BASICS OF REPRESENTATION THEORY

JOSHUA JOHNSON

ABSTRACT. In this paper, we will give a brief introduction to group theory and representation theory. We will primarily be following J.-P. Serre's *Linear Representations of Finite Groups*, focusing on irreducible representations and characters.

CONTENTS

1. INTRODUCTION

It is believed that the origins of group theory date back to Lagrange in the late 18th century, but it wasn't until the work of Cauchy and Galois in the mid-19th century that interest in the field truly began to take off. At its core, group theory is the study of symmetries, bijections of a set to itself, and group theory has become one of the cornerstones of modern mathematics. In this paper, we will study the symmetric group, as it serves as a reference for all other groups.

Sometimes, it is difficult to study a group in the abstract, so representation theory seeks to offer a solution. Representation theory is the study of groups through the lens of linear algebra, allowing us to observe how a group acts on a vector space while making use of all the standard theorems and tools that linear algebra provides. As such, basic knowledge of the fundamentals of abstract linear algebra is assumed throughout this paper.

Representation theory has applications in group theory as a whole. A famous example is Burnside's theorem. Burnside used representation theory in a relatively straightforward proof. It took over half a century to find a proof without using representation theory, and it is considerably more complicated. Representation theory also has applications in the physical sciences, appearing in quantum mechanics to find and exploit the symmetries of wave functions. While often overlooked, representation theory is a fascinating field of study, which we will now develop from the ground up.

Date: August 28, 2024.

2. Groups and Symmetries

Definition 2.1. A *group* G is a set together with a binary operation $* : G \times G \to G$ satisfying the following axioms:

(1) The binary operation is *associative*, i.e., for each $g_1, g_2, g_3 \in G$, $g_1 * (g_2 * g_3) = (g_1 * g_2) * g_3$.
(2) There exists an *identity* element $e \in G$ such that for all $g \in G$, $g * e = e * g = g$.
(3) For each $g \in G$, there exists an *inverse* $g^{-1} \in G$ such that $g * g^{-1} = g^{-1} * g = e$.

It can be easily shown that the identity and inverses are unique. For ease of notation, we will from now on omit the symbol $*$ and simply write $g_1 g_2$.

A particularly important family of groups is the *symmetric group* S_n of permutations on n letters under the binary operation of composition. A *permutation* is a bijection σ from $\{1, 2, \ldots, n\}$ to itself. The permutations of S_n are generally written in *cycle notation*, $(a_1 a_2 \ldots a_k)$ where $a_1, a_2, \ldots, a_k \in \{1, 2, \ldots, n\}$ for some $k \in \mathbb{N}$ and a_i is sent to $a_{i+1 \mod k}$ for all i such that $1 \leq i \leq k$. If an element does not appear in the cycle, it is left as is.

Fundamentally, groups can be thought of as the symmetries of some mathematical object, symmetries which are (fittingly) described by the symmetric group. To see how groups are symmetries, we must first introduce a few more concepts.

Definition 2.2. Let G and H be any two groups. The map $\varphi : G \to H$ is a *homomorphism* if for all $g_1, g_2 \in G$, $\varphi(g_1 g_2) = \varphi(g_1)\varphi(g_2)$. φ is an *isomorphism* if it is both a homomorphism and a bijection, and we say G is *isomorphic to* H, denoted by $G \cong H$.

It is a simple proposition that under any homomorphism φ, $\varphi(e_G) = e_H$, where e_G and e_H are the identity elements in G and H respectively, and $\varphi(g^{-1}) = \varphi(g)^{-1}$.

Often, it is useful to think of isomorphic groups as essentially being the same group with different labels. If G is isomorphic to H, then they have the same group structure and there is one-to-one correspondence between the elements of G and the elements of H. Informally, we can think that G is "basically the same as" H.

Definition 2.3. An *action* of a group G on a set A is a map from $G \times A$ to A (written as $g \cdot a$ for all $g \in G$ and $a \in A$) satisfying the following properties:

(1) $g_1 \cdot (g_2 \cdot a) = (g_1 g_2) \cdot a$ for all $g_1, g_2 \in G$ and $a \in A$.
(2) $e_G \cdot a = a$ for all $a \in A$.

For each fixed $g \in G$, there is a map $\sigma_g : A \to A$, $\sigma_g(a) = g \cdot a$ where

(1) for each fixed $g \in G$, σ_g is a permutation of the elements of A, and
(2) the map from G to $\mathrm{Sym}(A)$, where $\mathrm{Sym}(A)$ denotes the permutation group on the elements of A, defined by $g \mapsto \sigma_g$ is a homomorphism.

We may now state and prove the connection between groups and symmetries.

Theorem 2.4. *(Cayley's Theorem) Let G be a group. Then G is isomorphic to a subgroup of $Sym(G)$.*

Proof. Let $g \in G$. Define the function $f_g : G \to G$, $f_g(x) = gx$. Because g has an inverse g^{-1}, f_g also has an inverse $(f_g)^{-1} = f_{g^{-1}}$. Hence, f_g is injective, and since its domain is the same as its codomain, f_g is bijective. Thus, $f_g \in \mathrm{Sym}(G)$.

