ON THE THEORY OF FINITE GROUPS, THEIR LINEAR REPRESENTATIONS, AND BURNSIDE'S p^aq^b THEOREM

PRANAV SHANKAR

1. Introduction

Group theory came about in the early 1800s after mathematicians realized that the idea of the algebraic structures we now call groups made it easier for them to discover more in the various fields of math they were studying, including non-Euclidean geometry, modular arithmetic, and the algebra of quintic polynomials.

As group theory developed, mathematicians started to focus on how groups 'act' on sets, especially vector spaces. This focus spurred William Burnside (the first person to prove, and the namesake of, Burnside's Theorem.), and Ferdinand Frobenius to develop the theory of linear representations. This theory allowed mathematicians to use linear algebra to study nonlinear algebraic structures. Burnside himself used representation theory to find a relatively simple but ingenious proof of Burnside's Theorem, which mathematicians have struggled to find using only group theory (Group-theoretic proofs were found in the 1970s by David Goldschmidt, Helmut Bender, and Hiroshi Matsuyama, but they are much more complicated than Burnside's proof [Linman 2010]). Even though Burnside himself once thought representation theory was useless, it is nowadays applied practically, in crystallography, quantum mechanics, and more! [Bouchard]

This paper will focus on linear representations of finite groups. We will begin with a discussion of basic concepts related to groups and the functions between them. Then, we will discuss linear transformations, linear representations, characters, and, finally, prove Burnside's Theorem and discuss its implications.

2. The Basics of Groups

In this section, we will discuss groups, which you can think of as "alternative number systems", where the "numbers" can go from integers to rigid transformations and more!

Date: July 14, 2024.

2.1. **What is a Group?** Here, we will go over what a group is and some examples involving objects you are familiar with.

Definition 2.1. Consider a set G and a binary operation $*$. (Here, let $gh = g * h$) Then G is a *group* under $*$ if and only if the following 4 conditions (called *group axioms*) are satisfied :

1: For all $g, h \in G$, $gh \in G$. This condition on G is called *closure* under $*$.
2: For all $g, h, j \in G$, $g(hj) = (gh)j$. This condition on $*$ is called *associativity* over G.
3: There exists a unique element $1_G \in G$ such that $1_G * g = g * 1_G = g$. Such an element is called the *identity element* of G.
4: For all non-identity elements $h \in G$, there exists an element $h^{-1} \in G$ such that $hh^{-1} = h^{-1}h = 1_G$, where 1_G is the identity element of G. Then h^{-1} is called the *inverse* of h in G.

Example 2.1. One example of a group is the set $\mathbb{Z}$, under addition. This is a group because:

- The sum of any two integers is an integer, so $\mathbb{Z}$ is closed under addition.
- Addition is trivially associative over $\mathbb{Z}$.
- The unique integer y satisfying $y + z = z + y = z$ for all integers z is 0, so 0 is the additive identity element of $\mathbb{Z}$.
- For every nonzero integer z, the unique integer y satisfying $y + z = z + y = 0$ is $-z$, so $-z$ is the additive inverse of z in $\mathbb{Z}$.

Since $\mathbb{Z}$ is an infinite set, the group described above is an infinite group. In this paper, we are going to focus on finite groups. Finite groups can be thought of as sets of symmetries of a geometric object, under the operation of composition (i.e. applying one transformation after another). One simple example is D_4, the group of symmetries of a square.

The elements of the group are the 'do-nothing' transformation (id), rotations about the center at counterclockwise angles that are multiples of 90 degrees (r1, r2, r3), and reflections about the midlines and diagonals of the square (mv, mh, du, dn). Note that $r1$ is a 90-degree counterclockwise rotation. Diagrams of these transformations is shown in the below figure.

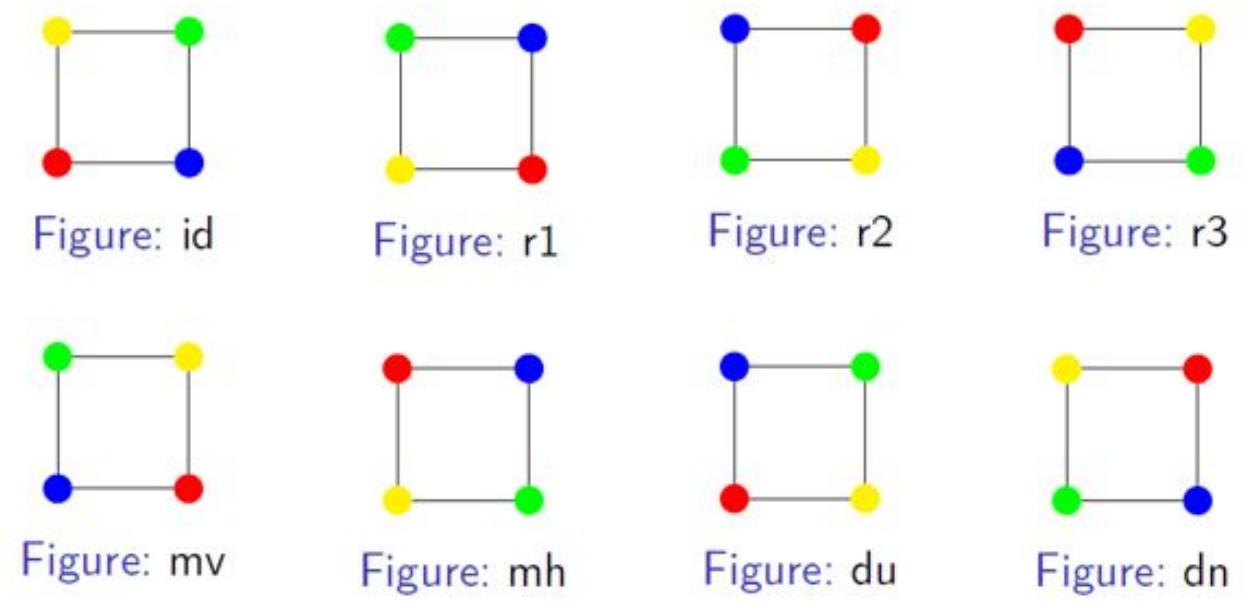

Figure 1. id and the rotations form a subgroup C_4 of D_4

Definition 2.2. A *subgroup* is a subset H of a group G that is itself a group under the same operation as G.

Example 2.2. Take the group D_4, explained above. and remove all the reflections. We get $\{id, r1, r2, r3\}$. The identity element is retained, and the inverse of any rotation is a rotation. We can also check that the resulting subset is closed under composition, and composition is associative over the subset. Thus, the subset of D_4 described is a subgroup of D_4. This subgroup is known as C_4.

Theorem 2.3. *(Subgroup Criterion) Let G be a finite group and let S be a subset of G. Then S is a subgroup of G if and only if, for all $x, y \in S$, $xy^{-1} \in S$.*

Proof. Let's first prove the "only if" part. Let S be a subgroup of G. Then, for all $a, b \in S$, $ab \in S$ and $b^{-1} \in S$ due to group axioms. Thus, plugging b^{-1} in for b in $ab \in S$ gives $ab^{-1} \in S$.

Now, let's prove the "if" part. Let S be a subset of G such that $cd^{-1} \in S$ for all $c, d \in S$. It is clear that the group operation of G is associative over S. We can plug $c = d$ into $cd^{-1} \in S$ to get $1_G \in S$, and we can plug $c = 1$ to get $d^{-1} \in S$ for all $d \in S$. Since $d^{-1} \in S$ for all $d \in S$, we can now replace d with d^{-1} to get $cd \in S$. Thus S is a subset of G that satisfies all the group axioms, so it is a subgroup of G. ■

This criterion is very useful because it greatly simplifies the task of proving that a subset of a given group satisfying certain conditions is a subgroup.

2.2. **Properties of Groups.** Here we will go over properties of groups. This will make you more familiar with how a group "works".

