BURNSIDE'S THEOREM

HANLEI WEN

Abstract. Between the late nineteenth century and early twentieth century, the classification of finite simple groups was a popular and significant topic in mathematics. Burnside's Theorem was not only a cornerstone of that exploration, but also a result that gave rise to an entire branch of mathematics – representation theory. This paper will not only elicit a proof of this theorem that combines results branching from multiple areas of mathematics, but will also use the proof of this theorem as a motivation to establish how the beautiful topic of representation theory came into being. In doing so, this paper is mostly self-contained: only linear algebra and fundamental algebraic knowledge are assumed.

Contents

1. Introduction

The general consensus is that abstract algebra – specifically group theory – was first discovered by Galois in the early 19th century. Following him and throughout the twentieth century, mathematicians found group theory to be a powerful tool worthy of investigation. Group theory allowed mathematicians to abstract away specific characteristics of particular groups and make conclusions about abstract groups in general. What was especially marvelous about groups was that there existed building blocks for finite groups just like how there existed buildings blocks – prime numbers – for the natural numbers. These building blocks were called

Date: August 28, 2024.

simple groups.

By the end of the 19th century, the search for a complete set of simple groups became one of the most popular topics in the mathematics world. In this fervent wave of exploration, English mathematician William Burnside proposed and proved in 1904 an important theorem – which we now call Burnside's Theorem – that greatly reduced the range of groups mathematicians had to consider. Not only so, Burnside's exploration and proof of the theorem gave birth to a whole new branch of mathematics – representation theory. This beautiful and significant theorem is stated as follows.

Theorem 1.1. *(Burnside's Theorem) Let G be a group of order $p^a q^b$ with p, q distinct primes. Then G is not simple unless it is of prime order.*

In this paper, we will trace how Burnside proved his renowned and deep theorem, journeying through group theory, representation and character theory and an excursion from number theory. Hence, to keep this paper relatively concise, we will assume familiarity with linear algebra and basic elements from abstract algebra – groups, subgroups, quotient groups, homomorphisms, definitions of rings and fields as well as Galois theory.

2. Important Group Theoretic Results

As a theorem that endeavors to simplify the classification of finite simple groups and one that is fundamentally rooted in group theory, any approach to Burnside's Theorem cannot do away with group theoretic results. In this section, I seek to build upon the basic definitions to present important results that are not only necessary for the proof of Burnside's Theorem, but are significant theorems in their own right. We begin with none other than the renowned Lagrange's Theorem.

2.1. **Lagrange's Theorem.**

The key motivating idea behind Lagrange's Theorem is that there is some relation between the order of a subgroup and the order of the original group. To formalize this idea, let us first make a definition.

Definition 2.1. Let G be a group and let H be a subgroup of G. We denote the order of G by $|G|$. Moreover, the number of left cosets of H in G is called the index of H and is denoted $|G : H|$.

Theorem 2.2. *(Lagrange's Theorem) Let G be a finite group and H be a subgroup of G. Then the order of H divides the order of G (i.e. $|H| \mid |G|$) and the number of left cosets of H is equal to $\frac{|G|}{|H|}$.*

Proof. Proven in a first course in group theory. See page 89 in [1]. □

It follows from Lagrange's Theorem that if $|G|$ is finite, then $|G : H| = \frac{|G|}{|H|}$. Note that this equality does not make sense if $|G|$ is infinite. Now, we can define the order $|x|$ of an element $x \in G$ to be the smallest positive integer n such that $x^n = 1$. A powerful corollary follows from Lagrange's Theorem.

Corollary 2.3. *If G is a finite group and $x \in G$, then $|x| \mid |G|$. In particular, $x^{|G|} = 1$.*

Proof. Left as an exercise for the reader. □

2.2. **Sylow's Theorem.** The full converse of Lagrange's Theorem is not true. That is, given a finite group G and a natural number n such that $n \bigm| |G|$, it is not guaranteed that G has a subgroup with order n. However, there are partial converses to Lagrange's Theorem. For our concern, we will be considering Sylow's Theorem, which is especially useful for proving that a group is not simple.

First, we prove a crucial result that is needed for Sylow's Theorem that is itself an important theorem too.

Theorem 2.4. *(Cauchy's Theorem) Let G be a finite abelian group and let p be a prime dividing $|G|$. Then G contains an element of order p.*

Proof. We proceed by induction on $|G|$. That is, we assume that the result is valid for all groups whose order is smaller than $|G|$ and then prove the result valid for G. The base case where $|G| = 1$ is trivial. Suppose $|G| > 1$. Then, there exists $x \in G$ such that $x \neq 1$.

Suppose $|G| = p$. By Lagrange's Theorem, $x^p = 1$. Moreover, $|x| \neq 1$. Hence, $|x| = p$ and we are done.

Now, suppose $|G| > p$. If p divides $|x|$, write $|x| = pn$. Then, notice that $(x^n)^k \neq 1$ for $k < p$ and $(x^n)^p = x^{(np)} = 1$ so $|x^n| = p$.

Finally, suppose p does not divide $|x|$. Let $N = \langle x \rangle$ where $\langle x \rangle = \{x^a \mid 0 \leq a < |x|\}$ is the subgroup generated by x. Since all subgroups of abelian groups are normal, N is normal. We can hence define the quotient group G/N. By Lagrange's Theorem, we have $|G/N| = \frac{|G|}{|N|}$. Now, $|N| = |x| > 1$ and p does not divide $|x|$ by assumption. Hence, $|G/N| < |G|$ and p must divide $|G/N|$. We can apply our inductive assumption to conclude that G/N contains an element $\bar{y} = yN$ of order p.

Now, $\bar{y} \neq \bar{1}$ since $\bar{y}$ has order $p > 1$ so $y \notin N$. But $\bar{y^p} = \bar{1}$ so $y^p \in N$. This means that $\langle y^p \rangle$ is a subgroup of N, and so $y \notin \langle y^p \rangle$ and hence $\langle y \rangle \neq \langle y^p \rangle$. Notice furthermore that $\langle y^p \rangle \subseteq \langle y \rangle$ and so we have that $|y^p| = |\langle y^p \rangle| < |\langle y \rangle| = |y|$.

Let $|y^p| = n$ and $|y| = m$. We want to show that p divides m. Assume for the sake of contradiction that p does not divide m. Since $y^{pn} = 1$, pn must be a multiple of m (if not we have $y^{pn} = y^{km+l} = y^l \neq 1$ where $l < m$). But

$$p \text{ does not divide } m \iff (p, m) = 1 \iff m \text{ divides } n \iff m \leq n$$

Hence, $|y| = m \leq n = |y^p|$. This contradicts the conclusion above. Therefore, p divides $m = |y|$ and so by the argument above applied to y, we are done. □

Now, there is another result necessary for Sylow's Theorem that demonstrates the relationship between the conjugacy class of g and the normalizer of g.

Lemma 2.5. *Let $g \in G$. Then,*

$$|Cl_g| = |G : N_G(g)|$$

Proof. Proven in a first course in group theory. See page 114 in [1]. □

We can finally move on to Sylow's Theorem and definitions that are specific to the theorem.

Definition 2.6. Let G be a group and p be a prime.

(1) A group of order p^α for some $\alpha \geq 1$ is called a p-group. Subgroups of G which are p-groups are also called p-subgroups.

(2) If G is a group of order $p^\alpha m$ where $p \nmid m$, then a subgroup of order p^α is called a Sylow p-subgroup of G.
(3) The set of Sylow p-subgroups of G will be denoted by $Syl_p(G)$.