Let $A = \{f_g | g \in G\} \subseteq \mathrm{Sym}(G)$ be a subgroup of $\mathrm{Sym}(G)$. We want to show that $G \cong A$. Consider the function $T : G \to A$, $T(g) = f_g$. T is a homomorphism, as for all $g, h \in G$,

$$(f_g \cdot f_h)(x) = f_g(f_h(x)) = f_g(hx) = g(hx) = (gh)x = f_{gh}(x),$$

and hence

$$T(g)T(h) = f_g \cdot f_h = f_{gh} = T(gh).$$

Furthermore, T is injective, as for all $g, g' \in G$,

$$f_g = f_{g'} \iff gx = g'x \iff g = g'$$

for all $x \in G$. Since the domain and codomain of T have the same cardinality, T is bijective and thus an isomorphism. Thus, $G \cong A$. □

Cayley's theorem shows that groups are "basically the same as" symmetries, so we may learn much about groups as a whole simply by studying the symmetric group S_n.

3. Representations

While we have given above the definition of a group G acting on a set A, it is not immediately obvious how it does so. Unless A is the trivial group containing only the identity element, there are multiple ways in which each element of G can permute the elements of A. How then does one describe the various ways in which G can act? This is the goal of representation theory, allowing us to use the tools of linear algebra to analyze group actions.

Definition 3.1. Let G be a finite group. The *general linear group* $GL(V)$ is the group of invertible $n \times n$ matrices over a complex vector space V. A *linear representation* of G is a homomorphism $\rho : G \to GL(V)$. We say that V is a *representation space*, or simply a *representation*, of G. The *degree* of ρ is $\dim(V) = n$. Note that both V and ρ may be referred to as representations.

For ease of notation, we will from now on write ρ_g rather than $\rho(g)$. Here are some simple examples:

- Let $GL(\mathbb{C}) = \mathbb{C}^*$ denote the multiplicative group of complex numbers, that is, $\mathbb{C} \setminus \{0\}$. A degree 1 representation is $\rho : G \to \mathbb{C}^*$, where ρ_g is an nth root of unity for all $g \in G$. If $\rho_g = 1$ for all $g \in G$, this is called the *trivial representation*.
- Specific to the symmetric group, another degree 1 representation is the *sign representation*, $\rho : S_n \to \mathbb{C}^*$, which sends cycles of odd length to 1 and cycles of even length to -1.
- Let $n = |G|$ and let V be a complex vector space of dimension n with basis $(e_h)_{h \in G}$ indexed by the elements of G. For $g \in G$, let $\rho_g : V \to V$ be the linear map that sends e_h to e_{gh}. Then ρ is called the *regular representation*.

There are many representations of a group, but some are, in a sense, more fundamental than others. Most representations are made from "smaller" building blocks called subrepresentations.

Definition 3.2. Let $\rho : G \to GL(V)$ be a representation. A subspace W of V is *stable* under G if for all $w \in W$, $\rho_g w \in W$ for all $g \in G$. The restriction $\rho^W : G \to GL(W)$ is a linear representation called a *subrepresentation*.

Recall from linear algebra that if V is a vector space with subspaces W and W', V is said to be the *direct sum* of W and W', denoted $V = W \oplus W'$, if each $v \in V$ can be written uniquely in the form $v = w + w'$, with $w \in W$ and $w' \in W'$. The mapping p which sends each $v \in V$ to its component $w \in W$ is called the *projection* of V onto W associated with the decomposition $V = W \oplus W'$. The following theorem shows that stable subspaces always come in pairs.

Theorem 3.3. *Let $\rho : G \to GL(V)$ be a linear representation of G in V and W a subspace of V stable under G. Then there exists a compliment W^0 of W which is stable under G.*

Proof. Let W' be an arbitrary complement of W in V, and let p be the corresponding projection of V onto W. Form the average p^0 of the conjugates of p by the elements of G:

$$p^0 = \frac{1}{|G|} \sum_{g \in G} \rho_g \cdot p \cdot \rho_g^{-1}.$$

Since p maps V into W and ρ_g preserves W, it follows that p^0 maps V into W. We have that $\rho_g^{-1} w \in W$ for $w \in W$, so $p \cdot \rho_g^{-1} w = \rho_g^{-1} w$, meaning $\rho_g \cdot p \cdot \rho_g^{-1} w = w$, and hence, $p^0 w = w$. Thus, p^0 is a projection of V into W corresponding to some complement W^0 of W. Furthermore, we may compute that

$$\rho_h \cdot p^0 \cdot \rho_h^{-1} = \frac{1}{|G|} \sum_{g \in G} \rho_h \cdot \rho_g \cdot p \cdot \rho_g^{-1} \cdot \rho_h^{-1} = \frac{1}{|G|} = \sum_{g \in G} \rho_{gh} \cdot p \cdot \rho_{gh}^{-1} = p^0,$$

which implies that

$$\rho_h \cdot p^0 = p^0 \cdot \rho_h$$

for all $h \in G$. Now consider $w \in W^0$ and $h \in G$. We then have $p^0 w = 0$, hence $p^0 \cdot \rho_h w = \rho_h \cdot p^0 w = 0$, implying that $\rho_h w \in W$. Thus, W^0 is stable under G. □

We say that a representation V is *irreducible* if it is nonzero and has no subspace that is stable under G besides V itself and the trivial subspace containing only 0. The following theorem gives a remarkable result concerning irreducible representations.

