on G/H if and only if H is transitive on the set of cosets not equal to H in G/H.

5. Show that the rank of σ is also the number of double cosets in $H\backslash G/H$ either directly or by using Mackey's Theorem.

Exercise 8.9. Use Frobenius reciprocity to give another proof that if ρ is an irreducible representation of G, then the multiplicity of ρ as a constituent of the regular representation is the degree of ρ.

Exercise 8.10. Let G be a group and H a subgroup. Suppose $\rho\colon H \to GL(V)$ is a representation. Let W be the vector space of all functions $f\colon G \to V$ such that $f(hg) = \rho(h)f(g)$ for all $g \in G$ and $h \in H$ equipped with pointwise operations. Define a representation $\varphi\colon G \to GL(W)$ by $\varphi_g(f)(g_0) = f(g_0 g)$. Prove that φ is a representation of G equivalent to $\operatorname{Ind}_H^G \rho$.

Chapter 9

Another Theorem of Burnside

In this chapter we give another application of representation theory to finite groups, again due to Burnside. The result is based on a study of real characters and conjugacy classes.

9.1 Conjugate representations

If $A = (a_{ij})$ is a matrix, then $\overline{A}$ is the matrix $(\overline{a_{ij}})$. One easily verifies that $\overline{AB} = \overline{A} \cdot \overline{B}$ and that if A is invertible, then so is $\overline{A}$ and moreover $\overline{A}^{-1} = \overline{A^{-1}}$. Hence if $\varphi\colon G \to GL_d(\mathbb{C})$ is a representation of G, then we can define the *conjugate representation* $\overline{\varphi}$ by $\overline{\varphi}_g = \overline{\varphi_g}$. If $f\colon G \to \mathbb{C}$ is a function, then define $\overline{f}$ by $\overline{f}(g) = \overline{f(g)}$.

Proposition 9.1.1. *Let $\varphi\colon G \to GL_d(\mathbb{C})$ be a representation. Then we have $\chi_{\overline{\varphi}} = \overline{\chi_\varphi}$.*

Proof. First note that if $A \in M_d(\mathbb{C})$, then

$$\mathrm{Tr}(\overline{A}) = \overline{a_{11}} + \cdots + \overline{a_{dd}} = \overline{a_{11} + \cdots + a_{dd}} = \overline{\mathrm{Tr}(A)}.$$

Thus $\chi_{\overline{\varphi}}(g) = \mathrm{Tr}(\overline{\varphi_g}) = \overline{\mathrm{Tr}(\varphi_g)} = \overline{\chi_\varphi(g)}$, as required. □

As a consequence, we observe that the conjugate of an irreducible representation is again irreducible.

Corollary 9.1.2. *Let $\varphi\colon G \to GL_d(\mathbb{C})$ be irreducible. Then $\overline{\varphi}$ is irreducible.*

Proof. Let $\chi = \chi_\varphi$. We compute

$$\langle \overline{\chi}, \overline{\chi} \rangle = \frac{1}{|G|} \sum_{g \in G} \chi(g)\overline{\chi(g)} = \frac{1}{|G|} \sum_{g \in G} \overline{\chi(g)}\chi(g) = \langle \chi, \chi \rangle = 1$$

and so $\overline{\varphi}$ is irreducible. □

Quite often one can use the above corollary to produce new irreducible characters for a group. However, the case when $\overline{\chi} = \chi$ is also of importance.

Definition 9.1.3 (Real character)**.** A character χ of G is called *real*[1] if $\chi = \overline{\chi}$, that is, $\chi(g) \in \mathbb{R}$ for all $g \in G$.

Example 9.1.4. The trivial character of a group is always real. The groups S_3 and S_4 have only real characters. On the other hand, if n is odd then $\mathbb{Z}_n$ has no non-trivial real characters.

Since the number of irreducible characters equals the number of conjugacy classes, there should be a corresponding notion of a "real" conjugacy class. First we make two simple observations.

Proposition 9.1.5. *Let χ be a character of a group G. Then $\chi(g^{-1}) = \overline{\chi(g)}$.*

Proof. Without loss of generality, we may assume that χ is the character of a unitary representation $\varphi\colon G \to U_n(\mathbb{C})$. Then

$$\chi(g^{-1}) = \mathrm{Tr}(\varphi_{g^{-1}}) = \mathrm{Tr}(\overline{\varphi_g}^T) = \mathrm{Tr}(\overline{\varphi_g}) = \overline{\mathrm{Tr}(\varphi_g)} = \overline{\chi(g)}$$

as required. □

Proposition 9.1.6. *Let g and h be conjugate. Then g^{-1} and h^{-1} are conjugate.*

Proof. Suppose $g = xhx^{-1}$. Then $g^{-1} = xh^{-1}x^{-1}$. □

So if C is a conjugacy class of G, then $C^{-1} = \{g^{-1} \mid g \in C\}$ is also a conjugacy class of G and moreover if χ is any character then $\chi(C^{-1}) = \overline{\chi(C)}$.

Definition 9.1.7 (Real conjugacy class)**.** A conjugacy class C of G is said to be *real* if $C = C^{-1}$.

The following proposition motivates the name.

[1] Some authors divide what we call real characters into two subclasses: real characters and quaternionic characters.

Proposition 9.1.8. *Let C be a real conjugacy class and χ a character of G. Then $\chi(C) = \overline{\chi(C)}$, that is, $\chi(C) \in \mathbb{R}$.*

Proof. If C is real then $\chi(C) = \chi(C^{-1}) = \overline{\chi(C)}$. □

An important result of Burnside is that the number of real irreducible characters is equal to the number of real conjugacy classes. The elegant proof we provide is due to Brauer and is based on the invertibility of the character table. First we prove a lemma.