Theorem 2.7. *(Sylow's Theorem) Let G be a group of order $p^\alpha m$, where p is a prime not dividing m. Then, Sylow p-subgroups of G exist, i.e., $Syl_p(G) \neq \emptyset$.*

Proof. We proceed by induction on $|G|$. If $|G| = 1$, there is nothing to prove. Assume inductively the existence of Sylow p-subgroups for all groups of order less than $|G|$. Now, consider the center of G, $Z(G)$. Recall that this is the set of all elements in G that commute with every element in G. It can be easily verified that $Z(G)$ is a subgroup of G.

Suppose that p divides $Z(G)$. By Cauchy's Theorem, $Z(G)$ has an element x of order p and hence a subgroup $N = \langle x \rangle$ of order p. Since $Z(G)$ is abelian, N is a normal subgroup. Hence, we can define $\bar{G} = G/N$. By Lagrange's Theorem, $|\bar{G}| = \frac{|G|}{|N|} = \frac{p^\alpha m}{p} = p^{\alpha-1}m$. By our inductive hypothesis, $\bar{G}$ has a subgroup $\bar{P}$ of order $p^{\alpha-1}$.

Let $P = \{x \in G \mid xN \in \bar{P}\}$. I claim that P is a subgroup of G. First, let $x \in P$. By definition, $x^{-1}NxN = (x^{-1}xN) = N = (xx^{-1})N = xNx^{-1}N \implies (xN)^{-1} = x^{-1}N$. Since $\bar{P}$ is closed under inverses, $x^{-1} \in P$. Next, let $x, y \in P$. Then, $xNyN = xyN \in P \implies xy \in P$.

Observe that N is a normal subgroup of P since $nN = N \in \bar{P}$ for all $n \in N$ and N is normal in G and so is also normal when restricted to P. Moreover, we have $P/N = \bar{P} \implies \frac{|P|}{|N|} = |\bar{P}| \implies |P| = |\bar{P}||N| = p^{\alpha-1} \cdot p = p^\alpha$. Hence, P is a Sylow p-subgroup of G.

We are left with the case where p does not divide $Z(G)$. Let $g_1, \dots, g_r$ be representatives of all distinct, non-central conjugacy classes of G. By Lemma 2.5 and the fact that the conjugacy classes partition G, we have

$$|G| = |Z(G)| + \sum_{i=1}^{r} |G : N_G(g_i)|$$

keeping in mind that each element in the center is a conjugacy class by itself. Now, since $|G|$ is divisible by p and $|Z(G)|$ is not divisible by p by assumption, there must be some i such that $|G : N_G(g_i)|$ is not divisible by p. Let $H = N_G(g_i)$. Then by Lagrange's Theorem,

$$|G : H| = \frac{|G|}{|H|} \implies |H| = \frac{|G|}{|G : H|} = \frac{p^\alpha m}{n} = p^\alpha k$$

for some $n, k \in \mathbb{N}$ where $(n, p) = 1$. Moreover, since $|G : H| = |Cl_{g_i}|$ and $g_i \notin Z(G)$, we have $|G : H| > 1 \implies |H| < |G|$. We can now apply our inductive hypothesis to obtain a Sylow p-subgroup of H which is in turn a Sylow p-subgroup of G. This completes the proof of Sylow's Theorem. □

Sylow's Theorem is very powerful when attempting to prove that a particular group is not simple.

Corollary 2.8. *Let G be an abelian group. Then G is simple if and only if it is of prime order.*

Proof. Left as an exercise for the reader. □

Another useful proposition can also be derived through the argument we used to prove Sylow's Theorem.

Proposition 2.9. *Let G be a finite group such that $|G| = p^\alpha$ for some prime p and $\alpha \geq 1$. Then $Z(G) \neq \{1\}$.*

Proof. Let $g_1, \ldots, g_r$ be representatives for all distinct non-central conjugacy classes. As we proved in Sylow's Theorem,

$$|G| = |Z(G)| + \sum_{i=1}^{r} |G : N_G(g_i)|$$

Now, $|G : N_G(g_i)| = \frac{|G|}{|N_G(g_i)|}$ and hence $|G : N_G(g_i)|$ divides $|G|$ for all $1 \leq i \leq r$. Moreover,

$$g_i \notin Z(G) \implies |N_G(g_i)| \neq |G| \implies |G : N_G(g_i)| \neq 1$$

Since $|G| = p^\alpha$ it follows that $|G : N_G(g_i)|$ is divisible by p for all $1 \leq i \leq r \implies \sum_{i=1}^{r} |G : N_G(g_i)|$ is divisible by p. Now, $|G|$ is also divisible by p. Hence, $|Z(G)|$ must also be divisible by p. In particular, $|Z(G)| \neq 1 \implies Z(G) \neq \{1\}$. This completes the proof. □

Observe that the center of the group is always normal, since for all $g \in Z(G)$ and for all $x \in G$, $xgx^{-1} = gxx^{-1} = g$ as g commutes with all $x \implies \forall x \in G,\ xZ(G)x^{-1} = Z(G) \implies Z(G)$ is normal. Therefore, if G is non-abelian and has prime power order, then $Z(G) \neq G$ (by non-abelian) and $Z(G) \neq \{1\}$ (by Proposition 2.9) $\implies G$ is not simple.

3. Representation Theory

3.1. **Introduction to Representations.** Although Burnside's Theorem can be proven with purely group theoretic results, it was originally proven by William Burnside in 1904 through representation theory. Intuitively, representation theory considers a group as an action upon a vector space. Through investigating how a group acts on the vector space (and therefore investigating the corresponding matrix), we can understand more about the structure of the group and its properties. We will prove Burnside's Theorem via this method.

From group theory, we intuitively define a homomorphism to be a structure preserving map. The following proposition elicits this property and will come into play in our proof of Burnside's Theorem.

Proposition 3.1. *Let $\phi : G \to H$ be a homomorphism and let N_H be a normal subgroup of H. Then $N_G = \{g \in G \mid \phi(g) \in N_H\}$ is a normal subgroup of G.*

Proof. First, let us prove that N_G is a subgroup. Evidently, $1 \in N_G$ as $\phi(1) = 1 \in N_H$. Suppose $g \in N_G$, then $\phi(g) \in N_H \implies \phi(g)^{-1} = \phi(g^{-1}) \in N_H \implies g^{-1} \in N_G$. Next, let $g, h \in N_G$. Then $\phi(g)\phi(h) = \phi(gh) \in N_H \implies gh \in N_G$.

Next, we prove that every element in G normalizes N_G. Take $g \in G$. For all $n \in N_G$, $\phi(gng^{-1}) = \phi(g)\phi(n)\phi(g^{-1}) \in \phi(g)N_H\phi(g)^{-1} = N_H \implies gng^{-1} \in N_G$. Hence, $gNg^{-1} \subseteq N$. Now, construct a map $f : N \to N$ defined by $f(n) = gng^{-1}$. $f(n_1) = f(n_2) \implies gn_1g^{-1} = gn_2g^{-1} \implies n_1 = n_2 \implies f$ is injective. Since the size of the domain of f equals the size of the codomain of f, f is also surjective. Hence, f permutes the elements in $N \implies gNg^{-1} = N$. □

Now we can build upon homomorphisms to define a specific type of homomorphism called a representation.

Definition 3.2. Let G be a group. A representation of the group G is a homomorphism $\phi : G \to GL(V)$ for some finite-dimensional non-zero vector space V.

The dimension of V is called the degree of ϕ. Since $\phi(g)$ is a linear map from V to V for $g \in G$, for simplicity we write ϕ_g for $\phi(g)$ and $\phi_g v$ for $\phi(g)(v)$. For our concern, we assume that the field acting on the vector space V is $\mathbb{C}$. Some of our following assertions may not hold if this were not the case (for example, our assumption that there exists an eigenvector).