Definition 2.4. The *order* of a finite group G, denoted by $|G|$ is the number of its elements. For example, the order of D_4 is 8 and the order of C_4 is 4.

Definition 2.5. A subset S of a group G not containing the identity element of G *generates* G if all elements of G can be expressed as a "product" of not necessarily distinct elements of S.

Example 2.3. For example, the set $r1, mv$ generates D_4.

Definition 2.6. A finite group is *cyclic* if it has a generating set with only 1 element.

Example 2.4. For example, the group C_4 is cyclic because it is generated by $r1$ The unique group with only 1 element, C_1 is also cyclic, and we call this group the *trivial group*

Definition 2.7. Let G be a finite group and let $g \in G$. Then, there is a unique cyclic subgroup generated by g. The order of this cyclic subgroup is called the *order* of g, and it's denoted by $|g|$. Equivalently, the $|g|$ is the smallest positive integer k such that $|g|^k = 1_G$.

Example 2.5. The order of the identity element of any group is 1, the order of any reflection in $D4$ is 2, the order of rotation $r2$ is 2, and the order of rotations $r1$ and $r3$ are both 4.

2.3. **Left Cosets and Lagrange's Theorem.** Here, we will go over what it means to multiply a group by an element (of that group or some other group), and what the results look like.

Definition 2.8. Let G be a finite group, H be a subgroup of G, and g be a fixed element of G. Then the *left coset* of H corresponding to g , denoted by gH, is the set of possible values of gh, where $h \in H$. The set of left cosets gH of H with respect to G is denoted by G/H(read: G mod H).

Example 2.6. Consider the group C_4. Then the set $\{id, r2\}$ is a group under composition, so it is a subgroup of C_4. Call this C_2. Then, the left coset of C_2 corresponding to $r1$ is $\{r1, r3\}$ because $r1 * id = r1$ and $r1 * r2 = r3$.

Theorem 2.9. *(Lagrange's Theorem). Let G be a finite group and H be a subgroup of G. Then, $|H|$ divides $|G|$.*

We first prove the following:

Lemma 2.10. *Each element of G is in exactly 1 left coset of H.*

Proof. Letting $h = 1_G$, we can see that every element of G is an element of some left coset of H. Now, assume there exists two cosets, let's say aH and bH, of H with at least 1 common element. Then, there are elements $h_1, h_2 \in H$ such that $ah_1 = bh_2$. Then $b = ah_1h_2^{-1}$. Now, let $c \in bH$. Then, $c = bh_3 = ah_1h_2^{-1}h_3 \in aH$. Thus, $bH \subset aH$. The same argument, but with b and a reversed, gives $aH \subset bH$. Thus, $aH = bH$, so we can't have an element of G in more than 1 left coset of H. Thus, every element of G is in exactly 1 left coset of H. ■

Lemma 2.11. *Let G be a group, and let H be a subgroup of G. Then, if $g \in H$, then $gH = H$. Equivalently, left-multiplying a group by an element of that group leaves the group invariant.*

Proof. Since H is a group, $gh \in H$ for all $g, h \in H$. Thus, $gH \subset H$. Now, let $h_1, h_2 \in H$ such that $gh_1 = gh_2$. Then $h_1 = h_2$ by left-multiplication by g^{-1}, so $|gH| = |H|$. Thus $gH = H$.

Since every element of G is an element of exactly 1 left coset of H (by Lemma 2.10), and the order of each of these cosets is $|H|$ by Lemma 2.11, the number of left cosets of H is $|G/H| = \frac{|G|}{|H|}$. Since this is a number of cosets, it is an integer. Thus, $|H|$ divides $|G|$. ■

3. Conjugacy and Commuting

In this section, we will discuss 2 important, related concepts in group theory, called conjugacy and commuting, both based on the order of arguments of the group operation.

3.1. Conjugacy and Normal Subgroups.

Here, we will get our first glimpse into how a group "acts" on itself. We will go over an action called conjugation, which acts as a gateway to many important theorems in group theory, including Cauchy's and Sylow's Theorems.

Definition 3.1. Two elements g_1 and g_2 of a finite group G are *conjugate* if there exists $h \in G$ such that $h^{-1}g_1h = g_2$.

Example 3.1. For example, in the group D_4, the elements $r1$ and $r3$ are conjugate because mv is its own inverse and $mv * r1 * mv = r3$.

Proposition 3.2. *Let G be a group. There exists a set of disjoint subsets of G such that:*

- *Every element of G is in one of these subsets.*
- *Any two elements of the same subset are conjugate.*
- *No two elements of different subsets are conjugate.*

These subsets are called the conjugacy classes *of G.*

Proof. It is obvious that conjugacy is symmetric (i.e., order of conjugacy doesn't matter). We can also see that conjugacy is reflexive (i.e. any element of G is conjugate with itself) as letting $h = 1$ gives $h^{-1}g_1h = g_1$.

Now, let $a, b, c \in G$ such that a is conjugate with both b and c. Then there exists elements $g_3, g_4 \in G$ such that $g_1ag_1^{-1} = b$ and $g_4bg_4^{-1} = c$. Then, $g_4g_3ag_3^{-1}g_4^{-1} = c$. But we note that $(g_4g_3)(g_3^{-1}g_4^{-1}) = (g_3^{-1}g_4^{-1})(g_4g_3) = 1_G$, so $g_3^{-1}g_4^{-1} = (g_4g_3)^{-1}$. Thus, $g_4g_3a(g_4g_3)^{-1} = c$, so a and c are conjugate. Hence, any two elements that are conjugate with any common element are conjugate, so conjugacy is transitive.

Thus, conjugacy is an equivalence relation on G, so it partitions G into equivalence classes satisfying the conditions outlined in the statement of the proposition. ■

Definition 3.3. Let G be a finite group. Then a subgroup H of G is *normal* if any element of G that is conjugate with an element of H is an element of H. Equivalently, a subgroup of G is normal if and only if it is also a union of conjugacy classes of G.

Example 3.2. We showed, in the previous section, that C_4 is a subgroup of D_4. In fact, C_4 is a normal subgroup of D_4; $\{id\}$, $\{r2\}$, $\{r1, r3\}$, $\{mh, mv\}$, and $\{du, dn\}$ are the conjugacy classes of D_4, and $C_4 = \{id, r1, r2, r3\}$ is just the union of the first 3 of these conjugacy classes.

Definition 3.4. A group G is *simple* if and only if it has no normal subgroups other than C_1 and G.

Example 3.3. The group C_2 is simple; its only subgroups are C_1 and C_2. However, C_4 isn't simple, as C_2 is a normal subgroup of C_4.

3.2. **Commuting, Abelian Groups, and Centralizers.** The order of group operands is a fundamental concept in group theory. In addition and multiplication, order doesn't matter, but that's not necessarily the case in group theory. However, there are some cases where the order of group operands doesn't matter, and we will discuss these cases here.

Definition 3.5. Let G be a finite group and let $g_1, g_2 \in G$. Then, g_1 and g_2 *commute* if $g_1g_2 = g_2g_1$.

Example 3.4. For example, in D_4, mv and mh commute because $mv * mh = mh * mv = r2$.

Definition 3.6. A group G is called *abelian* if any 2 elements of G commute. Equivalently, G is abelian if the group operation is commutative over G.

Example 3.5. The group C_4 is abelian because when you compose a pair of rotations, the order of the rotations doesn't matter. By this same logic, all cyclic groups are abelian (that's trivial for C_1 and C_2, and all other cyclic groups can be represented as rotational symmetry groups of regular polygons). However, D_4 is not abelian; $r1 * mv = du$ and $mv * r1 = dn \neq du$.