Lemma 9.1.9. *Let $\varphi\colon S_n \to GL_n(\mathbb{C})$ be the standard representation of S_n and let $A \in M_n(\mathbb{C})$ be a matrix. Then, for $g \in S_n$, the matrix $\varphi_g A$ is obtained from A by permuting the rows of A according to g and $A\varphi_g$ is obtained from A by permuting the columns of A according to g^{-1}.*

Proof. We compute $(\varphi_g A)_{g(i)j} = \sum_{k=1}^{n} \varphi(g)_{g(i)k} A_{kj} = A_{ij}$ since

$$\varphi(g)_{g(i)k} = \begin{cases} 1 & k = i \\ 0 & \text{else.} \end{cases}$$

Thus $\varphi_g A$ is obtained from A by placing row i of A into row $g(i)$. Since the representation φ is unitary, $A\varphi_g = (\varphi_g^T A^T)^T = (\varphi_{g^{-1}} A^T)^T$ the second statement follows from the first. □

Theorem 9.1.10 (Burnside)**.** *Let G be a finite group. The number of real irreducible characters of G equals the number of real conjugacy classes of G.*

Proof (Brauer). Let s be the number of conjugacy classes of G. Our standing notation will be that $\chi_1, \ldots, \chi_s$ are the irreducible characters of G and $C_1, \ldots, C_s$ are the conjugacy classes. Define $\alpha, \beta \in S_s$ by $\overline{\chi_i} = \chi_{\alpha(i)}$ and $C_i^{-1} = C_{\beta(i)}$. Notice that χ_i is a real character if and only if $\alpha(i) = i$ and similarly C_i is a real conjugacy class if and only if $\beta(i) = i$. Therefore, $|\mathrm{Fix}(\alpha)|$ is the number of real irreducible characters and $|\mathrm{Fix}(\beta)|$ is the number of real conjugacy classes. Notice that $\alpha = \alpha^{-1}$ since α swaps the indices of χ_i and $\overline{\chi_i}$.

Let $\varphi\colon S_s \to GL_s(\mathbb{C})$ be the standard representation of S_s. Then we have $\chi_\varphi(\alpha) = |\mathrm{Fix}(\alpha)|$ and $\chi_\varphi(\beta) = |\mathrm{Fix}(\beta)|$ so it suffices to prove $\mathrm{Tr}(\varphi_\alpha) = \mathrm{Tr}(\varphi_\beta)$. Let X be the character table of G. Then by Lemma 9.1.9 $\varphi_\alpha \mathsf{X}$ is obtained from X by swapping the rows of X corresponding to χ_i and $\overline{\chi_i}$ for each i. But this means that $\varphi_\alpha \mathsf{X} = \overline{\mathsf{X}}$. Similarly, $\mathsf{X}\varphi_\beta$ is obtained from X by swapping the columns of X corresponding to C_i and C_i^{-1} for each i. Since

$\chi(C^{-1}) = \overline{\chi(C)}$ for each conjugacy class C, this swapping again results in $\overline{\mathsf{X}}$. In other words,

$$\varphi_\alpha \mathsf{X} = \overline{\mathsf{X}} = \mathsf{X}\varphi_\beta.$$

But by the second orthogonality relations (Theorem 4.4.12) the columns of X form an orthogonal set of non-zero vectors and hence are linearly independent. Thus X is invertible and so $\varphi_\alpha = \mathsf{X}\varphi_\beta\mathsf{X}^{-1}$. We conclude $\mathrm{Tr}(\varphi_\alpha) = \mathrm{Tr}(\varphi_\beta)$, as was required. □

As a consequence we see that groups of odd order do not have non-trivial real irreducible characters.

Proposition 9.1.11. *Let G be a group. Then $|G|$ is odd if and only if G does not have any non-trivial real irreducible characters.*

Proof. By Theorem 9.1.10, it suffices to show that $\{1\}$ is the only real conjugacy class of G if and only if $|G|$ is odd. Suppose first G has even order. Then there is an element $g \in G$ of order 2. Since $g = g^{-1}$, if C is the conjugacy class of g, then $C = C^{-1}$ is real.

Suppose conversely that G contains a non-trivial real conjugacy class C. Let $g \in C$ and $N_G(g) = \{x \in G \mid xg = gx\}$ be the normalizer of g. Then $|C| = [G : N_G(g)]$. Suppose that $hgh^{-1} = g^{-1}$. Then

$$h^2gh^{-2} = hg^{-1}h^{-1} = (hgh^{-1})^{-1} = g$$

and so $h^2 \in N_G(g)$. If $h \in \langle h^2 \rangle$, then $h \in N_G(g)$ and so $g^{-1} = hgh^{-1} = g$. Hence in this case $g^2 = 1$ and so $|G|$ is even. If $h \notin \langle h^2 \rangle$, then h^2 is not a generator of $\langle h \rangle$ and so 2 divides the order of h. Thus $|G|$ is even. This completes the proof. □

From Proposition 9.1.11, we deduce a curious result about groups of odd order that doesn't seem to admit a direct elementary proof.

Theorem 9.1.12 (Burnside)**.** *Let G be a group of odd order and let s be the number of conjugacy classes of G. Then $s \equiv |G| \bmod 16$.*

Proof. By Proposition 9.1.11, G has the trivial character χ_0 and the remaining characters come in conjugate pairs $\chi_1, \chi_1', \ldots, \chi_k, \chi_k'$ of degrees $d_1, \ldots, d_k$. In particular, $s = 1 + 2k$ and

$$|G| = 1 + \sum_{j=1}^{k} 2d_j^2.$$

Since d_j divides $|G|$ it is odd and so we may write it as $d_j = 2m_j + 1$ for some non-negative integer m_j. Therefore, we have

$$\begin{aligned}|G| &= 1 + \sum_{j=1}^{k} 2(2m_j+1)^2 = 1 + \sum_{j=1}^{k}(8m_j^2 + 8m_j + 2)\\ &= 1 + 2k + 8\sum_{j=1}^{k} m_j(m_j+1) = s + 8\sum_{j=1}^{k} m_j(m_j+1)\\ &\equiv s \bmod 16\end{aligned}$$

since exactly one of m_j and $m_j + 1$ is even. □

Exercises

Exercise 9.1. Let G be a finite group.

1. Prove that two element $g, h \in G$ are conjugate if and only if $\chi(g) = \chi(h)$ for all irreducible characters χ.

2. Show that the conjugacy class C of an element $g \in G$ is real if and only if $\chi(g) = \overline{\chi(g)}$ for all irreducible characters χ.

Chapter 10

Representation Theory of the Symmetric Group

In this chapter, we construct the irreducible representations of the symmetric group S_n.