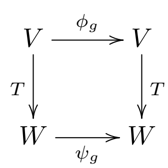
Suppose $\dim V = n$. Recall from linear algebra that the same linear transformation can be represented by different matrices when considered with respect to different bases. Hence, if instead of mapping into a linear transformation we think about mapping into a matrix, the differences caused by choosing a basis becomes problematic. This motivates us to make the following definition.

Definition 3.3. Two representations $\phi : G \to GL(V)$ and $\psi : G \to GL(W)$ are equivalent if there exists an isomorphism $T : V \to W$ such that

$$\psi_g = T\phi_g T^{-1}$$

for all $g \in G$, i.e. $\psi_g T = T\phi_g$ for all $g \in G$. In this case, we write $\phi \sim \psi$.

To further illustrate, consider the diagram



This diagram commutes, meaning that travelling along two different paths will yield the same result. For illustration, if we travel from top left to top right, we have $\phi_g = T\psi_g T^{-1}$; or from top left to bottom right, $\phi_g T = T\psi_g$. To gain some more intuition about the theory developed so far, let us consider some examples.

Example 3.4. (Trivial Representation) Let G be a group. The map $\phi : G \to \mathbb{C}^*$ where $\mathbb{C}^* = \mathbb{C} - \{0\}$ defined by

$$\phi(g) = 1, \ \forall g \in G$$

is called the trivial representation. Depending on convention, there could be different definitions of what a trivial representation is. For our concern in this exposition, the trivial representation is always the homomorphism $G \to \mathbb{C}*$ of degree 1.

Let us verify that the trivial representation is indeed a representation. For $g, h \in G$, notice that we have

$$\phi(gh) = 1 = 1 \cdot 1 = \phi(g) \cdot \phi(h)$$

as desired.

Example 3.5. (Representations of $\mathbb{Z}_n$) Define $\phi : \mathbb{Z}_n \to GL(\mathbb{C}^2) = GL_2(\mathbb{C})$ by

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

Notice that this is the standard rotation matrix in $\mathbb{R}^2$. Hence, $\phi_{\bar{m}+\bar{k}}$ is rotation anticlockwise by angle $\frac{2\pi(m+k)}{n}$ which is rotation anticlockwise by angle $\frac{2\pi m}{n}$ and then by $\frac{2\pi k}{n}$ which is $\phi_{\bar{m}}\phi_{\bar{k}}$, establishing that ϕ is indeed a representation. Now, let us define another representation $\psi : G \to GL_2(\mathbb{C})$ by

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

To verify that it is a representation, note that

$$\psi_{\bar{m}}\psi_{\bar{k}} = \begin{bmatrix} e^{\frac{2\pi mi}{n}} & 0 \\ 0 & e^{\frac{-2\pi mi}{n}} \end{bmatrix} \begin{bmatrix} e^{\frac{2\pi ki}{n}} & 0 \\ 0 & e^{\frac{-2\pi ki}{n}} \end{bmatrix} = \begin{bmatrix} e^{\frac{2\pi(m+k)i}{n}} & 0 \\ 0 & e^{\frac{-2\pi(m+k)i}{n}} \end{bmatrix} = \psi_{\bar{m}+\bar{k}}$$

Now, I claim that $\psi \sim \phi$. To show this, let

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

Then,

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

Finally, direct computation yields

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

establishing the equivalence.

Now, consider the subspace $\mathbb{C}e_1 = \{ce_1 \mid c \in \mathbb{C}\}$ of $\mathbb{C}^2$ where $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Notice that for all $v \in \mathbb{C}e_1$, $\psi_{\bar{m}}(v) = e^{\frac{2\pi mi}{n}}v \in \mathbb{C}e_1$. In other words, no matter what element of $\mathbb{Z}_n$ we take, its action upon the vector subspace $\mathbb{C}e_1$ yields an element still within $\mathbb{C}e_1$. This motivates the following definition.

Definition 3.6. Let $\phi : G \to GL(V)$ be a representation. A subspace $W \leq V$ is G-invariant if, for all $g \in G$ and $w \in W$, $\phi_g(w) \in W$.

This definition allows us to decompose a representation into smaller components.

Definition 3.7. Let there be representations $\phi^{(1)} : G \to GL(V_1)$ and $\phi^{(2)} : G \to GL(V_2)$. Then, we can define their direct sum $\phi^{(1)} \oplus \phi^{(2)} : G \to GL(V_1 \oplus V_2)$ by

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

Now, let us consider the matrix representation of direct sums. Let $\phi^{(1)} : G \to GL_n(\mathbb{C})$ and $\phi^{(2)} : G \to GL_m(\mathbb{C})$. Then, $\phi^{(1)} \oplus \phi^{(2)} : G \to GL_{m+n}(\mathbb{C})$ has block matrix of the form

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

Let us revisit our previous example.

Example 3.8. Define $\phi^{(1)} : G \to \mathbb{C}^*$ by $\phi^{(1)}_{\bar{m}} = e^{\frac{2\pi mi}{n}}$ and $\phi^{(2)} : G \to \mathbb{C}^*$ by $\phi^{(2)}_{\bar{m}} = e^{\frac{-2\pi mi}{n}}$. Then, we have $(\phi^{(1)} \oplus \phi^{(2)})_{\bar{m}} = \begin{bmatrix} e^{\frac{2\pi mi}{n}} & 0 \\ 0 & e^{\frac{-2\pi mi}{n}} \end{bmatrix}$, which is the representation we saw in Example 3.5.

Now, since we can decompose a representation into smaller representations, we can once again ask: are there basic building blocks for representations that form larger representations? This idea motivates the following definitions.

Definition 3.9.
(i) A representation $\phi : G \to GL(V)$ is said to be irreducible if the only G-invariant subspaces of V are $\{0\}$ and V.
(ii) A representation $\phi : G \to GL(V)$ is said to be completely reducible if $V = V_1 \oplus V_2 \oplus \ldots \oplus V_n$ where V_i are G-invariant subspaces and $\phi|_{V_i}$ is irreducible for all $1 \leq i \leq n$.
(iii) A representation $\phi : G \to GL(V)$ is said to be decomposable if there exist non-zero G-invariant subspaces V_1, V_2 with $V_1 \oplus V_2 = V$.

The definitions of equivalence we coined previously happen to work very well with the definitions of irreducibility, complete reducibility, and decomposability. The following lemma concretely demonstrates their relationship.

Lemma 3.10. *Let $\phi : G \to GL(V), \psi : G \to GL(W)$ be representations such that $\phi \sim \psi$. Then*
(i) ϕ is decomposable $\iff$ ψ is decomposable
(ii) ϕ is irreducible $\iff$ ψ is irreducible
(iii) ϕ is completely reducible $\iff$ ψ is completely reducible

Proof. The proof of this lemma is left as an exercise for the reader. □

3.2. **Maschke's Theorem.** It turns out that, just like how prime numbers form all possible numbers and how simple groups form all possible groups, every representation is formed from irreducible representations, i.e., every representation is completely reducible. This allows us to analyze and understand any representation through analyzing its irreducible components. Before stating and proving the theorem that gives us this result, however, we need to first develop some more theory.

Definition 3.11. An inner product on a vector space V is a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ that satisfies the following properties:
(i) $\langle v, c_1 w_1 + c_2 w_2 \rangle = c_1 \langle v, w_1 \rangle + c_2 \langle v, w_2 \rangle$
(ii) $\langle w, v \rangle = \overline{\langle v, w \rangle}$
(iii) $\langle v, v \rangle \geq 0$ and $\langle v, v \rangle = 0$ if and only if $v = 0$

A vector space V equipped with an inner product is called an inner product space.