Theorem 3.4. *Every representation is a direct sum of irreducible representations.*

Proof. Let V be a representation of G. The proof proceeds by induction on $\dim(V)$. Suppose $\dim(V) = 1$. It is a theorem from linear algebra that for any subspace W of V, $\dim(W) \leq \dim(V)$. Hence, W must have dimension zero or one, i.e., the only decomposition of V is the trivial decomposition $V = 0 \oplus V$. Thus, V is irreducible and a direct sum of irreducible representations.

Now suppose $\dim(V) = n$ for some $n \in \mathbb{N}$ and that a representation of dimension k is a direct sum of irreducible representaitons for all $k \in \mathbb{N}$ such that $1 \leq k < n$. If V is irreducible, then we are done. If not, then by Theorem 3.3, V has decomposition $V = V' \oplus V''$, where $\dim(V') < \dim(V)$ and $\dim(V'') < \dim(V)$. By the inductive hypothesis, V' and V'' are direct sums of irreducible representations, and so it follows that V is as well. □

Irreducible representations are truly at the heart of representation theory. As Theorem 3.4 tells us, every representation can be written as a direct sum of irreducible representations. Hence, just as we may study groups by restricting our focus to the symmetric group, we may study representations by restricting our focus to irreducible representations.

4. Characters

An important object of study in representation theory is the *character* of a representation, named as such because a representation is characterized by its character. We will now study the theory of characters and develop methods of obtaining new characters from old ones.

Definition 4.1. Let V be a complex vector space and a a linear map from V to itself with matrix (a_{ij}). The *trace* of (a_{ij}) is the sum along the main diagonal of (a_{ij}) and is given by

$$\mathrm{Tr}(a) = \sum_i a_{ii}.$$

It is a theorem from linear algebra that the trace is also the sum of the eigenvalues of (a_{ij}) (counted with their multiplicities), and does not depend on the choice of basis.

Now let $\rho : G \to GL(V)$ be a representation of a finite group. For each $g \in G$, put

$$\chi_\rho(g) = \mathrm{Tr}(\rho_g)$$

The complex-valued function χ_ρ is called the *character* of the representation ρ. The character of a representation has some useful properties which follow immediately from the definition. Note that we will denote the complex conjugate of a complex number z as $\overline{z}$.

Proposition 4.2. *If χ is the character of a representation ρ of degree n, then we have:*

(1) $\chi(e) = n$,
(2) $\chi(g^{-1}) = \overline{\chi(g)}$ *for* $g \in G$,
(3) $\chi(hgh^{-1}) = \chi(g)$ *for* $g, h \in G$.

Proof. First, the matrix of $\rho(e)$ is simply the $n \times n$ identity matrix, so it has trace n. Second, G is a finite group, so for every $g \in G$, g is of finite order and hence ρ_g is also of finite order. This means that there is some $N \in \mathbb{N}$ such that for all $g \in G$, the matrix of ρ_g^N is the $n \times n$ identity matrix. This implies that all eigenvalues of ρ_g are roots of unity, so their inverses are their complex conjugates. Hence,

$$\chi(g^{-1}) = \mathrm{Tr}(\rho_g^{-1}) = \sum_i \lambda_i^{-1} = \sum_i \overline{\lambda_i} = \overline{\mathrm{Tr}(\rho_g)} = \overline{\chi(g)}.$$

Finally, it is a theorem from linear algebra that for any square matrix A and invertible matrix B, $\mathrm{Tr}(BAB^{-1}) = \mathrm{Tr}(A)$. It then follows that

$$\chi(hgh^{-1}) = \mathrm{Tr}(\rho_{hgh^{-1}}) = \mathrm{Tr}(\rho_h\rho_g\rho_{h^{-1}}) = \mathrm{Tr}(\rho_h\rho_g\rho_h^{-1}) = \mathrm{Tr}(\rho_g) = \chi(g).$$

□

Characters are not difficult to compute when the corresponding representation is given. But we can save ourselves some work by obtaining new characters from ones that we already know. The following proposition shows how the character of a representation that is the direct sum of two representations with known characters may be obtained without any tedious computation.

Proposition 4.3. *Let $\rho^1 : G \to GL(V_1)$ and $\rho^2 : G \to GL(V_2)$ be two linear representations of G, and let χ_1 and χ_2 be their characters. Then the character χ of the direct sum representation $V_1 \oplus V_2$ is equal to $\chi_1 + \chi_2$.*

Proof. Let ρ_g^1 and ρ_g^2 be given in matrix form as R_g^1 and R_g^2 respectively. The representation $V_1 \oplus V_2$ is given by

$$R_g = \begin{pmatrix} R_g^1 & 0 \\ 0 & R_g^2 \end{pmatrix}.$$

It then follows that $\chi(g) = \mathrm{Tr}(R_g) = \mathrm{Tr}(R_g^1) + \mathrm{Tr}(R_g^2) = \chi_1(g) + \chi_2(g)$. □

We may gain a deeper understanding of the behavior of characters by examining how they interact with each other. To do this, we first introduce isomorphic representations and Schur's Lemma.