10.1 Partitions and tableaux

We begin with the fundamental notion of a partition of n. Simply speaking, a partition of n is a way of writing n as a sum of positive integers.

Definition 10.1.1 (Partition)**.** A *partition of* n is a tuple $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ of positive integers such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell$ and $\lambda_1 + \cdots + \lambda_\ell = n$. To indicate that λ is a partition of n, we write $\lambda \vdash n$.

For example, $(2, 2, 1, 1)$ is partition of 6 and $(3, 1)$ is partition of 4. Note that $(1, 2, 1)$ is not a partition of 4 since the second entry is bigger than the first.

There is a natural partition of n associated to any permutation $\sigma \in S_n$ called the *cycle type* of σ. Namely, $\mathrm{type}(\sigma) = (\lambda_1, \ldots, \lambda_\ell)$ where the λ_i are the lengths of the cycles of σ in decreasing order (with multiplicity). Here we must count cycles of length 1, which are normally omitted from the notation when writing cycle decompositions.

Example 10.1.2. Let $n = 5$. Then

$$\begin{aligned} \text{type}((1\ 2)(5\ 3\ 4)) &= (3,2) \\ \text{type}((1\ 2\ 3)) &= (3,1,1) \\ \text{type}((1\ 2\ 3\ 4\ 5)) &= (5) \\ \text{type}((1\ 2)(3\ 4)) &= (2,2,1). \end{aligned}$$

It is typically shown in a first course in group theory that two permutations are conjugate if and only if they have the same cycle type.

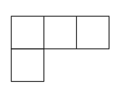
Theorem 10.1.3. *Let $\sigma, \tau \in S_n$. Then σ is conjugate to τ if and only if* $\text{type}(\sigma) = \text{type}(\tau)$.

It follows that the number of irreducible representations of S_n is the number of partitions of n. Thus we expect partitions to play a major role in the representation theory of the symmetric group. Our goal in this chapter is to give an explicit bijection between partitions of n and irreducible representations of S_n. First we need to deal with some preliminary combinatorics.

It is often convenient to represent partitions by a Tetris-like picture called a Young diagram.

Definition 10.1.4 (Young diagram)**.** If $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ is a partition of n, then the *Young diagram* (or simply *diagram*) of λ consists of n boxes placed into ℓ rows where the i^{th} row has λ_i boxes.

This definition is best illustrated with an example. If $\lambda = (3, 1)$, then the Young diagram is as follows.



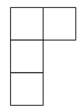
(10.1)

Conversely, any diagram consisting of n boxes arranged into rows such that the number of boxes in each row is non-increasing is the Young diagram of some partition of n.

Definition 10.1.5 (Conjugate partition)**.** If $\lambda \vdash n$, then the *conjugate partition* λ^T of λ is the partition whose Young diagram is the transpose of the diagram of λ, that is, the Young diagram of λ^T is obtained from the diagram of λ by exchanging rows and columns.

Again, a picture is worth one thousand words.

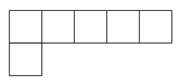
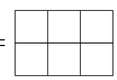
Example 10.1.6. If $\lambda = (3, 1)$, then its diagram is as in (10.1). The transpose diagram is



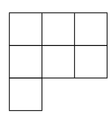
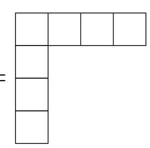
and so $\lambda^T = (2, 1, 1)$.

Next we want to introduce an ordering on partitions. Given two partitions λ and μ of n, we want to say that λ dominates μ, written $\lambda \trianglerighteq \mu$, if, for every $i \geq 1$, the first i rows of the diagram of λ contain at least as many boxes as the first i rows of μ.

Example 10.1.7. For instance, $(5, 1) \trianglerighteq (3, 3)$ as we can see from

$(5, 1) =$  and $(3, 3) =$ 

But neither $(3, 3, 1) \trianglerighteq (4, 1, 1, 1)$, nor $(4, 1, 1, 1) \trianglerighteq (3, 3, 1)$ because $(4, 1, 1, 1)$ has more elements in the first row, but $(3, 3, 1)$ has more elements in the first two rows.

$(3, 3, 1) =$  $(4, 1, 1, 1) =$ 

Let us formalize the definition. Observe that the number of boxes in the first i rows of $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ is $\lambda_1 + \cdots + \lambda_i$.

Definition 10.1.8 (Domination order)**.** Suppose that $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ and $\mu = (\mu_1, \ldots, \mu_m)$ are partitions of n. Then λ is said to *dominate* μ if

$$\lambda_1 + \lambda_2 + \cdots + \lambda_i \geq \mu_1 + \mu_2 + \cdots + \mu_i$$

for all $i \geq 1$ where if $i > \ell$, then we take $\lambda_i = 0$, and if $i > m$, then we take $\mu_i = 0$.

The domination order satisfies many of the properties enjoyed by $\geq$.

Proposition 10.1.9. *The dominance order satisfies:*

1. *Reflexivity:* $\lambda \trianglerighteq \lambda$;
2. *Anti-symmetry:* $\lambda \trianglerighteq \mu$ *and* $\mu \trianglerighteq \lambda$ *implies* $\lambda = \mu$;

3. *Transitivity:* $\lambda \trianglerighteq \mu$ *and* $\mu \trianglerighteq \rho$ *implies* $\lambda \trianglerighteq \rho$.

Proof. Reflexivity is clear. Suppose $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ and $\mu = (\mu_1, \ldots, \mu_m)$ are partitions of n. We prove by induction on n that $\lambda \trianglerighteq \mu$ and $\mu \trianglerighteq \lambda$ implies $\lambda = \mu$. If $n = 1$, then $\lambda = (1) = \mu$ and there is nothing to prove. Otherwise, by taking $i = 1$, we see that $\lambda_1 = \mu_1$. Call this common value $k > 0$. Then define partitions λ', μ' of $n - k$ by $\lambda' = (\lambda_2, \ldots, \lambda_\ell)$ and $\mu' = (\mu_2, \ldots, \mu_m)$. Since

$$\lambda_1 + \lambda_2 + \cdots + \lambda_i = \mu_1 + \mu_2 + \cdots + \mu_i$$

for all $i \geq 1$ and $\lambda_1 = \mu_1$, it follows that

$$\lambda_2 + \cdots + \lambda_i = \mu_2 + \cdots + \mu_i$$

for all $i \geq 1$ and hence $\lambda' \trianglerighteq \mu'$ and $\mu' \trianglerighteq \lambda'$. Thus by induction $\lambda' = \mu'$ and hence $\lambda = \mu$. This establishes anti-symmetry.