Then, we can define what a unitary linear transformation is – a linear transformation that, when applied to two vectors in a vector space, does not change their inner product. The following definition formalizes this idea.

Definition 3.12. Let V be an inner product space and let $T \in GL(V)$ be a linear transformation. Then, T is unitary if $\langle T(v), T(w) \rangle = \langle v, w \rangle$ for all $v, w \in V$. The vector space of all unitary linear transformations is denoted $U(V)$.

A correlated definition can be made for representations.

Definition 3.13. Let V be an inner product space. A representation $\phi : G \to GL(V)$ is called unitary if $\forall g \in G$, ϕ_g is unitary. That is, $\langle \phi_g(v), \phi_g(w) \rangle = \langle v, w \rangle$ for all $v, w \in V$. In other words, $\phi : G \to U(V)$.

One might ask, why define unitary representations? It turns out that they are incredibly useful as they satisfy two important properties that, together, will allow us to gain incredible understanding of representations. The two following propositions will each illuminate one important property.

Proposition 3.14. *Let $\phi : G \to GL(V)$ be a unitary representation. Then ϕ is either irreducible or decomposable.*

Proof. Suppose ϕ is not irreducible. Then ϕ must have a G-invariant subspace W of V. Define the orthogonal complement of W,

$$W^{\perp} = \{v \in V \mid \langle v, w \rangle = 0 \, \forall w \in W\}$$

By a theorem in linear algebra, if W is a proper subspace of V, then $W^{\perp}$ must also be a proper subspace of V. Moreover, we have $V = W \oplus W^{\perp}$. Hence, it suffices to prove that $W^{\perp}$ is also G-invariant. Now, this is equivalent to showing that $\langle v, \phi_g(w) \rangle = 0$ for all $v \in W, w \in W^{\perp}$ and $g \in G$. Now,

$$\begin{aligned} \langle v, \phi_g \rangle &= \langle \phi_{g^{-1}}(v), \phi_{g^{-1}} \phi_g(w) \rangle && \phi \text{ is unitary} \\ &= \langle \phi_{g^{-1}}(v), w \rangle && (\phi_{g^{-1}} \phi_g = \phi_{g^{-1}g} = \phi_1) \\ &= 0 && W \text{ is } G\text{-invariant} \implies \phi_{g^{-1}}(v) \in W \end{aligned}$$

Hence, $\phi_g(w) \in W^{\perp} \implies W^{\perp}$ is G-invariant, completing our proof. □

Proposition 3.15. *Every representation of a finite group G is equivalent to a unitary representation.*

Proof. We use an averaging trick that is a key concept in representation theory. Let $\phi : G \to GL(V)$ be a representation with $\dim V = n$. Choose a basis B for V and let $T : V \to \mathbb{C}^n$ taking coordinates with respect to B. Then by letting $\rho_g = T \phi_g T^{-1}$ we construct a new representation $\rho : G \to \mathbb{C}^n$ equivalent to ϕ. Let $\langle \cdot, \cdot \rangle$ be the standard inner product, then let us define a new inner product $(\cdot, \cdot)$ by

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

First, verifying that $(\cdot, \cdot)$ is indeed an inner product is easy and left to the reader. It remains to verify that ρ is a unitary representation under this inner product. Now, for $h \in G$

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

Notice that the map $f : G \to G$ defined by $f(g) = gh$ is bijective. Indeed, to prove injectivity, if $g_1 h = g_2 h$, applying h^{-1} on the right gives $g_1 = g_2$. To prove surjectivity, for all $k \in G$, $f(kh^{-1}) = kh^{-1}h = k$. Hence, sending g to gh simply permutes the elements. Therefore,

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

as desired. □

Combining Lemma 3.10, Proposition 3.14 and Proposition 3.15, we obtain that every representation of a finite group is either irreducible or decomposable. Note that the same conclusion cannot be made for infinite groups.

Theorem 3.16. *(Maschke's Theorem) Every representation of a finite group is completely reducible.*

Proof. Let $\phi : G \to GL(V)$ be a representation of a finite group G. We proceed by induction on $\dim V = n$. Let $n = 1$. Then, ϕ is irreducible since V has no non-zero, proper subspaces. Suppose the statement if true for $\dim V \leq n$. Let $\dim V = n+1$. From our conclusion above, ϕ is either irreducible or decomposable. If the former, we are done. Suppose ϕ is decomposable. Then, we have $V = V_1 \oplus V_2$ where V_1, V_2 are non-zero G-invariant subspaces. It follows that $\dim V_1, \dim V_2 < \dim V$, which by our inductive hypothesis implies that $\phi|_{V_1}, \phi|_{V_2}$ are completely reducible. Let $V_1 = U_1 \oplus \ldots \oplus U_m$ and $V_2 = W_1 \oplus \ldots \oplus W_k$ where U_i, W_j are G-invariant and $\phi|_{U_i}, \phi|_{W_j}$ are irreducible for all $1 \leq i \leq m, 1 \leq j \leq k$. Then, $V = U_1 \oplus \ldots \oplus U_m \oplus W_1 \oplus \ldots \oplus W_k$ and ϕ is completely reducible. □

Maschke's Theorem proves that for all representations, there is some decomposition into irreducible constituents. The natural question to ask then is whether this decomposition is unique. To answer this question, however, we need to develop some more representation theory.

3.3. **Schur's Lemma and Orthogonality Relations.** Similar to how we defined homomorphisms between groups, we can define homomorphisms between representations.

Definition 3.17. Let $\phi : G \to GL(V), \rho : G \to GL(W)$ be representations. A homomorphism between ϕ and ρ is a linear transformation $T : V \to W$ such that $T\phi_g = \rho_g T$ for all $g \in G$.

The set of all homomorphisms from ϕ to ρ is denoted $\text{Hom}_G(\phi, \rho)$. Observe that by definition, $\text{Hom}_G(\phi, \rho) \subseteq \text{Hom}(V, W) =$ the set of all linear transformations from V to W. Moreover, there are a few observations that we leave to the reader to verify.

(1) For any homomorphism $T : V \to W$, $\text{Im}T$ and $\text{Ker}T$ are G-invariant subspaces.
(2) If $\phi : G \to GL(V)$ and $\rho : G \to GL(W)$ are representations, then $\text{Hom}_G(\phi, \rho)$ is a subspace of $\text{Hom}(V, W)$.

Lemma 3.18. *(Schur's Lemma) Let ϕ, ρ be irreducible representations of G, and $T \in Hom_G(\phi, \rho)$. Then either T is invertible or $T = 0$. Consequently,*

(a) If $\phi \nsim \rho$, then $Hom_G(\phi, \rho) = 0$;
(b) If $\phi \sim \rho$, $T = \lambda I$ with $\lambda \in \mathbb{C}$ (i.e. T is a scalar matrix).

Proof. Let $\phi : G \to GL(V)$, $\rho : G \to GL(W)$ and let $T : V \to W$ be in $\text{Hom}_G(\phi, \rho)$. If $T = 0$, we are done; so assume $T \neq 0$. Since $\text{Ker}T$ is G-invariant and $T \neq 0 \implies \text{Ker}T = 0 \implies T$ is injective. Similarly, from $\text{Im}T$ is G-invariant, we have $\text{Im}T = W \implies T$ is surjective. Hence, T is bijective $\implies T$ is invertible.

For (a), assume $\text{Hom}_G(\phi, \rho) \neq 0$, then there exists an invertible $T \in \text{Hom}_G(\phi, \rho) \implies \phi \sim \rho$.