Proposition 3.7. *Let G be a finite abelian group. Then the conjugacy classes of G are just the 1-element subsets of G.*

Proof. Let g_1, g_2 be conjugate elements of an abelian group G. Then, there is an element $h \in G$ such that $h^{-1}g_1h = g_2$. However, since G is abelian, $g_1h = hg_1$, so $g_2 = h^{-1}g_1h = h^{-1}hg_1 = g_1$. Thus, any 2 conjugate elements of G are equal, so the conjugacy classes of G are 1-element subsets of G. ■

Corollary 3.8. *Let G be a finite abelian group. Then all subgroups of G are normal.*

Proof. By Proposition 3.7, the conjugacy classes of G are the single-element subsets of G. Thus, every subgroup (and, in fact, every nonempty subset) of G is a union of conjugacy classes of G. Hence, all subgroups of G are normal. ■

Proposition 3.9. *Let G be a finite group and H be a normal subgroup of G. Then, G/H is a subgroup of G. The group operation of such a subgroup is the operation $*$ such that $gJ * hJ = ghJ$ for all elements $g, h \in G$ and subgroups J of G.*

Proof. Note that there exists some subset $S$of G such that right–multiplying the elements of S by H gives the set G/H. Thus G/H is isomorphic to S. Let g_1, g_2 be elements of G. Then, $g_1H, g_2H \in G/H$. Since G is a group,

$$g_1g_2^{-1} \in G, so (g_1H)(g_2H)^{-1} = g_1g_2^{-1}H \in G/H.$$

Thus, by the subgroup criterion, S is a subgroup of G, so G/H is isomorphic* to a subgroup of G.

* See Section 4 on Group Homomorphisms. ■

Theorem 3.10. *(Cauchy's Theorem) Let G be a nontrivial finite abelian group. Then, for any prime p dividing $|G|$, there exists an element of G with order p.*

Proof. . Fix a prime number q and assume that the problem statement is not true. Then, there exists at least 1 finite abelian group with minimal order divisible by q that has no element of order q. Call one such group H. Then,every subgroup of H has an element of order q.

Take a nonidentity element h of H. Then h generates a unique cyclic subgroup J of H. Since H is abelian, J is a normal subgroup of H by Corollary 3.8, so, by Proposition 3.9, H/J is a subgroup of H. By Lagrange's Theorem, $|J||H/J| = |H|$. Since $|H|$ is divisible by q, and q is prime, at least one of $|J|$ and $|H/J|$ is divisible by q. WLOG, assume $|J|$ is divisible by q. Then, by our assumption in the first paragraph, J has an element j of order q. Then j generates a cyclic subgroup of J of order q. Since J is a subgroup of H, the cyclic subgroup generated by j is also a subgroup of H, so the order of j in H is also q. Hence, no group H satisfying the conditions in the first paragraph exists. Thus, by the well-ordering

principle, if G is abelian and $|G|$ is divisible by a prime p, then G has an element of order p. ■

Definition 3.11. Let G be a finite group and S be a subset of G. The *centralizer* of S in G, denoted by $C_G(S)$, is the set of elements of G that commute with every element of S.

Example 3.6. The centralizer of $\{r1, r2\}$ in D_4 is C_4 because id and all of the rotations in D_4 commute with both $r1$ and $r2$, but none of the reflections commute with $r1$.

Proposition 3.12. *Let G be a finite group and let S be a nonempty subset of G. Then $C_G(S)$ is a subgroup of G.*

Proof. Let $g, h \in C_G(S)$. Then, $gs = sg$ and $hs = sh$ for all $h \in C_G(S)$, so $gsg^{-1} = h^{-1}sh = s$. Thus $gh^{-1}shg^{-1} = (gh^{-1})s(gh^{-1})^{-1}) = s$, so $gh^{-1}s = sgh^{-1}$. Then $gh^{-1} \in C_G(S)$, so, by the Subgroup Criterion, $C_G(S)$ is a subgroup of G. ■

Definition 3.13. Let G be a finite group. The *center* of G, denoted by $C_G(G)$ is the set of elements of G that commute with every element of G. Equivalently, as sugggested by the notation, the center of G is the centralizer of G in itself. An element of the center of G is called *central* in G.

Example 3.7. The center of the group D_4 is C_2, because only id and $r2$ commute with every element of D_4 (Neither $r1$ nor $r3$ commute with any of the reflections in D_4).

Proposition 3.14. *Let G be a finite group. Then, every central element of G is in its own conjugacy class.*

Proof. Let $g \in C_G(G)$ and $h \in G$. Consider a conjugate $c = h^{-1}gh$ of g. Then, since g is central in G, $c = h^{-1}gh = h^{-1}hg = g.$, so g has no conjugates other than itself. Thus g is in its own conjugacy class. ■

Corollary 3.15. *Let G be a finite group. Then, $C_G(G)$ is a normal abelian subgroup of G.*

Proof. By Proposition 3.12, $C_G(S)$ is a subgroup of G for all subsets S of G. Thus $C_G(G)$ is a subgroup of G. Also, since all elements of $C_G(G)$ commute with all elements of G, they all commute with each other, so $C_G(G)$ is an abelian subgroup of G. By Proposition 3.14, every element of $C_G(G)$ is in its own conjugacy class of G, so $C_G(G)$ is a union of conjugacy classes of G. Hence $C_G(G)$ is a normal, abelian subgroup of G. ■

Theorem 3.16. *Let G be a finite group and let g be an element of G. Then, the conjugacy class of g (which we call $K_G(g)$) and $G/C_G(\{g\})$ have the same number of elements.*

Proof. Note that the set $G/C_G(\{g\})$ exists, because $C_G(\{g\})$ is a subgroup of G by Proposition 3.12. Consider the map f from $K_G(g)$ to $G/C_G(\{g\})$ such that $f(h^{-1}gh) = hC_G(\{g\})$ for all $h \in G$.

We first claim that f is a function. To prove this, let $h_1, h_2 \in G$ such that $h_1^{-1}gh_1 = h_2^{-1}gh_2$. Multiplying both sides by h_2 on the left and h_1^{-1} on the right gives $h_2h_1^{-1}g = gh_2h_1^{-1}$. Therefore, $h_2h_1^{-1}$ commutes with g and is thus in $C_G(g)$. Let $h_2h_1^{-1} = j$. Then $h_2 = jh_1$ and $j \in C_G(\{g\})$, so $h_1C_G(\{g\}) = h_2C_G(\{g\})$ by Lemma 2.11. Hence f is a function.

We then claim that f is injective. To prove this, let $h_1, h_2 \in G$ such that $h_1C_G(\{g\}) = h_2C_G(\{g\})$. Then there exist $g_1, g_2 \in C_G(\{g\})$ such that $h_1g_1 = h_2g_2$. Left-multiplying both sides by h_2^{-1} and right multiplying both sides by g_1^{-1} gives $h_2^{-1}h_1 = g_2g_1^{-1} \in C_G(\{g\})$. Therefore, $gh_2^{-1}h_1 = h_2^{-1}h_1g$. Left-multiplying both sides by h_2 and right multiplying both sides by h_1^{-1} gives $h_2gh_2^{-1} = h_1gh_1^{-1}$. Replacing h_1 and h_2 with their inverses gives $h_2^{-1}gh_2 = h_1^{-1}gh_1$. Thus, f is injective.

It is clear that f is surjective because any left coset of $C_G(\{g\})$ can be expressed as $jC_G(\{g\})$ for some $j \in G$, and for that j, $j^{-1}gj$ exists. Thus, f is a bijection so $K_G(g)$ and $G/C_G(\{g\})$ have the same number of elements. ■

Corollary 3.17. *(Class Equation). Let $g_1, g_2, \dots g_k$ be representatives of all the noncentral conjugacy classes of G. Then $|G| = C_G(G) + \sum_{j=1}^{k} |G/C_G(\{g_j\})|$.*

Proof. By Theorem 3.16, $|K_G(g)| = |G/C_G(\{g\})|$ for all $g \in G$. Thus,

$$C_G(G) + \sum_{j=1}^{k} |G/C_G(\{g_j\})| = \sum_{j=1}^{k} |K_G(g_j)| + C_G(G).$$

The RHS of this equation is just the order of the center of G plus the sum of the sizes of the noncentral conjugacy classes of G. This is simply $|G|$. ■

4. Group Homomorphisms and Solvability

In the previous sections, we discussed groups and their properties. Now, we will discuss functions between groups, Then we will discuss a special property of some groups known as solvability.