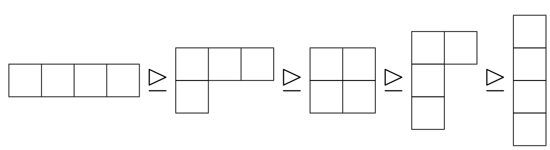
To obtain transitivity, simply observe that

$$\lambda_1 + \cdots + \lambda_i \geq \mu_1 + \cdots + \mu_i \geq \rho_1 + \cdots + \rho_i$$

and so $\lambda \trianglerighteq \rho$. □

Proposition 10.1.9 says that $\trianglerighteq$ is a *partial order* on the set of partitions of n.

Example 10.1.10.



Young tableaux are obtained from Young diagrams by placing the integers $1, \ldots, n$ into the boxes.

Definition 10.1.11 (Young tableaux)**.** If $\lambda \vdash n$, then a λ-*tableaux* (or *Young tableaux of shape* λ) is an array t of integers obtained by placing $1, \ldots, n$ into the boxes of the Young diagram for λ. There are clearly $n!$ λ-tableaux.

This concept is again best illustrated with an example.

Example 10.1.12. Suppose that $\lambda = (3, 2, 1)$. Then some λ-tableaux are as follows.

$$\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2} 6 \\ \cline{1-1}\end{array}\ ,\quad \begin{array}{|c|c|c|}\hline 3 & 1 & 2 \\ \hline 6 & 5 \\ \cline{1-2} 4 \\ \cline{1-1}\end{array}\ ,\quad \begin{array}{|c|c|c|}\hline 3 & 2 & 6 \\ \hline 1 & 5 \\ \cline{1-2} 4 \\ \cline{1-1}\end{array}$$

A rather technical combinatorial fact is that if t^λ is a λ-tableaux and s^μ is a μ-tableaux such that the integers in any given row of s^μ belong to distinct columns of t^λ, then $\lambda \trianglerighteq \mu$.

To prove this, we need the following proposition, which will be useful in its own right.

Proposition 10.1.13. *Let $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ and $\mu = (\mu_1, \ldots, \mu_m)$ be partitions of n. Suppose that t^λ is a λ-tableaux and s^μ is a μ-tableaux such that entries in the same row of s^μ are located in different columns of t^λ. Then we can find a λ-tableaux u^λ such that:*

1. *The j^{th} columns of t^λ and u^λ contain the same elements for $1 \leq j \leq \ell$;*
2. *The entries of the first i rows of s^μ belong to the first i rows of u^λ for each $1 \leq i \leq m$.*

Proof. For each $1 \leq r \leq m$, we construct a λ-tableaux t^λ_r such that:

(a) The j^{th} columns of t^λ and t^λ_r contain the same elements for $1 \leq j \leq \ell$;

(b) The entries of the first i rows of s^μ belong to the first i rows of t^λ_r for $1 \leq r \leq m$.

Setting $u^\lambda = t^\lambda_m$ will then complete the proof. The construction is by induction on r. Let us begin with $r = 1$. Let k be an element in the first row of s^μ and let $c(k)$ be the column of t^λ containing k. If k is in the first row of t^λ, we do nothing. Otherwise, we switch in t^λ the first entry in $c(k)$ with k. Because each element k of the first row of s^μ is in a different column of t^λ, the order in which we do this doesn't matter, and so there results a new λ-tableaux t^λ_1 satisfying properties 1 and 2.

Next suppose that t^λ_r with the desired two properties has been constructed for $1 \leq r \leq m-1$. Define t^λ_{r+1} as follows. Let k be an entry of row $r+1$ of s^μ and let $c(k)$ be the column in which k appears in t^λ_r. If k already appears in the first $r+1$ rows of t^λ_r, there is nothing to do. So assume that k does not appear in the first $r+1$ rows of t^λ_r. Notice that if row $r+1$ of t^λ_r does not intersect $c(k)$, then since the sizes of the rows are non-increasing, it follows that k already appears in the first r rows of t^λ_r. Thus we must have that $c(k)$ intersects row $r+1$ and so we can switch k with the element in row $r+1$ and column $c(k)$ of t^λ_r. Again, because each entry of row $r+1$ of s^μ is in a different column of t^λ, and hence of t^λ_r by property (a), we can do this for each such k independently. In this way, we have constructed t^λ_{r+1} satisfying (a) and (b). □

Let us illustrate how this works with an example.

Example 10.1.14. Suppose that t^λ and s^μ are given by

$$t^\lambda = \begin{array}{|c|c|c|c|c|}\hline 8 & 5 & 4 & 2 & 7 \\ \hline 1 & 3 \\ \cline{1-2} 6 \\ \cline{1-1}\end{array} \quad \text{and} \quad s^\mu = \begin{array}{|c|c|c|c|}\hline 1 & 2 & 3 & 4 \\ \hline 5 & 6 \\ \cline{1-2} 7 & 8 \\ \cline{1-2}\end{array}.$$

No two elements in the same row of s^μ belong to the same column of t^λ.

We construct t_1^λ by switching in t^λ each element appearing of the first row of s^μ with the element in its column of the first row of t^λ. So

$$t_1^\lambda = \begin{array}{|c|c|c|c|c|}\hline 1 & 3 & 4 & 2 & 7 \\ \hline 8 & 5 \\ \cline{1-2} 6 \\ \cline{1-1}\end{array}.$$

Now by switching 8 and 6, we obtain the λ-tableaux

$$t_2^\lambda = \begin{array}{|c|c|c|c|c|}\hline 1 & 3 & 4 & 2 & 7 \\ \hline 6 & 5 \\ \cline{1-2} 8 \\ \cline{1-1}\end{array}$$

which has every element in the first i rows of s^μ located in the first i rows of t_2^λ for $i = 1, 2, 3$. Hence we can take $u^\lambda = t_2^\lambda$.