For (b), assume that λ is an eigenvalue of the matrix of T in $\mathbb{C}$. Then, $\lambda I - T$ is not invertible by definition. Since $\mathrm{Hom}_G(\phi, \phi)$ is a subspace and $I \in \mathrm{Hom}_G(\phi, \phi)$, it follows that $\lambda I - T \in \mathrm{Hom}_G(\phi, \phi)$, but all non-zero T are invertible $\implies \lambda I - T = 0 \implies T = \lambda I$. □

Many important corollaries follow from Schur's lemma.

Corollary 3.19. *Let G be an abelian group. Then any irreducible representation of G has degree 1.*

Proof. Left as an exercise for the reader. □

Corollary 3.20. *Let $A \in GL_m(\mathbb{C})$ with $A^n = I$. Then A is diagonalizable and the eigenvalues of A are n^{th} roots of unity.*

Proof. Suppose $A^n = I$. Define a representation $\rho : \mathbb{Z}_n \to GL_m(\mathbb{C})$ by setting $\rho(\bar{k}) = A^k$. Let $\phi^{(1)}, \phi^{(2)}, \dots, \phi^{(m)}$ be a complete set of irreducible representations of $\mathbb{Z}_n$. By Corollary 3.19, $\phi^{(i)}$ has degree 1. Hence, there exists an isomorphism $T \in GL_m(\mathbb{C})$ such that $T^{-1}AT = D$ is diagonal. Then, we have $D^n = (T^{-1}AT)^n = T^{-1}A^nT = T^{-1}IT = I$. Suppose D has eigenvalues $\lambda_1, \dots, \lambda_m$. Then

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

$\implies \lambda_1^n = \ldots = \lambda_m^n = 1$ and hence the conclusion. □

Now, Schur's Lemma tells us more about what the matrix of T looks like. What about the matrix of a unitary, irreducible representation? Specifically, if $\phi : G \to GL_n(\mathbb{C})$ is a matrix, let $\phi_g = (\phi_{ij}(g))$. Then, we have n^2 function $\phi_{ij} : G \to \mathbb{C}$ associated to ϕ. It turns out that these functions have special properties when ϕ is unitary and irreducible.

Theorem 3.21. *(Schur's Orthogonality Relations) Suppose that $\phi : G \to U_n(\mathbb{C})$ and $\rho : G \to U_n(\mathbb{C})$ are inequivalent irreducible representations. Then*

(i) $\langle \rho_{kl}, \phi_{ij} \rangle = 0$

(ii) $\langle \phi_{kl}, \phi_{ij} \rangle = \begin{cases} \frac{1}{n} & \text{if } i = k \text{ and } j = l \\ 0 & \text{else} \end{cases}$

Proof. Let $T : V \to W$. Define

$$T^{\#} = \frac{1}{|G|} \sum_{g \in G} \rho_{g^{-1}} T \phi_g$$

First, we verify that $T^{\#} \in \mathrm{Hom}_G(\phi, \rho)$. By direct computation,

$$\begin{aligned} T^{\#}\phi_h &= \frac{1}{|G|} \sum_{g \in G} \rho_{g^{-1}} T \phi_{gh} \\ &= \frac{1}{|G|} \sum_{x \in G} \rho_{hx^{-1}} T \phi_x && (x = gh) \\ &= \rho_h \frac{1}{|G|} \sum_{x \in G} \rho_{x^{-1}} T \phi_x = \rho_h T^{\#} \end{aligned}$$

(1) By Schur's Lemma, if $\phi \nsim \rho$, then $\mathrm{Hom}_G(\phi, \rho) = 0 \implies T^\# = 0$.
If $\phi = \rho$, then similarly $T^\# = \lambda I$ for some λ. Then, $Tr(\lambda I) = \lambda \dim V = \lambda \deg \phi \implies \lambda = \frac{Tr(T^\#)}{\deg \phi}$. Finally, it follows from $Tr(AB) = Tr(BA)$ that $Tr(T^\#) = Tr(T)$. Hence, $T^\# = \frac{Tr(T)}{\deg(\phi)} I$.

(2) Let us consider $A^\#$ where $A = E_{ki} \in M_{mn}(\mathbb{C})$. Recall from linear algebra that if a matrix M is unitary, then $M^{-1} = \overline{M}^T$. Hence, $\rho_{lk}(g^{-1}) = \overline{\rho_{kl}(g)}$. Now,

$$
\begin{aligned}
A^\#_{lj} &= \frac{1}{|G|} \sum_{g \in G} (\rho_{g^{-1}} A \phi_g)_{lj} \\
&= \frac{1}{|G|} \sum_{g \in G} \left(\sum_{x,y} (\rho_{g^{-1}})_{lx} (A)_{xy} (\phi_g)_{yj} \right) \\
&= \frac{1}{|G|} \sum_{g \in G} \rho_{lk}(g^{-1}) \phi_{ij}(g) && (A)_{xy} \neq 0 \implies x = k, y = i \\
&= \frac{1}{|G|} \sum_{g \in G} \overline{\rho_{kl}(g)} \phi_{ij}(g) = \langle \rho_{kl}, \phi_{ij} \rangle
\end{aligned}
$$

By (2), the value of $\langle \rho_{kl}, \phi_{ij} \rangle$ is the value of $A^\#_{lj}$.
If $\phi \nsim \rho$ then $A^\# = 0$ by (1) gives us (i).
If $\phi = \rho$. Then $A^\# = \frac{Tr(E_{ki})}{\deg \phi} I$ by (1).

(1) If $k \neq i$ then E_{ki} has only 0s on the diagonal $\implies A^\# = 0$.
(2) If $l \neq j$, then since I_{lj} is only 0s on non-diagonals, $A^\#_{lj}$ must also be 0.
(3) If $i = k$ and $l = j$, then $Tr(E_{ki}) = 1$ and $A^\#_{lj} = \frac{Tr(E_{ki})}{n} \implies \langle \phi_{kl}, \phi_{ij} \rangle = \frac{1}{n}$

This gives us (ii) and completes the proof. □

Before recognizing the significance of this orthogonality relation, we need to make the following definition.

Definition 3.22. Let G be a group and define

$$C[G] = \{f \mid f : G \to \mathbb{C}\}$$

$C[G]$ is an inner product space with addition and scalar multiplication given by

$$(f_1 + f_2)(g) = f_1(g) + f_2(g)$$
$$(cf)(g) = c(f(g))$$

and inner product given by

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

$C[G]$ is called the group algebra of G.

Now, let $G = \{g_1, \ldots, g_n\}$. It can be easily verified that the functions $f_i : G \to \mathbb{C}$,

$$f_i(g_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

form a basis of $C[G]$ for $1 \leq i \leq n$. Hence, $\dim C[G] = |G|$. Moreover, from Schur's orthogonality relations, for any irreducible representation $\phi : G \to GL_n(\mathbb{C})$, the d^2 functions $\{\sqrt{d}\phi_{ij} \mid 1 \leq i, j \leq d\}$ form an orthonormal set in $C[G]$. In fact, the entries of all inequivalent irreducible representations form an orthonormal set. But any orthonormal set must have an order less than $\dim C[G] = |G|$. This means that there are at most $|G|$ classes of inequivalent irreducible representations (in particular, there are finitely many equivalence classes of irreducible representations).

3.4. **Character Theory.**

Definition 3.23. Let $\phi : G \to GL(V)$ be a representation. The character of ϕ is a function $\chi_\phi : G \to \mathbb{C}$ defined by

$$\chi_\phi(g) = Tr(\phi_g)$$

A character of an irreducible representation is called an irreducible character.