Definition 4.4. Let ρ^1 and ρ^2 be two representations of the same group G in vector spaces V_1 and V_2 respectively. These representations are said to be *isomorphic* if there exists a linear isomorphism $\tau : V_1 \to V_2$ which satisfies the identity

$$\tau \circ \rho_g^1 = \rho_g^2 \circ \tau$$

for all $g \in G$.

Proposition 4.5. *(Schur's Lemma) Let $\rho^1 : G \to GL(V_1)$ and $\rho^2 : G \to GL(V_2)$ be two irreducible representations of G, and let f be a linear mapping of V_1 into V_2 such that $\rho_g^2 \circ f = f \circ \rho_g^1$ for all $g \in G$. Then:*

1. *If ρ^1 and ρ^2 are not isomorphic, then we have $f = 0$.*
2. *If $V_1 = V_2$ and $\rho^1 = \rho^2$, then f is a homothety (i.e., a scalar multiple of the identity).*

Proof. The case $f = 0$ is trivial. Suppose then that $f \neq 0$, and let W_1 be its kernel (i.e., the set of $v \in V_1$ such that $fv = 0$). If $v \in W_1$, then $f\rho_g^1 v = \rho_g^2 f v = 0$, which implies that $\rho_g^1 v \in W_1$, and W_1 is stable under G. Since V_1 is irreducible, W_1 is either V_1 itself or 0. It cannot be that W_1 is V_1, as that would imply that $f = 0$. A similar argument shows that the image W_2 of f (the set of fv, for $v \in V$) is V_2. These two properties together show that f is an isomorphism, proving the contrapositive of the first part of the proposition.

Suppose now that $V_1 = V_2$, $\rho^1 = \rho^2$, and let λ be an eigenvalue of f; there exists at least one because the field of scalars is the field of complex numbers. Let $f' = f - \lambda$. Then the kernel of f' is nonzero. We also have that $\rho_g^2 \circ f = f \circ \rho_g^1$. The first part of the proposition shows that these properties can only be true if $f' = 0$, that is, if $f = \lambda$. □

Keeping the above hypothesis, we obtain a few corollaries.

Corollary 4.6. *Let f be a linear mapping of V_1 onto V_2, and put:*

$$f^0 = \frac{1}{|G|} \sum_{h \in G} (\rho_h^2)^{-1} f \rho_h^1.$$

Then:

1. *If ρ^1 and ρ^2 are not isomorphic, then we have $f^0 = 0$.*
2. *If $V_1 = V_2$ and $\rho^1 = \rho^2$, then f^0 is a homothety of ratio $(1/n)\,Tr(f)$, with $n = \dim(V_1)$.*

Proof. We may compute

$$(\rho_g^2)^{-1} f^0 \rho_g^1 = \frac{1}{|G|}\sum_{h\in G}(\rho_g^2)^{-1}(\rho_h^2)^{-1} f \rho_h^1 \rho_g^1 = \frac{1}{|G|}\sum_{h\in G}(\rho_{hg}^2)^{-1} f \rho_{gh}^1 = f^0,$$

and hence $\rho_g^2 f^0 = f^0 \rho_g^1$. By Schur's Lemma, in the first case, $f^0 = 0$, and in the second case, $f^0 = \lambda$. Moreover, we have

$$\mathrm{Tr}(f^0) = \frac{1}{|G|}\sum_{h\in G}\mathrm{Tr}((\rho_h^1)^{-1} f \rho_h^1) = \mathrm{Tr}(f),$$

and since $\mathrm{Tr}(f) = n\cdot\lambda$, we obtain $\lambda = (1/n)\cdot \mathrm{Tr}(f)$. □

Now suppose we write ρ^1 and ρ^2 in matrix form:

$$\rho_h^1 = (r_{i_1 j_1}(h)), \quad \rho_h^2 = (r_{i_2 j_2}(h)).$$

The linear map f is defined by a matrix $(x_{i_2 i_1})$, and f^0 is defined by $(x^0_{i_2 i_1})$. The definition of f^0 yields

$$x^0_{i_2 i_1} = \frac{1}{|G|}\sum_{h,j_1,j_2} r_{i_2 j_2}(h^{-1}) x_{j_2 j_1} r_{j_1 i_1}(h).$$

The right hand side is a linear form with respect to $x_{j_2 j_1}$; in the first case, this form vanishes for all systems of values of the $x_{j_2 j_1}$, meaning its coefficients are zero. Thus:

Corollary 4.7. *In case (1), we have:*

$$\frac{1}{|G|}\sum_{h,j_1,j_2} r_{i_2 j_2}(h^{-1}) r_{j_1 i_1}(h) = 0$$

for arbitrary i_2, i_2, j_1, j_2.