Our first use of Proposition 10.1.13 is to establish the following combinatorial criterion for domination.

Lemma 10.1.15 (Dominance lemma)**.** *Let λ and μ be partitions of n and suppose that t^λ and s^μ are tableaux of respective shapes λ and μ. Moreover, suppose that integers in the same row of s^μ are located in different columns of t^λ. Then $\lambda \trianglerighteq \mu$.*

Proof. Let $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ and $\mu = (\mu_1, \ldots, \mu_m)$. By Proposition 10.1.13 we can find a λ-tableaux u^λ such that, for $1 \leq i \leq m$, the entries of the first i rows of s^μ are in the first i rows of u^λ. Then since $\lambda_1 + \cdots + \lambda_i$ is the number of entries in the first i rows of u^λ and $\mu_1 + \cdots + \mu_i$ is the number of entries in the first i rows of s^μ, it follows that $\lambda_1 + \cdots + \lambda_i \geq \mu_1 + \cdots + \mu_i$ for all $i \geq 1$ and hence $\lambda \trianglerighteq \mu$. $\square$

10.2 Constructing the irreducible representations

If $X \subseteq \{1, \ldots, n\}$, we identify S_X with those permutations in S_n that fix all elements outside of X. For instance, $S_{\{2,3\}}$ consists of $\{Id, (2\ 3)\}$.

Definition 10.2.1 (Column stabilizer)**.** Let t be a Young tableaux. Then the *column stabilizer* of t is the subgroup of S_n preserving the columns of t. That is, $\sigma \in C_t$ if and only if $\sigma(i)$ is in the same column as i for each $i \in \{1, \ldots, n\}$.

Let us turn to an example.

Example 10.2.2. Suppose that

$$t = \begin{array}{|c|c|c|}\hline 1 & 3 & 7 \\ \hline 4 & 5 \\ \cline{1-2} 2 & 6 \\ \cline{1-2}\end{array}\,.$$

Then $C_t = S_{\{1,2,4\}} S_{\{3,5,6\}} S_{\{7\}} \cong S_{\{1,2,4\}} \times S_{\{3,5,6\}} \times S_{\{7\}}$. So, for example, $(1\ 4), (1\ 2\ 4)(3\ 5) \in C_t$. Since $S_{\{7\}} = \{Id\}$, it follows $|C_t| = 3! \cdot 3! = 36$.

The group S_n acts transitively on the set of λ-tableaux by applying $\sigma \in S_n$ to the entries of the boxes. The result of applying $\sigma \in S_n$ to t is denoted σt. For example, if

$$t = \begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\ \hline 2 \\ \cline{1-1}\end{array}$$

and $\sigma = (1\ 3\ 2)$, then

$$\sigma t = \begin{array}{|c|c|c|}\hline 3 & 2 & 4 \\ \hline 1 \\ \cline{1-1}\end{array}\,.$$

Let us define an equivalence relation $\sim$ on the set of λ-tableaux by putting $t_1 \sim t_2$ if they have the same entries in each row. For example,

$$\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2}\end{array} \sim \begin{array}{|c|c|c|}\hline 3 & 1 & 2 \\ \hline 5 & 4 \\ \cline{1-2}\end{array}$$

since they both have $\{1, 2, 3\}$ in the first row and $\{4, 5\}$ in the second row.

Definition 10.2.3 (Tabloid)**.** A $\sim$-equivalence class of λ-tableaux is called a *λ-tabloid* or a *tabloid of shape* λ. The tabloid of a tableaux t is denoted $[t]$. The set of all tabloids of shape λ is denote T^λ. Denote by T_λ the tabloid with $1, \ldots, \lambda_1$ in row 1, $\lambda_1 + 1, \ldots, \lambda_1 + \lambda_2$ in row 2 and in general with $\lambda_1 + \cdots + \lambda_{i-1} + 1, \ldots, \lambda_1 + \cdots + \lambda_i$ in row i. In other words T_λ is the tabloid corresponding to the tableaux which has j in the j^{th} box.

For example, $T_{(3,2)}$ is the equivalence class of

$$\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2}\end{array}\,.$$

Our next proposition shows that the action of S_n on λ-tableaux induces a well-defined action of S_n on tabloids of shape λ.

Proposition 10.2.4. *Suppose that $t_1 \sim t_2$ and $\sigma \in S_n$. Then $\sigma t_1 \sim \sigma t_2$. Hence there is a well-defined action of S_n on T^λ given by putting $\sigma[t] = [\sigma t]$ for t a λ-tableaux.*

Proof. To show that $\sigma t_1 \sim \sigma t_2$, we must show that i, j are in the same row of σt_1 if and only if they are in the same row of σt_2. But i, j are in the same row of σt_1 if and only if $\sigma^{-1}(i)$ and $\sigma^{-1}(j)$ are in the same row of t_1, which occurs if and only if $\sigma^{-1}(i)$ and $\sigma^{-1}(j)$ are in the same row of t_2. But this occurs if and only if i, j are in the same row of σt_2. This proves that $\sigma t_1 \sim \sigma t_2$. From this it is easy to verify that $\sigma[t] = [\sigma t]$ gives a well-defined action of S_n on T^λ. □

The action of S_n on λ-tabloids is transitive since it was already transitive on λ-tableaux. Suppose that $\lambda = (\lambda_1, \ldots, \lambda_\ell)$. The stabilizer S_λ of T_λ is

$$S_\lambda = S_{\{1,\ldots,\lambda_1\}} \times S_{\{\lambda_1+1,\ldots,\lambda_1+\lambda_2\}} \times \cdots \times S_{\{\lambda_1+\cdots+\lambda_{\ell-1}+1,\ldots,n\}}.$$

Thus $|T^\lambda| = [S_n : S_\lambda] = n!/\lambda_1! \cdots \lambda_\ell!$.

For a partition λ, set $M^\lambda = \mathbb{C}T^\lambda$ and let $\varphi^\lambda \colon S_n \to GL(M^\lambda)$ be the associated permutation representation.