4.1. **Homomorphisms.**

Here we are going to talk about functions between groups that work 'neatly' with the operations of the groups. These fuctions are called homomorphisms.

Definition 4.1. For groups G and H with group operations $*$ and $\Diamond$, respectively, a *homomorphism* from G to H is a function f from G to H such that $f(g_1)\Diamond f(g_2) = f(g_1 * g_2)$ for all $g_1, g_2 \in G$.

Definition 4.2. Let G and H be groups. The *nullspace* of a homomorphism f from G to H, denoted by $Null(f)$ is the set of all $g \in G$ such that $f(g) = 1_H$.

Definition 4.3. Let G and H be groups. The *range* of a homomorphism f, from G to H, denoted by $Rng(f)$, is the set of all $h \in H$ such that there exists some $g \in G$ such that $f(g) = h$.

Definition 4.4. An *isomorphism* is a bijective homomorphism. If there exists an isomorphism between two groups, these groups are *isomorphic*. Isomorphic groups are the "same" group, up to relabeling of elements (i.e. they have the same order and same structure). In the theorems and proofs below, we say that groups are equal if and only if they are isomorphic.

Example 4.1. Consider the groups $S_3 = \{123, 132, 231, 213, 312, 321\}$, the group of permutations of the set $\{1, 2, 3\}$, and $D_3 = \{id, r1, r2, fA, fB, fC\}$, the symmetry group of an equilateral triangle ABC, where id is the "do-nothing" transformation, r1 and r2 are 120-degree counterclockwise and clockwise rotations about the center of the triangle, fA is the reflection about the A-altitude, and fB and fC are defined similarly. These groups are isomorphic; any function f from S_3 to D_3 such that no two inputs map to the same output, $f(123) = id$ and $f(231)$ and $f(312)$ map to $r1$ and $r2$ in some order, is an isomorphism from S_3 to D_3. Thus $S_3 = D_3$.

We now prove several theorems about group isomorphisms.

Theorem 4.5. *(1st Isomorphism Theorem) Let G and H be groups and let f be a homomomorphism from G to H. Then $G/Null(f) = Rng(f)$.*

Proof. We first prove the following lemma.

Lemma 4.6. *$Null(f)$ is a normal subgroup of G.*

Proof. Since f is a homomorphism, it is compatible with the group operation of G. Let $g, j \in Null(f)$. Then $f(gj^{-1}) = f(g)f(j^{-1}) = f(g)(f(j))^{-1} = 1_H(1_H^{-1}) = 1_H$. Thus, $gj^{-1} \in Null(f)$, so $Null(f)$ satisfies the Subgroup Criterion. Thus $Null(f)$ is a subgroup of G.

Now, let $c = jnj^{-1}$ be a conjugate of an element of $Null(f)$, where $j \in G$ and $n \in Null(f)$. Then $f(n) = 1_H$ so

$$f(c) = f(j)f(n)f(j^{-1}) = f(j)f(j^{-1}) = 1_H.$$

Thus $c \in Null(f)$, so $Null(f)$ is a normal subgroup of G. ■

Define a map m from $G/Null(f)$ to Rng(f) such that $m(gNull(f)) = f(g)$ for all $g \in G$. Since f is a bijection, m is clearly a function (since every $g \in G$ maps to some $f(g)$ under f), and it is surjective (since every $f(g)$ has a $g \in G$ that maps to it). Now, let $f(g_1) = f(g_2)$ for $g_1, g_2 \in G$. Then $g_1 = g_2$ by bijectivity of F, so $g_1Null(f) = g_2Null(f)$. Thus, m is a bijection.

Now, let $g_1, g_2 \in G$. Then, $m(g_2Null(f)g_1Null(f)) = m(g_2g_1Null(f)) = f(g_2g_1) = f(g_1)f(g_2)$.

Thus, m is a homomorphism. Since we already established that m is bijective, it is an isomorphism. Thus $G/Null(f) = Rng(f)$. ■

Definition 4.7. Let G be a finite group. The *normalizer* of a subset S of G, denoted by $N_G(S)$, is the set of elements h such that $hSh^{-1} = S$ (ie S is fixed under conjugation by h).

Proposition 4.8. *For any group G and subset S of G, $N_G(S)$ is a subgroup of G.*

Proof. Let $g, h \in N_G(S)$. Then, $gSg^{-1} = hSh^{-1} = S$. Then, we can see that $h^{-1}Sh = h^{-1}hSh^{-1}h = S$. Thus,

$$(gh^{-1})S(gh^{-1})^{-1} = gh^{-1}Shg^{-1} = gSg^{-1} = S.$$

Hence $N_G(S)$ satisfies the Subgroup Criterion, so it's a subgroup of G. ■

Theorem 4.9. *(2nd Isomorphism Theorem) Let A and B be subgroups of a group G such that A is also a subgroup of $N_G(B)$, and let AB is the set of values of ab for some $a, b \in G$. Then,*

- *AB is a subgroup of G*
- *B is a normal subgroup of AB*
- *$A \cap B$ is a normal subgroup of A*
- *$AB/B = A/(A \cap B)$.*

Proof. We first prove that AB is a subgroup of G. It is clear that $AB \subset G$. Let $x, y \in AB$. Then, there exists $x_1, y_1 \in A$ and $x_2, y_2 \in B$ such that $x = x_1x_2$ and $y = y_1y_2$. Then

$$xy^{-1} = x_1x_2(y_1y_2)^{-1} = x_1x_2y_2^{-1}y_1^{-1}.$$

Since A is a subgroup of $N_G(B)$,

$$x_1Bx_1^{-1} = y_1By_1^{-1} = B, so x_1^{-1}x_2x_1 \in B.$$

Therefore $x_2x_1 = x_1x_1^{-1}x_2x_1 \in AB$, so $BA \subset AB$. Thus $BA = AB$. Now, notice that $x_2y_2^{-1}y_1^{-1} \in BA$, so $x_2y_2^{-1}y_1^{-1} \in AB$. Thus, $xy^{-1} = x_1x_2y_2^{-1}y_1^{-1} \in AAB = AB$. Therefore $AB \subset G$ that satisfies the Subgroup Criterion, so it is a subgroup of G.

Note that B is a subgroup of AB because B is a group and $1_GB = B$. Since B is a group, it is fixed under conjugation by any element of itself, so B is a subgroup of $N_G(B)$. Since A is also a subgroup of $N_G(B)$ and $N_G(B)$ is a group (by Proposition 4.8), AB is a subgroup of $N_G(B)$. Thus B is fixed under conjugation by any element of AB, so B is a normal subgroup of AB.

It is clear that $A \cap B \subset A$. Now, let $j_1, j_2 \in A \cap B$. Then, since A and B are both groups $j_1j_2^{-1} \in A$ and $j_1j_2^{-1} \in B$. Thus, $j_1j_2^{-1} \in A \cap B$, so by the Subgroup Criterion, $A \cap B$ is a subgroup of A. Since A is a subgroup of $N_G(B)$, B is fixed under conjugation by any element of A, so B (which includes $A \cap B$) is fixed under conjugation by any element of A. Thus $A \cap B$ is a normal subgroup of A.