Observe that since $\phi_1 = I$, then $\chi_\phi(1) = n = \deg \phi$. Moreover, since traces have the property that $Tr(ABC) = Tr(CAB)$ it can easily verified that (1) if $\phi \sim \rho$ then $\chi_\phi = \chi_\rho$ and (2) $\chi_\phi(g) = \chi_\phi(hgh^{-1})$.

The definition of characters gives rise to another variant of Schur's Orthogonality Relations, as stated in the following theorem.

Theorem 3.24. *(First Orthogonality Relations) Let ϕ, ρ be irreducible representations of G. Then*

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

Proof. We may assume WLOG that ϕ and ρ are unitary since χ is constant on equivalent irreducible representations. Then, we have

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

The conclusion follows from Schur's orthogonality relations. □

We are now very close to proving that a representation can be uniquely decomposed into irreducible representations. Before doing so, let us introduce some notation.

Definition 3.25. If V is a vector space and ϕ a representation, let $mV = V \oplus V \oplus \ldots \oplus V$ and $m\phi = \phi \oplus \phi \oplus \ldots \oplus \phi$. If $\rho = m_1\phi^{(1)} \oplus m_2\phi^{(2)} \oplus \ldots \oplus m_s\phi^{(s)}$, then m_i is called the multiplicity of $\phi^{(i)}$ in ρ.

This idea of multiplicity is not entirely well-defined currently since ρ might have different decompositions. If, however, we can compute the multiplicity of an irreducible constituent from ρ, then we can determine that decomposition is indeed unique.

Theorem 3.26. *Let $\phi^{(1)}, \ldots, \phi^{(s)}$ be a complete set of representatives of the equivalence classes of irreducible representations of G and let*

$$\rho \sim m_1\phi^{(1)} \oplus \ldots \oplus m_s\phi^{(s)}$$

Then $m_i = \langle \chi_{\phi^{(i)}}, \chi_\rho \rangle$.

Proof. First, observe that if $\phi = \rho \oplus \psi$, then the block matrix of ϕ has the form

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

and hence $\chi_\phi = Tr(\phi) = Tr(\rho) + Tr(\psi) = \chi_\rho + \chi_\psi$. Therefore, since characters are constant on equivalent representations, we have $\chi_\rho = m_1\chi_{\phi^{(1)}} + \ldots + m_s\chi_{\phi^{(s)}}$ which by first orthogonality relations gives us

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

as desired. □

Now, since we have a finite number of equivalence classes of reducible representations and for any ρ, we can find the multiplicity of any irreducible representation by considering characters, it follows that the decomposition of ρ into irreducible characters is unique up to choices of representatives of equivalence classes. Let us now define a special representation.

Definition 3.27. Let X be a finite set. Define $\mathbb{C}X$ by

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

Moreover, let two elements be considered equal if all coefficients are equal and define addition by pairwise addition of coefficients and scalar multiplication by scalar multiplication on coefficients.

X is synthetically a basis for $\mathbb{C}X$. We can now define what is called a regular representation.

Definition 3.28. Let G be a finite group. The regular representation of G is the homomorphism $L : G \to GL(\mathbb{C}G)$ defined by

$$L_g \sum_{h \in G} c_h h = \sum_{h \in G} c_h gh$$

In other words, the regular representation acts on G via left multiplication. This representation is particularly important as a result of its character and its irreducible constituents.

Proposition 3.29. *The character of the regular representation L is given by*

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

Proof. Let $G = \{g_1, \dots, g_n\}$. Then $L_g g_j = g g_j$. Thus if $[L_g]$ is the matrix with respect to this ordering of G, then we have

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

In particular,

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

from which the conclusion follows. □

Proposition 3.30. *Let G be a finite group. Let $\phi^{(1)}, \dots, \phi^{(s)}$ be a complete set of representatives for distinct equivalence classes of irreducible representations. Moreover, define $d_i = \deg \phi^{(i)}$. Then we have*

$$L_g \sim d_1 \phi^{(1)} \oplus \ldots \oplus d_s \phi^{(s)}$$

Proof. This is easily verifiable by computation. For simplicity, let $\chi_i = \chi_{\phi^{(i)}}$.

$$\langle \chi_i, \chi_L \rangle = \frac{1}{|G|} \sum_{g \in G} \overline{\chi_i(g)} \chi_L(g) = \frac{1}{|G|} \overline{\chi_i(1)} |G| = d_i$$

where the second equality follows from Proposition 3.29. □

4. Burnside's Theorem

We finally have enough theoretical knowledge from both group theory and representation theory to begin proving Burnside's Theorem. Being a theorem about prime numbers, however, Burnside's Theorem still requires some more knowledge – from number theory.

4.1. **Algebraic Integers.**

The subject of algebraic integers is crucially important for the proof of Burnside's Theorem. We will develop some foundation here.

Definition 4.1. A complex number is said to be an algebraic integer if it is the root of a monic polynomial with integer coefficients.

Observe that n^{th} roots of unity are always algebraic integers by definition and eigenvalues of matrices with integer coefficients are also algebraic integers as they are the solutions to the characteristic polynomial of the matrix.

Proposition 4.2. *A rational number r is an algebraic integer if and only if it is an integer.*

Proof. Let $r = \frac{m}{n}$ with $(m, n) = 1$. If r is an algebraic integer, then it is the solution to some polynomial $p(x) = x^k + a_{n-1}x^{k-1} + \ldots + a_0$. Then, we have

$$\frac{m}{n}^k + a_{n-1}\frac{m}{n}^{k-1} + \ldots + a_0 = 0$$
$$\implies m^k + a_{n-1}m^{k-1}n + \ldots a_0 n^k = 0$$
$$\implies m^k = -n(a_{n-1}m^{k-1} \ldots + a_0 n^{k-1})$$

Hence, $n \mid m \implies n = 1 \implies r = m \in \mathbb{Z}$ as desired. □

An important property of algebraic integers is that they form an algebraic ring.

Lemma 4.3. *$y \in \mathbb{C}$ is an algebraic integer if and only if there exists $y_1, \ldots, y_t \in \mathbb{C}$, not all zero, such that*

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

with $a_{ij} \in \mathbb{Z}$ for all $1 \leq i \leq t$.

Proof. ($\Rightarrow$) Let y be a root of $p(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0$ and take $y_i = y^{i-1}$. Then for $1 \leq i \leq n-1$ we have $yy_i = y^i = y_{i+1}$ and for y_n we have $yy_n = y^n = -a_0 - a_1 y_1 - \ldots a_{n-1} y_{n-1}$.

($\Leftarrow$) Let $A = (a_i j)$ 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 $AY = yY$ with $Y \neq 0$ by assumption and so y is an eigenvalue of the $t \times t$ matrix of integer coefficients $\Longrightarrow$ y is an algebraic integer. □

Proposition 4.4. *Algebraic integers form a ring. Moreover, algebraic integers are closed under complex conjugacy.*

Proof. It suffices to show the algebraic integers are closed under taking inverses, products, sums and complex conjugates.