In the second case, we have $f^0 = \lambda$, i.e., $x^0_{i_2 i_1} = \lambda\delta_{i_2 i_1}$, where $\delta_{i_2 i_1}$ denotes the Kronecker symbol which is equal to 1 if $i_1 = i_2$ and 0 otherwise. With $\lambda = (1/n)\mathrm{Tr}(f) = (1/n)\sum \delta_{j_2 j_1} x_{j_2 j_1}$, we obtain the equality:

$$\frac{1}{|G|}\sum_{h,j_1,j_2} r_{i_2 j_2}(h^{-1}) x_{j_2 j_1} r_{j_1 i_1}(h) = \frac{1}{n}\sum_{j_1,j_2}\delta_{i_1 i_2}\delta_{j_2 j_1} x_{j_2 j_1}.$$

Equating the coefficients of $x_{j_2 j_1}$, we arrive at:

Corollary 4.8. *In case (2), we have:*

$$\frac{1}{|G|}\sum_{h,j_1,j_2} r_{i_2 j_2}(h^{-1}) r_{j_1 i_1}(h) = \frac{1}{n}\delta_{i_2 i_1}\delta_{j_2 j_1} = \begin{cases} 1/n & \text{if } i_1 = i_2 \text{ and } j_1 = j_2, \\ 0 & \text{otherwise.}\end{cases}$$

In their current forms, these corollaries are a bit unwieldy to use. Let us introduce some notation to help. If ϕ and ψ are functions on G, set

$$\langle \phi, \psi\rangle = \frac{1}{|G|}\sum_{h\in G}\phi(h^{-1})\psi(h) = \frac{1}{|G|}\sum_{h\in G} = \phi(h)\psi(h^{-1}) = \langle \psi, \phi\rangle.$$

Note that $\langle \phi, \psi\rangle$ is linear in ϕ and ψ. With this, corollaries 4.7 and 4.8 become

$$\langle r_{i_2,j_2}, r_{j_1,i_1}\rangle = 0 \quad \text{and} \quad \langle r_{i_2,j_2}, r_{j_1,i_1}\rangle = \frac{1}{n}\delta_{i_2 i_1}\delta_{j_2 j_1}$$

respectively. Suppose furthermore that the matrices $(r_{ij}(h))$ are unitary, i.e., that $r_{ij}(h^{-1}) = \overline{r_{ij}(h)}$; this can always be achieved by a suitable choice of basis. Then these corollaries are nothing more than orthogonality relations for the following scalar product:

$$(\phi|\psi) = \frac{1}{|G|} \sum_{h \in G} \phi(h)\overline{\psi(h)}$$

In particular, if χ is the character of an irreducible representation of G, we have $\chi(h^{-1}) = \overline{\chi(h)}$ (Proposition 4.2), so that $(\phi|\chi) = \langle \phi, \chi \rangle$ for all functions ϕ on G. At last, we may state and prove the orthogonality relations for characters.

Theorem 4.9. *(Orthogonality relations)*

(1) If χ is the character of an irreducible representation, we have $(\chi|\chi) = 1$ (i.e., the norm of χ is 1).
(2) If χ and χ' are the characters of two nonisomorphic irreducible representations, we have $(\chi|\chi') = 0$ (i.e., χ and χ' are orthogonal).

Proof. Let ρ be an irreducible representation with character χ and matrix form $\rho_h = (r_{ij}(h))$. Then $\chi(h) = \sum r_{ii}(h)$, so

$$(\chi|\chi) = \langle \chi, \chi \rangle = \sum_{i,j} \langle r_{ii}, r_{jj} \rangle.$$

By Corollary 4.8, we have $\langle r_{ii}, r_{jj} \rangle = \delta_{ij}/n$, where n is the degree of ρ. Thus,

$$(\chi|\chi) = \frac{1}{n} \sum_{i,j} \delta_{ij} = n/n = 1.$$

The second part of the theorem is proved analogously using Corollary 4.7. □

The orthogonality relations are useful, as they allow us verify whether a character is irreducible. They lead us to a result regarding the isomorphisms of irreducible representations.

Corollary 4.10. *Let V be a linear representation of G, with character ϕ, and suppose V decomposes into a direct sum of irreducible representations:*

$$V = W_1 \oplus \cdots \oplus W_k.$$

Then, if W is an irreducible representation with character χ, the number of W_i isomorphic to W is equal to the scalar product $(\phi|\chi) = \langle \phi, \chi \rangle$.

Proof. Let χ_i be the character of W_i. By Proposition 4.3, we have

$$\phi = \chi_1 + \cdots + \chi_k.$$

Then $(\phi|\chi) = (\chi_1|\chi) + \cdots + (\chi_k|\chi)$. By the orthogonality relations, each term in this sum is either 1 or 0 depending on whether the corresponding W_i is isomorphic to W. This proves the theorem. □

There are two more important relations that reduce the work in computing characters, but to understand where they come from, we must first state and prove a theorem concerning the characters of the regular representation. Recall that the regular representation has a basis $(e_h)_{h \in G}$ such that $\rho_g e_h = e_{gh}$.