Now, let f be a map from A to AB/B such that $f(a) = aB$ for all $a \in A$ and $b \in B$. It is clear that f is a surjective function because a uniquely determines aB and every element of AB/B is a left coset $abB = aB$ (these cosets are equal by Lemma 2.11) that corresponds to at least one a. Therefore, $Rng(f) = AB/B$. Now, let $a_1, a_2 \in A$. Then, $f(a_1)f(a_2) = a_1Ba_2B = a_1a_2B = f(a_1a_2)$, so f is a homomorphism. The nullspace of this homomorphism is the set of all a such that $aB = B$. By Lemma 2.11, this equation is satisfied if $a \in A \cap B$. However, if $a \notin B$, then $a * 1_G = a \notin B$. Thus, the set of values of A that satisfy $aB = B$ is precisely $A \cap B$, so $Null(f) = A \cap B$. Thus f is a homomorphism with domain A, range AB/B, and nullspace $A \cap B$, so, by the 1st Isomorphism Theorem, $AB/B = A/(A \cap B)$. ■

Theorem 4.10. *(3rd Isomorphism Theorem) Let H and K be normal subgroups of a group G such that H is a subgroup of K. Then K/H is a normal subgroup of G/H and $(G/H)/(K/H) = G/K$.*

Proof. Let $k_1, k_2 \in K$. Then $k_1k_2^{-1} \in K$ since K is a group, so $k_1k_2^{-1} \in K$. Thus $(k_1H)(k_2H)^{-1} = k_1k_2^{-1}H \in K/H$, so K/H is a subgroup of G/H by the subgroup criterion. Now, let $c = (jH)^{-1}(k_1H)(jH)$ be a conjugate of k_1H in G/H, where $j \in G$. Then $c = (j^{-1}k_1j)H \in K/H$ since K is a normal subgroup of G. Thus K/H is a normal subgroup of G/H.

Now consider a map f from G/H to G/K such that $f(gH) = gK$. Let g_1, g_2 be elements of G such that $g_1H = g_2H$. Then, right-multiplying both sides by g_2^{-1} gives $g_1g_2^{-1}H = H$, so $g_1g_2^{-1} \in H$. Since H is a subgroup of K, $g_1g_2^{-1} \in K$, so $g_1K = g_2K$. Thus, f is a function. Also, since $f(g_1Hg_2H) = f(g_1g_2H) = g_1g_2K = (g_1K)(g_2K) = f(g_1H)f(g_2H)$, f is a homomorphism. We can also see that f is surjective because there exists a g corresponding to each left coset gK, and thus there exists a gH. By surjectivity, $Rng(f) = G/K$. Now, note that $Null(f)$ is the set of all gH such that $gK = K$. However, the set of all g that $gK = K$ is K. Therefore, $Null(f) = K/H$. Thus, f is a homomorphism with domain G/H, range G/K, and nullspace K/H. Hence, by the 1st Isomorphism Theorem, $(G/H)/(K/H) = G/K$. ■

Theorem 4.11. *(4th Isomorphism Theorem) Let G be a finite group and let N be a normal subgroup of G. Denote the set of subgroups of G containing N by $I_G(N)$, and let $A, B \in I_G(N)$. Then, $|I_G(N)|$ is equal to the number of subgroups of G/N.*

Proof. Let $H \in I_G(N)$ Since N is fixed under conjugation by elements of G, it is fixed under conjugation by elements of H. Let $hN \in H/N$. Since H is a subgroup of G, $h \in G$, so $hN \in G/N$. Thus H/N is a subgroup of G/N.

Now, let J be a group such that H is a normal subgroup of J such that J/N is a subgroup of G/N. Then every left coset jN for some $j \in G$ is a left coset of N wit respect to G, so it can be expressed as gN for some $g \in G$. Hence, $J \subset G$, and, since J is a group, J is a subgroup of G.

From the last 2 paragraphs, every element of $I_G(N)$ corresponds to a subgroup of G/N,

and vice versa. Thus, there's a bijection between the elements of $I_G(N)$ and the subgroups of G/N, and the theorem follows. ■

Theorem 4.12. *(1st Sylow Theorem) Let G be a group of order p^am, where p, a and m are positive integers , p is prime, and m is not divisible by p. Then, G has a subgroup of order p^a. Such a subgroup is called a* Sylow p-subgroup *of G.*

Proof. Assume that this theorem is not true. Then, there exists a group H of order p^am, where p, a, and m satisfy the properties given in the theorem statement. that has no Sylow p-subgroup, but whose subgroups all have Sylow p-subgroups. Then, the only subgroup of H whose order is divisible by p^a is H itself. We now take 2 cases.

CASE 1: p divides $|C_H(H)|$. Since $|C_H(H)|$ is an abelian group (by Corollary 3.15), by Cauchy's Theorem, it has a cyclic, normal subgroup J of order p. Then H/J is a subgroup of H of order $p^{a-1}m$, which has a Sylow p-subgroup S of order p^{a-1}. By the 4th Isomorphism Theorem, there exists a subgroup T of H such that $T/J = S$. Then $|T| = |J||S| = p * p^{a-1} = p^a$. Thus T is a Sylow p-subgroup of H, contradiction.

CASE 2: p does not divide $|C_H(H)|$. Then, by the Class Equation, $|H| = C_H(H) + \sum_{j=1}^{k} |H/C_H\{h_j\})$, where $h_1, h_2 \dots h_k$ Since $|H|$ is divisible by p but $|C_H(H)|$ isn't, there is some non-central element $j \in H$ such that $|H/C_H(\{j\})|$ is not divisible by p. Then, $|C_H(\{j\})|$ is divisible by p^a. If $|C_H(\{j\})| \neq H$, $|C_H(\{j\})|$,it would have a Sylow p-subgroup by our assumption about H, and so would H. Thus $|C_H(\{j\})| = H$. Hence j is central in H, contradiction.

Therefore, no group H satisfying our assumption exists. The theorem thus follows by the well-ordering principle. ■

4.2. **Group Solvability.** Burnside's Theorem states that a group is solvable if and only if it satisfies a certain condition. In this section, we will explore the meaning of the counter-intuitive but transformative concept of group solvability, which was discovered as a means to prove the impossibility of a solution to a general 5th degree polynomial equation.

Definition 4.13. A group G is *solvable* if there exists, for some positive integer k, a sequence of subgroups $C_1 = G_1, G_2, \dots G_k = G$ such that, for all j such that $1 \leq j \leq k-1$, G_j is a normal subgroup of G_{j+1} and G_{j+1}/G_j is abelian.

Example 4.2. All abelian groups A are solvable, as the sequence C_1, A satisfies all the conditions outlined in the definition above. D_4 is a non-abelian solvable group: the sequence C_1, C_4, D_4 satisfies all the conditions outlined above, as D_4/C_4 is the abelian group C_2. A_5, the group of permutations of $\{1, 2, 3, 4, 5\}$ with an even number of inversions, is not solvable, and this fact is used to prove that there is no solution in radicals to the general 5th degree polynomial.

Proposition 4.14. *Let G be a group and N be a normal subgroup of G. If N and G/N are both solvable, then so is G.*

Proof. Since N is solvable, there exist subgroups $C_1 = N_1, N_2, \dots N_k = N$ such that, for all j such that $1 \leq j \leq k-1$, N_j is a normal subgroup of N_{j+1} and N_{j+1}/N_j is abelian. Also, since G/N is solvable, there exists subgroups $C_1 = P_1, P_2, \dots P_k = G/N$ such that, for all j such that $1 \leq j \leq k-1$, P_j is a normal subgroup of P_{j+1} and P_{j+1}/P_j is abelian. By the 4th Isomorphism Theorem, there exist subgroups $N = Q_1, Q_2, \dots Q_k = G$ of G such that, for all j such that $1 \leq j \leq k-1$ $P_j = Q_j/N$ for all j, Q_j is a normal subgroup of Q_{j+1} and $Q_{j+1}/Q_j = P_{j+1}/P_j$. Thus, for all j such that $1 \leq j \leq k-1$, Q_{j+1}/Q_j is abelian. Thus, in the sequence of groups $C_1 = P_1, P_2 \dots P_k = G/N = Q_1, Q_2 \dots Q_k = G$, every group except for G is a normal subgroup of the next group, and the quotient group is abelian. Thus, G is solvable. ■

Theorem 4.15. *Let G be a finite group whose order is a power of a prime. Then G is solvable.*

Proof. We first prove the following lemma.