Example 10.2.5. Suppose that $\lambda = (n-1, 1)$. Then two λ-tableaux are equivalent if and only if they have the same entry in the second row. Thus T^λ is in bijection with $\{1, \ldots, n\}$ and φ^λ is equivalent to the standard representation. On the other hand, if $\lambda = (n)$, then there is only one λ-tabloid and so φ^λ is the trivial representation.

If $\lambda \neq (n)$, then φ^λ is a non-trivial permutation representation of S_n and hence is not irreducible. Nonetheless, it contains a distinguished irreducible constituent that we now seek to isolate.

Definition 10.2.6 (Polytabloid)**.** Let $\lambda, \mu \vdash n$. Let t be a λ-tableaux and define a linear operator $A_t \colon M^\mu \to M^\mu$ by

$$A_t = \sum_{\pi \in C_t} \operatorname{sgn}(\pi) \varphi^\mu_\pi.$$

In the case $\lambda = \mu$, the element

$$e_t = A_t[t] = \sum_{\pi \in C_t} \operatorname{sgn}(\pi) \pi[t]$$

of M^λ is called the *polytabloid* associated to t.

Our next proposition shows that the action of S_n on λ-tableaux is compatible with the definition of a λ-tabloid.

Proposition 10.2.7. *If $\sigma \in S_n$ and t is a λ-tableaux, then $\varphi^\lambda_\sigma e_t = e_{\sigma t}$.*

Proof. First we claim that $C_{\sigma t} = \sigma C_t \sigma^{-1}$. Indeed, if X_i is the set of entries of column i of t, then $\sigma(X_i)$ is the set of entries of column i of σt. Since τ stabilizes X_i if and only if $\sigma\tau\sigma^{-1}$ stabilizes $\sigma(X_i)$, the claim follows. Now we compute

$$\begin{aligned}\varphi^\lambda_\sigma A_t &= \sum_{\pi \in C_t} \operatorname{sgn}(\pi)\varphi^\lambda_\sigma \varphi^\lambda_\pi \\ &= \sum_{\tau \in C_{\sigma t}} \operatorname{sgn}(\sigma^{-1}\tau\sigma)\varphi^\lambda_\sigma \varphi^\lambda_{\sigma^{-1}\tau\sigma} \\ &= A_{\sigma t}\varphi^\lambda_\sigma\end{aligned}$$

where we have made the substitution $\tau = \sigma\pi\sigma^{-1}$.

Thus $\varphi^\lambda_\sigma e_t = \varphi^\lambda_\sigma A_t[t] = A_{\sigma t}\varphi^\lambda_\sigma[t] = A_{\sigma t}[\sigma t] = e_{\sigma t}$. This completes the proof. □

We can now define our desired subrepresentation.

Definition 10.2.8 (Sprecht representation)**.** Let λ be a partition of n. Define S^λ to be the subspace of M^λ spanned by the polytabloids e_t with t a λ-tableaux. Proposition 10.2.7 implies that S^λ is S_n-invariant. Let $\psi^\lambda \colon S_n \to GL(S^\lambda)$ be the corresponding subrepresentation. It is called the *Sprecht representation* associated to λ.

Remark 10.2.9. The e_t are not in general linearly independent. See the next example.

Our goal is to prove that the ψ^λ form a complete set of irreducible representations of S_n. Let's look at an example.

Example 10.2.10 (Alternating representation)**.** Consider the partition $\lambda = (1, 1, \ldots, 1)$ of n. Since each row has only one element, λ-tableaux are the same thing as λ-tabloids. Thus φ^λ is equivalent to the regular representation of S_n. Let t be a λ-tableaux. Because t has only one column, trivially $C_t = S_n$. Thus

$$e_t = \sum_{\pi \in S_n} \operatorname{sgn}(\pi)\pi[t].$$

We claim that if $\sigma \in S_n$, then $\varphi^\lambda_\sigma e_t = \operatorname{sgn}(\sigma)e_t$. Since we know that $\varphi^\lambda_\sigma e_t = e_{\sigma t}$ by Proposition 10.2.7, it will follow that $S^\lambda = \mathbb{C}e_t$ and that ψ^λ is equivalent to the degree one representation $\operatorname{sgn} \colon S_n \to \mathbb{C}^*$.

Indeed,

$$\begin{aligned}\varphi^\lambda_\sigma e_t &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi)\varphi^\lambda_\sigma\varphi^\lambda_\pi[t]\\ &= \sum_{\tau\in S_n} \operatorname{sgn}(\sigma^{-1}\tau)\varphi^\lambda_\tau[t]\\ &= \operatorname{sgn}(\sigma)e_t\end{aligned}$$

where we have performed the substitution $\tau = \sigma\pi$.

The proof that the ψ^λ are the irreducible representations of S_n proceeds via a series of lemmas.

Lemma 10.2.11. *Let $\lambda, \mu \vdash n$ and suppose that t^λ is a λ-tableaux and s^μ is a μ-tableaux such that $A_{t^\lambda}[s^\mu] \neq 0$. Then $\lambda \trianglerighteq \mu$. Moreover, if $\lambda = \mu$, then $A_{t^\lambda}[s^\mu] = \pm e_{t^\lambda}$.*

Proof. We use the dominance lemma. Suppose that we have two elements i, j that are in the same row of s^μ and the same column of t^λ. Then $(i\ j)[s^\mu] = [s^\mu] = Id[s^\mu]$ and thus

$$(\varphi^\mu_{Id} - \varphi^\mu_{(i\ j)})[s^\mu] = 0. \tag{10.2}$$

Let $H = \{Id, (i\ j)\}$. Then H is a subgroup of C_{t^λ}; let $\sigma_1, \ldots, \sigma_k$ be a complete set of left coset representatives for H in C_{t^λ}. Then we have

$$\begin{aligned}A_{t^\lambda}[s^\mu] &= \sum_{\pi\in C_{t^\lambda}} \operatorname{sgn}(\pi)\varphi^\mu_\pi[s^\mu]\\ &= \sum_{r=1}^{k}\left(\operatorname{sgn}(\sigma_r)\varphi^\mu_{\sigma_r} + \operatorname{sgn}(\sigma_r(i\ j))\varphi^\mu_{\sigma_r(i\ j)}\right)[s^\mu]\\ &= \sum_{r=1}^{k}\operatorname{sgn}(\sigma_r)\varphi^\mu_{\sigma_r}(\varphi^\mu_{Id} - \varphi^\mu_{(i\ j)})[s^\mu]\\ &= 0\end{aligned}$$

where the last equality uses (10.2). This contradiction implies that the elements of each row of s^μ are in different columns of t^λ. The dominance lemma (Lemma 10.1.15) now yields that $\lambda \trianglerighteq \mu$.