Theorem 4.11. *The character r_G of the regular representation is given by:*

$$r_G(g) = \begin{cases} |G| & \text{if } g = e, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. In the case $g = e$, we have that the matrix of ρ_g is the identity matrix of dimension $|G|$, which has trace $|G|$. If $g \neq e$, then the main diagonal of the matrix of ρ_g is all zeroes, as no basis vector is left unpermuted. Hence, the matrix has trace zero. □

Corollary 4.12. *Every irreducible representation W_i is contained in the regular representation with multiplicity equal to its degree n_i.*

Proof. By Corollary 4.10, the number of irreducible representations isomorphic to W_i that are contained in the regular representation is $\langle r_G, \chi_i \rangle$. We then have

$$\langle r_G, \chi_i \rangle = \frac{1}{|G|} \sum_{g \in G} r_G(g^{-1}) \chi_i(g) = \frac{1}{|G|} \cdot |G| \cdot \chi_i(e) = \chi_i(e) = n_i.$$

□

At last, we may state and prove the remaining two relations for irreducible characters.

Corollary 4.13. *(Degree relations)*

(1) The degrees n_i satisfy the relation $\sum_{i=1}^{h} n_i^2 = |G|$.
(2) If $g \in G$ and $g \neq e$, we have $\sum_{i=1}^{h} n_i \chi_i(g) = 0$.

Proof. By Corollary 4.12, each irreducible representation is included in the regular representation with multiplicity equal to its degree. This implies that every irreducible character is included in the character of the regular representation with the same multiplicity. That is, $r_G(g) = \sum_{i=1}^{h} n_i \chi_i(g)$ for all $g \in G$. Letting $g = e$, we obtain (1), and letting $g \neq e$, we obtain (2). □

The degree relations can be used to compute the last irreducible character when all other irreducible characters are known. A fully worked example will be given in the next section.

5. Class Functions

Recall from Proposition 4.2 that characters are invariant under conjugation. They are an example of *class functions*. A class function f is a function which satisfies $f(vuv^{-1}) = f(u)$ for all u, v in the domain of f. We will examine class functions acting on the symmetric group S_n as our motivating example.

Proposition 5.1. *Let f be a class function on G, and let $\rho : G \to GL(V)$ be a linear representation of G. Let ρ_f be the linear mapping of V into itself defined by:*

$$\rho_f = \sum_{h \in G} f(h) \rho_h.$$

If V is irreducible of degree n and character χ, then ρ_f is a homothety of ratio λ given by

$$\lambda = \frac{1}{n} \sum_{h \in G} f(h) \chi(h) = \frac{|G|}{n} (f | \overline{\chi}).$$

Proof. We may compute

$$\rho_g^{-1}\rho_f\rho_g = \sum_{h\in G} f(h)\rho_g^{-1}\rho_h\rho_g = \sum_{h\in G} f(h)\rho_{h^{-1}gh}.$$

Using the substitution $u = h^{-1}gh$, we obtain

$$\rho_g^{-1}\rho_f\rho_g = \sum_{u\in G} f(huh^{-1})\rho_u = \sum_{u\in G} f(u)\rho_u = \rho_f.$$

Hence, $\rho_f\rho_g = \rho_g\rho_f$. By the second part of Schur's Lemma, this shows that ρ_f is a homothety. The trace of ρ_f is $n\lambda$, but it is also $\sum_{h\in G} f(h)\mathrm{Tr}(\rho_h) = \sum_{h\in G} f(t)\chi(h)$. Thus,

$$\lambda = \frac{1}{n}\sum_{h\in G} f(h)\chi(h) = \frac{|G|}{n}(f|\overline{\chi}).$$

□

We can now use this proposition in a proof regarding the vector space H of class functions on G.

Theorem 5.2. *The characters $\chi_1, \dots, \chi_h$ form an orthonornal basis of H.*

Proof. By the orthogonality relations, the χ_i form an orthonormal set of vectors in H, i.e., they are linearly independent. To show that they are a basis of H, we need to show that they span H. To do this, we will show that every element of H that is orthogonal to the $\overline{\chi_i}$ is the zero function.

Let f be such a function, and for each representation ρ, let $\rho_f = \sum_{h\in G} f(h)\rho_h$. Since f is orthogonal to the $\overline{\chi_i}$, by Proposition 5.1, it must be that ρ_f is zero when ρ is irreducible. Since ρ decomposes as a direct sum of irreducible representations, it follows that ρ_f is always zero. In particular, in the regular representation, the image of the basis vector e_1 under ρ_f is given by

$$\rho_f e_1 = \sum_{h\in G} f(h)\rho_h e_1 = \sum_{h\in G} f(h)e_h = 0.$$

This implies that $f(h) = 0$ for all $h \in G$, i.e., that $f = 0$. □

Class functions gives rise to the study of *conjugacy classes*, which are equivalence classes of G, the elements of which are related by conjugation. For example, consider the permutations (12) and (23) in S_3. (23) may be obtained from (12) by the following conjugation:

$$(123)(12)(123)^{-1} = (123)(12)(132) = (123)(13) = (23)$$

Hence, we say (12) and (23) belong to the same conjugacy class. In fact, it is a theorem from combinatorics that all cycles of a given length form a conjugacy class. We state and prove this below.