Lemma 4.16. *The center of G, a group whose order is a power of a fixed prime p, is nontrivial.*

Proof. Assume otherwise. Let $g_1, g_2 \dots g_n$ are the representatives of the non-central conjugacy classes of G. Since all the g_j are non-central, $G/C_G(\{g_j\})$ is a nontrivial subgroup of G for all integers j such that $1 \leq j \leq n$, so $\sum_{j=1}^{n} |G/C_G(\{g_j\})|$ is divisible by p. Now, when we take the class equation, $|G| = C_G(G) + \sum_{j=1}^{n} |G/C_G(\{g_j\})|$, we see that the RHS isn't divisible by p. Thus $|G|$ isn't divisible by p, contradicting our assumption about G. ■

Let q be a fixed prime, and let k be the minimum positive integer such that there exists an unsolvable group H with order q^k. We now take 2 cases:

CASE 1: H is not simple. Then H has a normal subgroup J that isn't C_1 or H. Then J and H/J are both subgroups of H whose orders can be expressed as q^m, where m is an integer such that $1 \leq m \leq k-1$. Thus, by our assumption about H, J and H/J are both solvable, so, by Proposition 4.14, H is solvable, contradiction.

CASE 2: H is simple. Consider $C_H(H)$, which, by Corollary 3.15, is a normal subgroup of H. Since H is simple, either $C_H(H) = H$ or $C_H(H) = C_1$. However, the first equation would imply that H is abelian and thus solvable. Therefore, $C_H(H) = C_1$, which contradicts Lemma 4.16.
Therefore, there is no group H satisfying our assumption. Hence, by the well-ordering principle, any group of prime-power order is solvable. ■

5. Linear Transformations

Representation theory relies on both group theory and linear algebra. This section is an overview of the linear algebra needed in group representation theory. Linear algebra is the study of vectors, vector spaces, and linear transformations and their representations as matrices.

Definition 5.1. Take a set F, and consider 2 binary operations, which we call "addition" $(+)$ and "multiplication" $(*)$. (This may or may not be the addition and multiplication you learned in elementary school.) Then F is a *field* if F is an abelian group under addition (with 0 as the additive identity), the nonzero elements of F form an abelian group under multiplication, and, for all $a, b, c \in F$, $a*(b+c) = a*b+a*c$ and $(a+b)*c = a*c+b*c$. The axiom involving the last two equations is called *distributivity* of multiplication over addition.

Example 5.1. The sets $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are fields under ordinary addition and multiplication. The set of algebraic numbers (i.e. roots of polynomials with integer coefficients) is also a field. $\mathbb{Z}$, however, is not a field because it isn't a group under multiplication (for example, the multiplicative inverse of 2 isn't an integer).

Definition 5.2. Let F be a field, whose elements are called "scalars" with field addition denoted by $+$ and field multiplication denoted by $*$. Consider a set S, whose elements are called "vectors", and 2 commutative binary operations, "vector addition" $\oplus$ that takes sends vector-vector pairs to vectors, and "scalar multiplication" $\otimes$, that sends scalar-vector (or vector-scalar) pairs to vectors. Then S is a *vector space* over F if it satisfies the following conditions:

- S is an abelian group under vector addition. The identity element of this group is called the *zero vector* , denoted by $\mathbf{0}$.
- Scalar multiplication is compatible with field multiplication, i.e., for all ordered triples $(a, b, \mathbf{c})$, where a and b are scalars and $\mathbf{c}$ is a vector, $(ab) \otimes \mathbf{c} = a(b \otimes \mathbf{c})$.
- For all vectors v, $1_F \otimes v = v \otimes 1_F = v$, where 1_F is the multiplicative identity of F.
- Scalar multiplication is distributive over vector addition, i.e. for all vectors $\mathbf{a}, \mathbf{b}$ and scalars c, $(\mathbf{a} \oplus \mathbf{b}) \otimes c = \mathbf{a} \otimes c \oplus \mathbf{b} \otimes c$ and $\mathbf{a} \otimes (\mathbf{b} \oplus c) = \mathbf{a} \otimes \mathbf{b} \oplus \mathbf{a} \otimes c$.
- Scalar multiplication is distributive over field addition, i.e. for all vectors $\mathbf{a}$ and scalars b,c, $\mathbf{a} \otimes (b + c) = \mathbf{a} \otimes b + \mathbf{a} \otimes c$

Example 5.2. $\mathbb{C}$, the set of ordered triples of real numbers, and the set of real square matrices of a given size are all vector spaces over $\mathbb{R}$. The set of polynomials with degree at most 3 (including the zero polynomial, which is defined to have degree -1) with complex coefficients is a vector space over $\mathbb{C}$.

Definition 5.3. Let $S = \mathbf{v}_1, \mathbf{v}_2 \dots \mathbf{v}_m$ be a subset of a vector space V over a field F. Then S is *linearly dependent* if there exist scalars $c_1, c_2 \dots c_n \in F$, not all 0, such that $\sum_{j=1}^{n} c_j \mathbf{v}_j = \mathbf{0}$ (i.e. if $\mathbf{0}$ can be expressed as a linear combination of vectors in S whose coefficients are not all 0. If no such $c_1, c_2 \dots c_n$ exist, S is *linearly independent*.

Definition 5.4. Let V be a vector space over F. Then, a subset $S = \{\mathbf{v}_1, \mathbf{v}_2 \dots \mathbf{v}_n\}$ of V is a *basis* for V if and only if S is linearly independent and every vector in V can be expressed

as $\sum_{j=1}^{n} c_j \mathbf{v}_j$ for some scalars $c_1, c_2 \dots c_n \in F$.. Equivalently, S is a basis for V if and only if every vector in V can be represented as a linear combination of vectors in S in exactly 1 way. The elements of a basis are called *basis vectors*. The *standard basis* of a vector space is the ordered basis corresponding to the set of ordered n-tuples of elements of F with all entries 0 except for a single 1, and, for all k such that $1 \leq k \leq n$, the kth vector has a 1 in the kth entry.

Theorem 5.5. *All bases of any given finitely-generated vector space have the same size (and by the same logic, all bases of any given infinitely-generated vector space have the same cardinality). The size of such a basis of a vector space V is called the* dimension *of V, denoted by* $\dim V$. *A vector space with dimension n is called* n-dimensional.

Proof. We prove the following key lemma.

Lemma 5.6. *Let B and C be subsets of V such that B is linearly independent and C generates V. Then $|B| \leq |C|$.*

Proof. Let $B = \{\mathbf{b}_1, \mathbf{b}_2 \dots \mathbf{b}_n\}$ and $C = \{\mathbf{c}_1, \mathbf{c}_2 \dots \mathbf{c}_m\}$, and assume $n > m$. Since C generates V, every element of V can be expressed as a linear combination of elements of C. Thus, adding an element of B to C gives a linearly dependent set. Therefore, the set $\{\mathbf{c}_1, \mathbf{c}_2 \dots \mathbf{c}_m, \mathbf{b}_1\}$ generates V. Using the same argument multiple times and inducting based on that shows there exists some proper subset of B that generates V. But this implies that B is linearly dependent, a contradiction. ■

Let X and Y be arbitrary bases for V. By definition, X and Y are both linearly independent and generate V. By Lemma 5.6, $|X| \leq |Y|$ and $|Y| \leq |X|$. Thus $|X| = |Y|$. Thus all bases of a given finitely-generated vector space have the same size. ■

Definition 5.7. Let f be a function between from V to W where V and W are vector spaces over the same field F Then f is a *linear transformation* if and only if all three of the following axioms are satisfied (here $+$ represents vector addition and $*$ represents scalar multiplication)

- f is compatible with vector addition: $f(\mathbf{a} + \mathbf{b}) = f(\mathbf{a}) + f(\mathbf{b})$ for all $\mathbf{a}, \mathbf{b} \in V$.
- f is compatible with scalar multiplication: $cf(\mathbf{a}) = f(c\mathbf{a})$ for all $\mathbf{a} \in V$ and $c \in F$.