Next suppose that $\lambda = \mu$. Let u^λ be as in Proposition 10.1.13. The fact that the columns of t^λ and u^λ have the same elements implies that the unique permutation σ with $u^\lambda = \sigma t^\lambda$ actually belongs to C_{t^λ}. On the other hand, for all $i \geq 1$, the first i rows of s^μ belong to the first i rows of u^λ. But since $\lambda = \mu$, this implies $[u^\lambda] = [s^\mu]$. Indeed, the first row of s^μ is contained in the first row of u^λ, but they have the same number of boxes. So these rows contain the same elements. Suppose by induction, that each of the first i rows of u^λ and s^μ have the same elements. Then since each element of the first $i+1$ rows of s^μ belongs to the first $i+1$ rows of u^λ, it follows from the inductive hypothesis that each element of row $i+1$ of s^μ belongs to row $i+1$ of u^λ. Since these tableaux both have shape λ, it follows that they have the same $(i+1)^{st}$ row. We conclude that $[u^\lambda] = [s^\mu]$.

It follows that

$$\begin{aligned} A_{t^\lambda}[s^\mu] &= \sum_{\pi \in C_{t^\lambda}} \operatorname{sgn}(\pi)\varphi^\lambda_\pi[s^\mu] \\ &= \sum_{\tau \in C_{t^\lambda}} \operatorname{sgn}(\tau\sigma^{-1})\varphi^\lambda_\tau\varphi^\lambda_{\sigma^{-1}}[u^\lambda] \\ &= \operatorname{sgn}(\sigma^{-1}) \sum_{\tau \in C_{t^\lambda}} \operatorname{sgn}(\tau)\tau[t^\lambda] \\ &= \pm e_{t^\lambda} \end{aligned}$$

where in the second equality we have performed the change of variables $\tau = \pi\sigma$. This completes the proof. □

The next lemma continues our study of the operator A_t.

Lemma 10.2.12. *Let t be a λ-tableaux. Then the image of the operator $A_t \colon M^\lambda \to M^\lambda$ is $\mathbb{C}e_t$.*

Proof. From the equation $e_t = A_t[t]$, it suffices to show that the image is contained in $\mathbb{C}e_t$. To prove this, it suffices to check on basis elements $[s] \in T^\lambda$. If $A_t[s] = 0$, there is nothing to prove; otherwise, Lemma 10.2.11 yields $A_t[s] = \pm e_t \in \mathbb{C}e_t$. This completes the proof. □

Recall that $M^\lambda = \mathbb{C}T^\lambda$ comes equipped with an inner product for which T^λ is an orthonormal basis and that, moreover, the representation φ^λ is unitary with respect to this product. Furthermore, if t is a λ-tableaux, then

$$A_t^* = \sum_{\pi \in C_t} \operatorname{sgn}(\pi)(\varphi^\lambda_\pi)^* = \sum_{\tau \in C_t} \operatorname{sgn}(\tau)\varphi^\lambda_\tau = A_t$$

where the penultimate equality is obtained by setting $\tau = \pi^{-1}$ and using that φ is unitary. Thus A_t is self-adjoint.

The key to proving that the ψ^λ are the irreducible representations of S_n is the following theorem.

Theorem 10.2.13 (Subrepresentation theorem). *Let λ be a partition of n and suppose that V is an S_n-invariant subspace of M^λ. Then either $S^\lambda \subseteq V$ or $V \subseteq (S^\lambda)^\perp$.*

Proof. Suppose first that there is a λ-tableaux t and a vector $v \in V$ such that $A_t v \neq 0$. Then by Lemma 10.2.12 and S_n-invariance of V, we have $0 \neq A_t v \in \mathbb{C}e_t \cap V$. It follows that $e_t \in V$. Hence, for all $\sigma \in S_n$, we have $e_{\sigma t} = \varphi^\lambda_\sigma e_t \in V$. Because S_n acts transitively on the set of λ-tableaux, we conclude that $S^\lambda \subseteq V$.

Suppose next that, for all λ-tableaux t and all $v \in V$, one has $A_t v = 0$. Then we have

$$\langle v, e_t\rangle = \langle v, A_t[t]\rangle = \langle A_t^* v, [t]\rangle = \langle A_t v, [t]\rangle = 0$$

because $A_t^* = A_t$ and $A_t v = 0$. As t and v were arbitrary, this shows that $V \subseteq (S^\lambda)^\perp$, completing the proof. □

As a corollary we see that S^λ is irreducible.

Corollary 10.2.14. *Let $\lambda \vdash n$. Then $\psi^\lambda \colon S_n \to GL(S^\lambda)$ is irreducible.*

Proof. Let V be a proper S_n-invariant subspace of S^λ. Then by Theorem 10.2.13, we have $V \subseteq (S^\lambda)^\perp \cap S^\lambda = \{0\}$. This yields the corollary. □

We have thus constructed, for each partition λ of n, an irreducible representation of S_n. The number of conjugacy classes of S_n is the number of partitions of n. Hence if we can show that $\lambda \neq \mu$ implies that $\psi^\lambda \nsim \psi^\mu$, then it will follow that we have found all the irreducible representations of S_n.

Lemma 10.2.15. *Suppose that $\lambda, \mu \vdash n$ and let $T \in \mathrm{Hom}_{S_n}(\varphi^\lambda, \varphi^\mu)$. If $S^\lambda \nsubseteq \ker T$, then $\lambda \trianglerighteq \mu$. Moreover, if $\lambda = \mu$, then $T|_{S^\lambda}$ is a scalar multiple of the identity map.*

Proof. Theorem 10.2.13 implies that $\ker T \subseteq (S^\lambda)^\perp$. So, for any λ-tableaux t, it follows that $0 \neq Te_t = TA_t[t] = A_t T[t]$, where the last equality uses that T commutes with $\varphi^\lambda(S_n)$ and the definition of A_t. Now $T[t]$ is a linear combination of μ-tabloids and so there exists a μ-tabloid $[s]$ such that $A_t[s] \neq 0$. But then $\lambda \trianglerighteq \mu$ by Lemma 10.2.11.