Lemma 5.3. *Let $\sigma, \tau \in S_n$ where σ is the k-cycle $(a_1a_2\dots a_k)$. Then*

$$\tau\sigma\tau^{-1} = (\tau(a_1)\tau(a_2)\dots\tau(a_k)).$$

Proof. Consider $\tau(a_i)$ such that $1 \le i \le k$. Then $\tau^{-1}\tau(a_i) = a_i$ and $\sigma(a_i) = a_{i+1 \bmod k}$. We then have $\tau\sigma\tau^{-1}(\tau(a_i)) = \tau(a_{i+1 \bmod k})$. Now consider j such that $j \in \{1, 2, \dots, n\}$, but $j \neq a_i$ for any i. Then $\sigma(j) = j$ because j is not in the

k-cycle that defines σ. Hence, $\tau\sigma\tau^{-1}(\tau(j)) = \tau(j)$. Thus, $\tau\sigma\tau^{-1}$ fixes any number not of the form $\tau(a_i)$ for some i, and we have

$$\tau\sigma\tau^{-1} = (\tau(a_1)\tau(a_2)\dots\tau(a_k)).$$

□

Recall that any permutation $\sigma \in S_n$ can be written as a product of disjoint cycles. Suppose we write σ in this way with cycles of length $k_1, k_2, \dots, k_\ell$ where $k_1 \geq k_2 \geq \cdots \geq k_\ell$ and where we include 1's in the list for fixed points. The sequence $(k_1 k_2 \dots k_\ell)$ is called the *cycle-type* of σ. For example, suppose $\sigma \in S_5$ and $\sigma = (12)(35)$. Then the cycle-type of σ is $(2, 2, 1)$. Note that $\sum_{i=1}^{\ell} k_i = n$ as every element of $\{1, 2, \dots, n\}$ is either fixed or appears in some cycle. We may now prove the theorem concerning conjugacy of permutations.

Theorem 5.4. *The conjugacy classes of S_n are determined by cycle-type. That is, if σ has cycle-type $(k_1 k_2 \dots k_\ell)$, then any conjugate of σ has cycle-type $(k_1 k_2 \dots k_\ell)$, and if ρ is any other element of S_n with cycle-type $(k_1 k_2 \dots k_\ell)$, then σ is conjugate to ρ.*

Proof. Suppose $\sigma \in S_n$ has cycle-type $(k_1 k_2 \dots k_\ell)$ so that we may write σ as a produce of disjoint cycles $\sigma = \alpha_1 \alpha_2 \dots \alpha_\ell$ where α_i is a k_i-cycle. Let $\tau \in S_n$. We then have

$$\tau\sigma\tau^{-1} = \tau\alpha_1\alpha_2\dots\alpha_\ell\tau^{-1} = (\tau\alpha_1\tau^{-1})(\tau\alpha_2\tau^{-1})\dots(\tau\alpha_\ell\tau^{-1}).$$

Now for each i such that $1 \leq i \leq \ell$, α_i is a k_i-cycle, and by lemma 4.6, $\tau\alpha_i\tau^{-1}$ is also a k_i-cycle. For any $i, j \in \{1, 2, \dots, \ell\}$ such that $i \neq j$, α_i and α_j are disjoint, and so $\tau\alpha_i\tau^{-1}$ and $\tau\alpha_j\tau^{-1}$ are also disjoint since τ is an injective function. Hence, the above prodcut is $\tau\sigma\tau^{-1}$ written as a product of disjoint cycles, and each $\tau\alpha_i\tau^{-1}$ is a k_i-cycle. Thus, any conjugate of σ has cycle-type $(k_1 k_2 \dots k_\ell)$.

Now let $\sigma, \rho \in S_n$ both be of cycle-type $(k_1 k_2 \dots k_\ell)$, and we will show that they are conjugate. Let σ and ρ be written as products of disjoint cycles as

$$\sigma = \alpha_1\alpha_2\dots\alpha_\ell \quad \text{and} \quad \rho = \beta_1\beta_2\dots\beta_\ell$$

where α_i and β_i are k_i-cycles. For each i, let us write

$$\alpha_i = (a_{i1}a_{i-2}\dots a_{ik_i}) \quad \text{and} \quad \beta_i = (b_{i1}b_{i2}\dots b_{ik_i})$$

Now define τ by $\tau(a_{ij}) = b_{ij}$ for every i, j such that $1 \leq i \leq \ell$ and $1 \leq j \leq k_i$. By Proposition 4.6, $\tau\alpha_i\tau^{-1} = \beta_i$, so it follows that

$$\tau\sigma\tau^{-1} = (\tau\alpha_1\tau^{-1})(\tau\alpha_2\tau^{-1})\dots(\tau\alpha_\ell\tau^{-1}) = \beta_1\beta_2\dots\beta_\ell = \rho.$$

Thus, any two elements in S_n of the same cycle-type belong to the same conjugacy class. □

As class functions, characters are constant on each conjugacy class, but as it turns out, there is a surprising connection between irreducible characters and conjugacy classes.

Theorem 5.5. *The number of irreducible representations of G (up to isomorphism) is equal to the number of conjugacy classes of G.*

Proof. Let $C_1, \dots, C_k$ be the distinct classes of G. A class function f is constant on each of the classes, taking value λ_i on C_i, and these can be chosen arbitrarily. As a result, the dimension of H is equal to k, but as $\chi_1 \dots, \chi_h$ is a basis of H, this implies that $h = k$. Hence, the number of irreducible representations of G is equal to the number of conjugacy classes. $\square$

As there are the same number of conjugacy classes and irreducible representations, it is useful to organize character information in a square table called a *character table* of G. The columns are labeled with the conjugacy classes of G, and the rows are labeled with the irreducible representations. The entries of the table are the characters of the representations evaluated on each conjugacy class.