Equivalently, a linear transformation is a homomorphism between vector spaces that is also compatible with scalar multiplication. If $W = V$, then the transformation is called an *linear operator* on V.

Definition 5.8. Let $B = \{\mathbf{b}_1, \mathbf{b}_2 \dots \mathbf{b}_n\}$ be an ordered basis for a given n-dimensional vector space V over a field F., and let T be a linear transformation from V to a m-dimensional vector space W. Let $C = \{\mathbf{c}_1, \mathbf{c}_2 \dots \mathbf{c}_n\} =$ Then the *matrix representation* of T with respect to B and C, denoted $[T]_B^C$ is the unique matrix $m * n$ matrix A (whose entries, which are elements of F, are denoted by a_{jk} where $j, k \in Z$, $1 \leq j \leq m$ and $1 \leq k \leq n$.) such that $T(b_k) = \sum_{j=1}^{m} a_{jk} \mathbf{c}_j$ for all integers k such that $1 \leq k \leq n$.

Definition 5.9. Let L be a linear transformation from a vector space V to a vector space W Then L is *invertible* if and only if it is bijective.

Proposition 5.10. *Let L be a linear transformation from vector space V to vector space W., where both vector spaces are over the same field F. Then, $Null(L)$ and $Rng(L)$ are subspaces of W.*

Proof. To prove that a subset of a vector space is a subspace, we only need to show that it is closed under vector addition and scalar multiplication. Now, let $\mathbf{x}, \mathbf{y} \in Null(L)$ and let $c \in F$. Then $L(\mathbf{x}+\mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y}) = 0$ and $L(c\mathbf{x}) = cL(\mathbf{x}) = 0$, so $\mathbf{x}+\mathbf{y}$ and cx are in $Null(L)$. Thus $Null(L)$ is closed under vector addition and scalar multiplication, so it's a subspace of V.

Let $w, z \in Rng(L)$. Then, there exist $\mathbf{a}, \mathbf{b} \in V$ such that $L(\mathbf{a}) = \mathbf{w}$ and $L(\mathbf{b}) = \mathbf{z}$. Then $L(c\mathbf{a}) = c\mathbf{w}$ and $L(\mathbf{a}+\mathbf{b}) = \mathbf{w}+\mathbf{z}$. Thus $c\mathbf{w}, \mathbf{w}+\mathbf{z} \in Rng(L)$. Hence $Rng(L)$ is closed under vector addition and scalar multiplication, so it's a subspace of V. ■

Proposition 5.11. *A linear operator is invertible if and only if it is injective.*

Proof. The "only if" portion is obvious. To prove the "if" portion, we let L be an injective linear operator on vector space V. Now assume for the sake of contradiction that there exists an element of V that isn't in $Rng(L)$. Then, L sends V to a proper subset of V.. By Proposition 5.10, this proper subset is a subspace of V, so its dimension is less than V. However, since L is injective, every output is mapped to at most 1 input, so $dim(Rng(L)) \geq dim(V)$, which is a contradiction. Thus, $Rng(L) = V$, so L is surjective. Hence L is bijective, and thus invertible. ■

Definition 5.12. Let L be a linear operator on vector space V over field F. For any vector-scalar pair $(\mathbf{v}, c)$, if $L(\mathbf{v}) = c\mathbf{v}$, $\mathbf{v}$ is called an *eigenvector* of L, and c is called an *eigenvalue* of L.

Definition 5.13. Let L be a linear operator on a vector space V. Then L is called *diagonalizable* if and only if there exists a basis B for V such that $[L]_B^B$ is a diagonal matrix.

Theorem 5.14. *A linear operator L over a vector space V is diagonalizable if and only if there is a set of eigenvectors of L that form a basis for V.*

Proof. For the matrix representation $[L]_B^B$ with entries a_{jk} of L with respect to a basis $B = \{\mathbf{b}_1, \mathbf{b}_2 \dots \mathbf{b}_n\}$ to be diagonal, for all j such that $1 \leq j \leq n$, $T(\mathbf{b}_j) = \sum_{j=1}^{n} a_{jj}\mathbf{b}_j$. Since the diagonal entries a_{jj} are scalars, the $\mathbf{b}_j$ are eigenvectors for all j such that $1 \leq j \leq n$. Thus B is a basis for V consisting of eigenvectors of L.

Conversely, if B was a basis for V consisting only of eigenvectors of L, and the eigenvalues of L are $c_1, c_2 \dots c_n$, the entries a_{jk} of $[L]_B^B$ would satisfy the system $T(\mathbf{b}_k) = \sum_{j=1}^{n} a_{jk}\mathbf{b}_j$

for all j, k between 1 and n. Replacing k with j gives $T(\mathbf{b}_j) = \sum_{j=1}^{n} a_{jj}\mathbf{b}_j$, so the diagonal entries of $[L]_B^B$ are the eigenvalues of L. ■

6. Intro to Representation Theory

The previous sections discussed the background in group theory and linear algebra that we need to understand the representation theory of finite groups.

Definition 6.1. Let V be a vector space over $\mathbb{C}$.. Then the *general linear group* of V, denoted by $GL(V)$, is the set of invertible linear operators on V, combined with the binary operation of function composition.

Definition 6.2. Let G be a finite group. A *representation* of G is a homomorphism between G to $GL(V)$, where V is a vector space over $\mathbb{C}$. We usually denote the images of representations using subscripts rather than parentheses, e.g. r_g. However, we may sometimes use $r(g)$ to avoid nesting subscripts.

Example 6.1. One representation of D_4 sending it to the general linear group of ordered pairs of complex numbers is the function r such that $r(id)$ is the identity transformation I (which keeps every vector in its domain the same), $r(r1)$ sends (a, b) to $(-b, a)$, $r(mv)$ sends (a, b) to $(-a, b)$, and all the other values are chosen to make r a homomorphism. The degree of this representation is 2 because the set of ordered pairs of complex numbers is a complex vector space of dimension 2. Also, for any finite group G, there's a linear representation of G sends all elements to I; that's called the *trivial representation* of G.

Definition 6.3. Let V and W be complex vector spaces, and let r and s be representations of a finite group G, sending G to $GL(V)$ and $GL(W)$, respectively. Then r and s are *isomorphic* if and only if there exists an invertible linear transformation f such that $f \circ r_g = s_g \circ f$ for all $g \in G$.

Definition 6.4. Let G be a group, V be a complex vector space, and r be a linear representation sending G to $GL(V)$. Then a subspace W of V is called *invariant* with respect to r if and only if, for all $g \in G$ and $\mathbf{w} \in W$, $r_g(w) \in W$.

Definition 6.5. Let V be a complex vector space and let W be a subspace of V (i.e. a subset of V that's a complex vector space over the same field and under the same operations of vector addition and scalar multiplication). Let r be a linear representation sending a finite group G to $GL(V)$, and let r^W be the restriction of r to W (i.e. the same representation, except the set of ordered pairs (g, r_g), is reduced to make the range of the representation $GL(W)$.). If W is invariant with respect to r, then r^W is a linear representation, and it's called a *subrepresentation* of r.

Definition 6.6. Let G be a finite group, V be a complex vector space, and r be a linear representation sending G to $GL(V)$. Then, r is *irreducible* if and only if it is nontrivial but has no nontrivial subrepresentations. Equivalently, r is irreducible if and only if V has no

proper, invariant subspaces with respect to r other than the trivial vector space (which contains only $\mathbf{0}$.)

Definition 6.7. Let r and s be linear representations of a finite group G over complex vector spaces V and W, respectively. Then, the *direct sum* of r and s, denoted by $r \oplus s$, is the unique representation such that $(r \oplus s)_g(\mathbf{x} + \mathbf{y}) = (r_g(\mathbf{x}), s_g(\mathbf{y}))$ for all $g \in G$ and $\mathbf{x}, \mathbf{y} \in V, W$, respectively. This representation sends G to $GL(V \oplus W)$, and the vector space $V \oplus W$ is called the *direct sum* of V and W.