Suppose now that $\lambda = \mu$. Then

$$Te_t = A_t T[t] \in \mathbb{C}e_t \subseteq S^\lambda$$

by Lemma 10.2.12. Thus T leaves S^λ invariant. Since S^λ is irreducible, Schur's lemma implies $T|_{S^\lambda} = cI$ for some $c \in \mathbb{C}$. □

As a consequence we obtain the following result.

Lemma 10.2.16. *If* $\mathrm{Hom}_{S_n}(\psi^\lambda, \varphi^\mu) \neq 0$*, then* $\lambda \trianglerighteq \mu$*. Moreover, if* $\lambda = \mu$*, then* $\dim \mathrm{Hom}_{S_n}(\psi^\lambda, \varphi^\mu) = 1$*.*

Proof. Let $T\colon S^\lambda \to M^\mu$ be a non-zero homomorphism of representations. Then we can extend T to $M^\lambda = S^\lambda \oplus (S^\lambda)^\perp$ by putting $T(v + w) = Tv$ for elements $v \in S^\lambda$ and $w \in (S^\lambda)^\perp$. This extension is a homomorphism of representations because $(S^\lambda)^\perp$ is S_n-invariant and so

$$T(\varphi^\lambda_\sigma(v + w)) = T(\varphi^\lambda_\sigma v + \varphi^\lambda_\sigma w) = T\varphi^\lambda_\sigma v = \varphi^\mu_\sigma Tv = \varphi^\mu_\sigma T(v + w).$$

Clearly $S^\lambda \nsubseteq \ker T$ and so $\lambda \trianglerighteq \mu$ by Lemma 10.2.15. Moreover, if $\lambda = \mu$, then T must be a scalar multiple of the inclusion map by Lemma 10.2.15 and so $\dim \mathrm{Hom}_{S_n}(\psi^\lambda, \varphi^\mu) = 1$. □

We can now prove the main result.

Theorem 10.2.17. *The Sprecht representations* ψ^λ *with* $\lambda \vdash n$ *form a complete set of inequivalent irreducible representations of* S_n*.*

Proof. All that remains is to show that $\psi^\lambda \sim \psi^\mu$ implies $\lambda = \mu$. But $\psi^\lambda \sim \psi^\mu$, implies that $0 \neq \mathrm{Hom}_{S_n}(\psi^\lambda, \psi^\mu) \subseteq \mathrm{Hom}_{S_n}(\psi^\lambda, \varphi^\mu)$. Thus $\lambda \trianglerighteq \mu$ by Lemma 10.2.16. A symmetric argument shows that $\mu \trianglerighteq \lambda$ and so $\lambda = \mu$ by Proposition 10.1.9. This establishes the theorem. □

In fact, we can deduce more from Lemma 10.2.16.

Corollary 10.2.18. *Suppose* $\mu \vdash n$*. Then* ψ^μ *appears with multiplicity one as an irreducible constituent of* φ^μ*. Any other irreducible constituent* ψ^λ *of* φ^μ *satisfies* $\lambda \trianglerighteq \mu$*.*

Exercises

Exercise 10.1. Verify that the relation $\sim$ on λ-tableaux is an equivalence relation.

Exercise 10.2. Verify that the action in Proposition 10.2.4 is indeed an action.

Exercise 10.3. Prove that if $\lambda = (n - 1, 1)$, then the corresponding Sprecht representation of S_n is equivalent to the augmentation subrepresentation of the standard representation of S_n.

Exercise 10.4. Compute the character table of S_5.

Bibliography

[1] Tullio Ceccherini-Silberstein, Fabio Scarabotti, and Filippo Tolli. *Harmonic analysis on finite groups*, volume 108 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2008. Representation theory, Gelfand pairs and Markov chains.

[2] Charles W. Curtis and Irving Reiner. *Representation theory of finite groups and associative algebras*. Wiley Classics Library. John Wiley & Sons Inc., New York, 1988. Reprint of the 1962 original, A Wiley-Interscience Publication.

[3] Giuliana Davidoff, Peter Sarnak, and Alain Valette. *Elementary number theory, group theory, and Ramanujan graphs*, volume 55 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2003.

[4] Persi Diaconis. *Group representations in probability and statistics*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11. Institute of Mathematical Statistics, Hayward, CA, 1988.

[5] Persi Diaconis. A generalization of spectral analysis with application to ranked data. *Ann. Statist.*, 17(3):949–979, 1989.

[6] Larry Dornhoff. *Group representation theory. Part A: Ordinary representation theory*. Marcel Dekker Inc., New York, 1971. Pure and Applied Mathematics, 7.

[7] William Fulton. *Young tableaux*, volume 35 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1997. With applications to representation theory and geometry.

[8] William Fulton and Joe Harris. *Representation theory*, volume 129 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991. A first course, Readings in Mathematics.

[9] Marshall Hall, Jr. *The theory of groups.* The Macmillan Co., New York, N.Y., 1959.

[10] G. D. James. *The representation theory of the symmetric groups*, volume 682 of *Lecture Notes in Mathematics.* Springer, Berlin, 1978.

[11] Gordon James and Adalbert Kerber. *The representation theory of the symmetric group*, volume 16 of *Encyclopedia of Mathematics and its Applications.* Addison-Wesley Publishing Co., Reading, Mass., 1981. With a foreword by P. M. Cohn, With an introduction by Gilbert de B. Robinson.

[12] Bruce E. Sagan. *The symmetric group*, volume 203 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 2001. Representations, combinatorial algorithms, and symmetric functions.

[13] Jean-Pierre Serre. *Linear representations of finite groups.* Springer-Verlag, New York, 1977. Translated from the second French edition by Leonard L. Scott, Graduate Texts in Mathematics, Vol. 42.

[14] Barry Simon. *Representations of finite and compact groups*, volume 10 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 1996.

Index