Let α be an algebraic integer. Suppose it is the solution to the monic polynomial $p(x)$. Then either $p(-x)$ or $-p(-x)$ is a monic polynomial and α is a solution to both.
Moreover, $p(\bar{\alpha}) = \overline{p(\alpha)} = 0$.
Now, suppose y, y' are algebraic integers. Choose $y_1, \ldots, y_t \in \mathbb{C}$ not all 0 and $y'_1, \ldots, y'_s \in \mathbb{C}$ not all 0 as defined in Lemma 4.3.
Then, we have

$$(y + y')y_i y'_j = yy_i y'_j + y' y'_j y_i = \sum_{k=1}^{t} a_{ik} y_k y'_j + \sum_{k=1}^{s} a_{jk} y'_k y_i$$

and

$$(yy')y_i y'_j = (yy_i)(y' y'_j) = \left(\sum_{k=1}^{t} a_{ik} y_k\right)\left(\sum_{l=1}^{s} a_{jl} y'_l\right) = \sum_{k,l} a_{ik} a_{jl} y_k y'_l$$

both of which are sums of $y_i y'_j$ which implies that both $y + y'$ and yy' are algebraic integers by Lemma 4.3. □

The following proposition suggests the relevance of number theory in our investigation.

Proposition 4.5. *Let χ be a character of a finite group G. Then $\chi(g)$ is an algebraic integer for all $g \in G$.*

Proof. Let $\phi : G \to GL_m(\mathbb{C})$ be a representation. Let $|G| = n$. Then, $g^n = 1$ and so $\phi_g^n = \phi_{g^n} = I$. By Corollary 3.22, the eigenvalues $\lambda_1, \ldots, \lambda_m$ of ϕ_g are all n^{th} roots of unity. By Proposition 4.4, $\chi(g) = Tr(\phi_g) = \lambda_1 + \ldots + \lambda_m$ is an algebraic integer for all $g \in G$. □

The following thereom further consolidates this relevance.

Theorem 4.6. *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 $\frac{h}{d}\chi_\phi(g)$ is an algebraic integer.*

Proof. Let $C_1, \dots, C_s$ be the conjugacy classes of G and let $|C_i| = h_i$.
(1) Consider $T_i = \sum_{x \in C_i} \phi_x$. Observe that

$$\phi_g T_i \phi_{g^{-1}} = \sum_{x \in C_i} \phi_{gxg^{-1}} = \sum_{y \in C_i} \phi_y = T_i$$

since an action of g upon G is a permutation. Hence, $T_i\phi_g = \phi_g T_i \implies T_i \in \mathrm{Hom}_G(\phi, \phi)$. By Schur's lemma, $T_i = \lambda I$. Now,

$$d\lambda = Tr(T_i) = \sum_{x \in C_i} Tr(\phi_x) = \sum_{x \in C_i} \chi(x) = h_i \chi_\phi(x)$$

$\implies \lambda = \frac{h_i}{d}\chi_\phi(x) \implies T_i = \frac{h_i}{d}\chi_\phi(x) I.$

(2) Now, observe that $T_i T_j = \sum_{x \in C_i} \phi_x \sum_{x \in C_j} \phi_y = \sum_{g \in G} a_{ijg}\phi_g$. We claim that a_{ijg} depends only on the conjugacy class of g. Let $X_g = \{(x, y) \in C_i \times C_j \mid xy = g\}$. We have $|X_g| = a_{ijg}$. Let g' be conjugate to g so that $g' = kgk^{-1}$. Construct a map $\psi : X_g \to X_{g'}$ defined by $\psi(x, y) = (kxk^{-1}, kyk^{-1})$. The inverse map $\psi' : X_{g'} \to X_g$ defined by $\psi'(x', y') = (k^{-1}xk, k^{-1}yk)$ is obvious $\implies \psi$ is a bijection $\implies |X_g| = |X_{g'}|$. Hence, we have $\sum_{g \in G} a_{ijg}\phi_g = \sum_{k=1}^{s} \sum_{g \in C_k} a_{ijk}\phi_g = \sum_{k=1}^{s} a_{ijk} T_k$.

The conclusion follows from (1), (2) and Lemma 4.3. □

The following lemma concerning n^{th} roots of unity is significant in many situations.

Lemma 4.7. *Let $\lambda_1, \dots, \lambda_d$ be n^{th} roots of unity. Then*

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

and equality holds if and only if $\lambda_1 = \ldots = \lambda_d$.

Proof. By the triangle inequality, $|\lambda_1 + \ldots + \lambda_d| \leq |\lambda_1| + \ldots + |\lambda_d|$. Now, consider $v, w \in \mathbb{R}^2$, observe that $||v + w||^2 = ||v||^2 + 2||v||||w|| \cos\theta + ||w||^2 \leq (||v|| + ||w||)^2$ where equality is reached if and only if $\cos\theta = 1 \implies \theta = 0 \implies v$ and w are scalar multiples. But $|\lambda_1| = \ldots = |\lambda_d| = 1 \implies$ equality holds if and only if $\lambda_1 = \ldots = \lambda_d$ as desired. □

Now, let us borrow some definitions and results from field, ring and Galois theory. Let $w = e^{\frac{2\pi i}{n}}$ and let us denote by $\mathbb{Q}[w]$ the smallest subfield of $\mathbb{C}$ containing w. Moreover, let us define Γ to be the group of all field automorphisms $\sigma : \mathbb{Q}[w] \to \mathbb{Q}[w]$ such that $\sigma(r) = r$ for all $r \in \mathbb{Q}$. Then the following lemma holds.

Lemma 4.8. *Let $p(z)$ be a polynomial with rational coefficients. If $\alpha \in \mathbb{Q}[w]$ is a root of p, then so is $\sigma(\alpha)$ for all $\sigma \in \Gamma$.*

Proof. Suppose $p(z) = a_k z^k + \ldots + a_0$. Then,

$$\begin{aligned} p(\sigma(\alpha)) &= a_k\sigma(\alpha)^k + \ldots + a_0 \\ &= \sigma(a_k\alpha^k + \ldots + a_0) \qquad \sigma \text{ is a field automorphism and } \sigma(a_i) = a_i \\ &= \sigma(0) = 0 \end{aligned}$$

□

It immediately follows that if α is an n^{th} root of unity, then so is $\sigma(\alpha)$ and if α is an algebraic integer, then so is $\sigma(\alpha)$. Another result from Galois theory is crucially important.

Theorem 4.9. *Let $a \in \mathbb{Q}[w]$. Then $\sigma(a) = a$ for all $\sigma \in \Gamma$ if and only if $a \in \mathbb{Q}$.*

Proof. Proven in a course in field theory. □

Now, it is proven in a course on rings and fields that the field $\mathbb{Q}[w]$ has dimension $\phi(n)$ as a $\mathbb{Q}$-vector space, where $\phi(n)$ is the number of positive integers less than n that are coprime to n (known as the Euler-totient function). It follows from the fundamental theorem of Galois theory that $|\Gamma| = \phi(n)$. What we really require, however, is $|\Gamma|$ is finite. With this result, the following corollary makes sense.

Corollary 4.10. *Let $\alpha \in \mathbb{Q}$. Then $\Pi_{\sigma\in\Gamma}\sigma(\alpha) \in \mathbb{Q}$.*

Proof. Let $\tau \in \Gamma$. Then,

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

The conclusion follows by Theorem 4.9. □

4.2. **Burnside's Theorem and Proof.** The following technical lemma brings what we have taken from Galois theory together with representation theory to produce a result that is essential for the proof of Burnside's Theorem.

Lemma 4.11. *Let G be a group with order n and let C be a conjugacy class of G. Let $\phi : G \to GL_d(\mathbb{C})$ be an irreducible representation and assume $|C| = h$ is relatively prime to d. Then either*

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

Proof. It suffices to show $\phi_g = \lambda I$ for some $g \in C \implies X(g) = 0$.
From Corollary 4.5 and Theorem 4.6, $\frac{h}{d}\chi(g)$ and $\chi(g)$ are both algebraic integers. Since $(h,d) = 1$, find k, j such that $kh + jd = 1$. Let

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

Let $\lambda_1, \ldots, \lambda_d$ be the eigenvalues of ϕ_g. They are n^{th} roots of unity but not all the same since ϕ_g is diagonalizable but not scalar. Hence, by Lemma 4.7, $\alpha = \left|\frac{\chi(g)}{d}\right| < 1$.