As an example, let us construct the character table of S_3. To begin, Theorem 5.4 tells us that the conjugacy classes of S_3 are given by the cycles of length 1,2, and 3 respectively, giving $\{1\}$, $\{(12), (13), (23)\}$, $\{(123), (132)\}$. It becomes cluttered to list every element of every conjugacy class, so for convenience we will choose a representative from each class to put in the table. With three conjugacy classes, Theorem 5.5 tells us that there are three irreducible representations. The setup for the table is then:

S_3	e	(12)	(123)
χ_1			
χ_2			
χ_3			

As with any representation, we have the trivial representation, which sends all elements to 1. Thus, they each have character 1.

S_3	e	(12)	(123)
χ_1	1	1	1
χ_2			
χ_3			

We verify that the first orthogonality relation is satisfied:

$$(\chi_1|\chi_1) = \frac{1}{6}[1 + 3(1) + 2(1)] = \frac{6}{6} = 1.$$

Note that the scalar product is taken over all elements of S_3, so we must weight the table entries by the size of the corresponding conjugacy class.

Another representation is the sign representation which sends cycles of odd length to 1 and cycles of even length to -1. Hence, we may fill in the second tow.

S_3	e	(12)	(123)
χ_1	1	1	1
χ_2	1	-1	1
χ_3			

We can check that this representation is irreducible by using the second orthogonality relation:

$$(\chi_1|\chi_2) = \frac{1}{6}[1(1) + 3(-1) + 2(1)] = 0.$$

Now, how to go about filling in the last row? We could guess and check until we happen to stumble across a representation that satisfies the necessary relations.

However, we may save ourselves the trouble and simply use the relations to compute it directly. Using the first degree relation, we know that $\chi_3(e) = n_3$ must satisfy

$$1 + 1 + n_3^2 = 6.$$

The only nonnegative integer satisfying this is $n_3 = 2$. We can now use the second degree relation to compute the rest:

$$\begin{aligned} 1 - 1 + 2\chi_3((12)) = 0 &\iff \chi_3((12)) = 0, \\ 1 + 1 + 2\chi_3((123)) = 0 &\iff \chi_3((123)) = -1. \end{aligned}$$

Thus, the completed table is

S_3	e	(12)	(123)
χ_1	1	1	1
χ_2	1	-1	1
χ_3	2	0	-1

To conclude, let us examine an application of character tables in the space of class functions H. A character table is nothing but a square matrix. Let g_i denote some representative from the conjugacy class C_i. Then the character table is given by

$$\begin{bmatrix} \chi_1(g_1) & \chi_1(g_2) & \dots & \chi_1(g_h) \\ \chi_2(g_1) & \chi_2(g_2) & \dots & \chi_2(g_h) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_h(g_1) & \chi_h(g_2) & \dots & \chi_h(g_h) \end{bmatrix} = \begin{bmatrix} \chi_1^T \\ \chi_2^T \\ \vdots \\ \chi_h^T \end{bmatrix},$$

where each $\chi_i^T = (\chi_i(g_1), \chi_i(g_2), \dots, \chi_i(g_h))^T$ is a row vector.

Now, consider the set of functions $\delta_1, \dots, \delta_h$, where $\delta_i(g)$ is 1 if $g \in C_i$ and 0 otherwise. As there is one δ_i for each conjugacy class C_i, by Theorem 5.5, the size of this set is equal to the dimension of H. Furthermore, each class function f can be written as a linear combination of the δ_i's:

$$f = \sum_{i=1}^{h} f \cdot \delta_i = f \cdot \left(\sum_{i=1}^{h} \delta_i \right)$$

Hence, the δ_i's span H. By the 2-out-of-3 theorem from linear algebra, this is sufficient to show that the δ_i's are a basis of H.

Now, suppose we want to change basis from the δ_i's to the χ_i's. We desire a linear transformation $T : H \to H$ such that $T(\delta_i) = \chi_i$. The matrix of T is then

$$\begin{bmatrix} T(\delta_1) & \dots & T(\delta_h) \end{bmatrix} = \begin{bmatrix} \chi_1 & \dots & \chi_h \end{bmatrix}.$$

But this is exactly the transpose of the character table. Thus, the character table gives us a simple way to compute the change of basis matrix in H.

Acknowledgments

I would like to thank my mentor Jessica Cao for introducing me to representation theory, giving me the idea for this paper, and providing me with so many useful study materials, feedback, and support. I would also like to thank Peter May for organizing the REU and providing feedback on my work. Lastly, I would like to thank Aden Chen for proofreading this paper.

References

[1] J.-P. Serre. Linear Representations of Finite Groups. New York: Springer 1977.
[2] W. Fulton and J. Harris. Representation Theory: A First Course. New York: Springer 1991.
[3] D. Dummit and R. Foote. Abstract Algebra. John Wiley and Sons Inc. 2004
[4] https://www.math.wm.edu/ vinroot/430S13SymConj.pdf
[5] https://en.wikipedia.org/wiki/Cayley