Note also that $\alpha \in \mathbb{Q}[w]$. Let $\sigma \in \Gamma$. $\sigma(\alpha)$ is an algebraic integer and $\sigma(\chi(g))$ is a sum of n^{th} roots of unity, not all equal. By the same reason as above, $\sigma(\alpha) = \left|\frac{\sigma(\chi(g))}{d}\right| < 1$.

Now, let $q = \Pi_{\sigma\in\Gamma}\sigma(\alpha)$. q is an algebraic integer with

$$|q| = |\Pi_{\sigma\in\Gamma}\sigma(\alpha)| = \Pi_{\sigma\in\Gamma}|\sigma(\alpha)| < 1$$

But $q \in \mathbb{Q}$ by Corollary 4.10 $\implies q \in \mathbb{Z}$. But $|q| < 1 \implies q = 0 \implies \sigma(\alpha) = 0$ for some $\sigma \implies \alpha = 0 \implies \chi(g) = 0$. □

We are finally able to put everything together and come close to a proof to Burnside's Theorem.

Lemma 4.12. *Let G be a finite non-abelian group. Suppose there is a conjugacy class $C \neq \{1\}$ of G such that $|C| = p^t$ with p prime and $t \geq 0$. Then G is not simple.*

Proof. Assume for the sake of contradiction that G is simple. Let $\phi^{(1)}, \ldots, \phi^{(s)}$ be a complete set of representatives of irreducible representatives of G, d_i be the corresponding degrees and χ_i the corresponding characters. WLOG let $\phi^{(1)}$ be the trivial representation.

G is simple $\implies \ker\phi^{(k)} = \{1\}$ for $k > 1$ (since the kernel of any homomorphism is normal and $\ker\phi^{(k)} = G \implies \phi^{(k)}$ is trivial). Therefore $\phi^{(k)}$ is injective for $k > 1$ and since G is non-abelian and $\mathbb{C}^*$ is abelian, we have $d_k > 1$ for $k > 1$. Also, as $Z(G)$ is normal and G is non-abelian, we have $Z(G) = \{1\}$. Hence, C has more than one element and so $t > 0$.

Let $g \in C, k > 1$. Let Z_k be the set of all elements of G such that $\phi_g^{(k)}$ is scalar and let $H = \{\lambda I_{d_k} \mid \lambda \in \mathbb{C}^*\}$. H is evidently a subgroup of $GL_{d_k}(\mathbb{C})$ contained within the center and is therefore normal. By Proposition 3.1, Z_k, the pre-image of H under $\phi^{(k)}$, is a normal subgroup of G. If $Z_k = G$, then $d_k = 1$, but we previously showed that $d_k > 1$. Hence, $Z_k = \{1\}$ by simplicity of G. By Lemma 4.11, for $g \neq 1$ and $p \nmid d_k$, we have $\chi(g) = 0$.

Now, let L be the regular representation of G. By Proposition 3.30, $L \sim d_1\phi^{(1)} \oplus d_2\phi^{(2)} \oplus \ldots \oplus d_s\phi^{(s)}$. Since $g \neq 1$, by Proposition 3.29

$$\begin{aligned} 0 = \chi_L(g) &= d_1\chi_1(g) + \ldots + d_s\chi_s(g) \\ &= 1 + \sum_{k=2}^{s} d_k\chi_k(g) \\ &= 1 + \sum_{p \mid d_k} d_k\chi_k(g) \\ &= 1 + pz \end{aligned}$$

where z is an algebraic integer. Therefore $z = -\frac{1}{p}$ is an algebraic integer and hence an integer by Proposition 4.2. This implies that $p = \pm 1$, which is a contradiction. □

We finally have all the components we need to prove the final goal of our exposition – Burnside's Theorem. For convenience, we will state the theorem once again.

Theorem 4.13. *(Burnside's Theorem) Let G be a group of order $p^a q^b$ with p, q distinct primes. Then G is not simple unless it is of prime order.*

Proof. The case for abelian groups is covered by Corollary 2.8. Assume G is non-abelian. If $a = 0$ or $b = 0$, then by Proposition 2.9, $Z(G)$ is non-trivial (and not equal to G since G is non-abelian). Hence, G is not simple.

Assume $a, b > 1$. By Sylow's Theorem, G has a subgroup H of order q^b. Once again by Proposition 2.9, $Z(H)$ is non-trivial. Let $g \neq 1 \in Z(H)$ and let $N_G(g) = \{x \in G \mid xg = gx\}$ be the normalizer of $g \in G$. Since $g \in Z(H)$, for all $h \in H$, $gh = hg \implies h \in N_G(g)$. Hence $H \subseteq N_G(g)$. By Lagrange's Theorem, we have

$$p^a = |G : H| = \frac{|G|}{|H|} = \frac{|G|}{|N_G(g)|} \cdot \frac{|N_G(g)|}{|H|} = |G : N_G(g)||N_G(g) : H|$$

since N_G is a subgroup of G and H is a group (and hence a subgroup of N_G). Hence, $|G : N_G(g)| = p^t$ for some $t \geq 0$. By Lemma 2.5, we have $|Cl_g| = p^t$. Apply Lemma 4.12 to obtain that G is not simple. The proof is complete. □

Finally, one might encounter a restatement of the theorem:

Theorem 4.14. *(Burnside's Theorem) Let G be a group of order $p^a q^b$ with p, q distinct primes. Then G is solvable.*

As our final remark, we will show that this second statement of Burnside's Theorem is actually stronger. In order to do so, we first need to understand the definition of solvability, a concept that arose out of Galois theory.

Definition 4.15. Let G be a group. G is solvable if there exists a list of groups $G_1, G_2, \ldots, G_n$ that satisfy

$$\{1\} = G_1 \trianglelefteq G_2 \trianglelefteq \ldots \trianglelefteq G_n = G$$

and G_{i+1}/G_i is abelian for all $1 \leq i \leq n-1$.

Remark 4.16. Theorem 4.14 is a stronger statement than Theorem 4.13.

Proof. Suppose G is solvable. If G is not simple then we are done. Let G be simple. Suppose for the sake of contradiction that G is not of prime order. Since G is simple, the only possible normal subgroup chain (called a normal series) is $\{1\} \trianglelefteq G$. Since $G/\{1\} = G$, G must be abelian. By Corollary 2.8, G must be of prime order. This is a contradiction.

To show that the converse is not true, consider the group S_5. It is verifiable that S_5 has a proper normal subgroup A_5, so S_5 is not simple. On the other hand, neither A_5 nor S_5 are abelian, so S_5 is not solvable either. Hence, not being simple does not imply being solvable, so the converse cannot hold. □

Acknowledgements

I am deeply indebted to my mentor, Pranjal Warade, for her patient guidance in my investigation of abstract algebra, representation theory and Burnside's Theorem and for her helpful feedback on my draft. I would also like to express thanks towards Professor Justin Campbell for his enlightening talks on the subject of representation and character theory. Finally, I would like to thank Professor Peter May for his hard work in administering and organizing the program and for granting my last minute request to switch to the algebra cohort. This paper could not have been written without their help.

References

[1] Dummit, David S. and Foote, Richard M. Abstract Algebra. John Wiley & Sons, Inc. 2004.
[2] Linman, Julie. Burnside's Theorem. Oregon State University. 2010.
[3] Steinberg, Benjamin. Representation Theory of Finite Groups. School of Mathematics and Statistics, Carleton University. 2009.