Theorem 6.8. *(Maschke's Theorem) Let r be a linear representation of a finite group G sending G to the general linear group of a complex vector space V, and let s be a subrepresentation of r. Then there exists a subrepresentation t of r such that $r = t \oplus s$.*

See [Magnell 2022] for a proof of this theorem.

Corollary 6.9. *Every linear representation of a group G can be decomposed into a direct sum of irreducible representations of G.*

Proof. We strongly induct on $\dim V$. In the base case, V is 1-dimensional. Then r is irreducible, so the base case is true.

Now, assume that the theorem is true for all proper subrepresentations of r. Then, there exists a nontrivial proper G-invariant subspace W of V, wand r^W is a subrepresentation of r. By Maschke's theorem, there's a nontrivial subrepresentation t of r such that $r = r^W \oplus t$. By our inductive assumption, both r^W and t can be decomposed into irreducibles. Since r is the direct sum of representations that can be decomposed into irreducibles, r can be decomposed into irreducibles. ■

Theorem 6.10. *(Schur's Lemma) Let r and s be irreducible representations from G to $GL(V)$ and $GL(W)$, respectively, and let f be a linear transformation from V to W that's compatible with the group operation of G. Then f is a scalar multiple of I, and $f = 0$ if r and s aren't isomorphic.*

Proof. Let $g \in G$ and let $\mathbf{n} \in Null(f)$. Then $f(g\mathbf{n}) = gf(\mathbf{n}) = \mathbf{0}$, so $g\mathbf{n} \in Null(f)$. Hence $Null(f)$ is a G-invariant of V. However, since r is irreducible, $Null(f) = 0$ or $Null(f) = V$. Hence f is either an isomorphism or 0.. Thus, if r and s aren't isomorphic, $f = 0$..
Since $\mathbb{C}$ is algebraically closed, the eigenvalues of f are all complex, so there exists a complex number z such that $f - zI$ is not invertible. Since s is irreducible, $f - zI = 0$, so $f = zI$. ■

7. Characters

One of the most important concepts in representation theory is that of the character. The character of a linear representation is a function associated with the representation, closely associated with the trace of a linear operator or square matrix.

Definition 7.1. Let L be a linear operator. Then the *trace* of L, denoted by $tr(L)$, is the sum of the eigenvalues of L.

Example 7.1. The trace of an n-dimensional identity transformation is n.

Definition 7.2. Let r be a linear representation sending a finite group G to $GL(V)$ for some complex vector space V. Then the *character* of r, denoted by X_r (or sometimes $X(r)$ to avoid nested subscripts) , is the function that sends g to $Tr(r_g)$ for all $g \in G$.

Example 7.2. The character of a trivial representation is 1, because it's equal to the trace of I, and the trivial representation of any group is 1-dimensional

Theorem 7.3. *Let r be a linear representation sending a group G to $GL(V)$ for some complex vector space V. Then $X_r(g)$ is a sum of $X_r(1)$ roots of unity.*

Proof. We first prove the following lemma.

Lemma 7.4. *The minimal polynomial of r_g has distinct roots, and these roots are $|g|$th roots of unity, for all $g \in G$.*

Proof. Note that $(r_g)^{|g|} = r_{g^{|g|}} = r_{1_G} = I$. Thus $P_{min}(r_g)$ is a factor of $x^{|g|} - 1$ whose degree is that of r_g, so its roots are $X_r(1)$ distinct $|g|$th roots of unity. ■

Since r_g has a minimal polynomial with distinct roots, it's diagonalizable, and its eigenvalues are $|g|$th roots of unity. Thus the trace of r_g, which is $X_r(g)$, is the sum of $X_r(1)$ $|g|$th roots of unity. ■

Proposition 7.5. *The character of a direct sum of linear representations is the sum of the characters of the representations.*

Proof. Let r and s be linear representations sending a finite group G to $GL(V)$ and $GL(W)$, where V and W are complex vector spaces. Let $g \in G$, let z be an eigenvalue of $(r \oplus s)_g$, and let $\mathbf{v} \in V$ and $\mathbf{v} \in W$ such that $\mathbf{v} + \mathbf{w}$ is a corresponding eigenvector. Then,

$$(r \oplus s)_g(\mathbf{v} + \mathbf{w}) = (r_g(\mathbf{v}, s_g(\mathbf{w}) = z(\mathbf{v}, \mathbf{w}).$$

From there, we can see that the possible values of z are the eigenvalues of r_g and the eigenvalues of s_g. We can get the eigenvalues of r_g by setting $w = 0$, and we can get the eigenvalues of s_g by setting $v = 0$. Thus, the sum of the eigenvalues of $r \oplus s$, which is $X_{r \oplus s}$, is just $X_r + X_s$. ■

Definition 7.6. Let V be a complex vector space and f be a function sending pairs of vectors to scalars. Then f is an *inner product* if

- It is conjugate-symmetric (i.e $f(\mathbf{x},\mathbf{y}) = \overline{f(\mathbf{y},\mathbf{x})}$ for all vectors $\mathbf{x},\mathbf{y} \in V$.)
- It is linear in the first argument (i.e. $af(\mathbf{x},\mathbf{z}) + f(\mathbf{y},\mathbf{z}) = f(a\mathbf{x}+\mathbf{y},\mathbf{z})$ for all vectors $\mathbf{x},\mathbf{y},\mathbf{z} \in V$ and complex numbers a.) This, combined with conjugate-symmetry, gives that f is conjugate-linear in the second argument (i.e. $\overline{a}f(\mathbf{x},\mathbf{y})+f(\mathbf{x}+\mathbf{z}) = f(\mathbf{x}, a\mathbf{y}+\mathbf{z})$.
- It is positive definite (i.e. $f(\mathbf{x},\mathbf{x}) \in \mathbb{R}_{>0}$ for all nonzero $\mathbf{x}$ in V. for all vectors $\mathbf{x},\mathbf{y},\mathbf{z} \in V$ and complex numbers a.)

Definition 7.7. With respect to an inner product f, the *norm* of a vector $\mathbf{v}$ is $\sqrt{f(\mathbf{v},\mathbf{v})}$.

Definition 7.8. With respect to an inner product f, vectors $\mathbf{v}$ and $\mathbf{w} \in V$ are *orthogonal* if $f(\mathbf{v},\mathbf{w}) = 0$. If $\mathbf{v}$ and $\mathbf{w}$ are also unit vectors (i.e. have norm 1), they are *orthonormal*.

Definition 7.9. Let j and h be characters of linear representations of a group G. Let

$$\langle h, j\rangle = \frac{\sum_{g\in G} \overline{h(g)} j(g)}{|G|}.$$

Then $\langle h, j\rangle$ is an inner product, called the *Schur inner product* of G.

Proposition 7.10. *The characters of irreducible representations of a group G are orthonormal with respect to the Schur inner product of G.*

Proof. See [Linman, 2010] for a proof of this theorem. ■

Definition 7.11. The *regular representation* $r_R(G)$ of a group G is the representation sending $g \in G$ to the function f such that $f(h) = gh$ for all $h \in G$. Then $X_{r_R(G)}(1_G) = |G|$ and $X_{r_R(G)}(j) = 0$ for all nonidentity $j \in G$. This is because any image of a regular representation permutes a group, so the sum of the eigenvalues (trace) of such an image is the number of fixed points. The identity element maps to the identity transformation, which fixes $|G|$. All other elements map to permutations that fix no elements.

Proposition 7.12. *Let G be a finite group. Then,*

$$r_R(G) = \oplus_{j=1}^{k} r_j^{\oplus dim V_j},$$

where k is the number of irreducible representations of G and the r_j's are the distinct irreducible representations of G, and, for all j such that $1 \le j \le k$, r_j sends G to $GL(V_j)$.

Proof. Let

$$r_R(G) = \oplus_{m=1}^{k} r_m^{\oplus a_m},$$

where the a_m's are nonnegative integers. Then, by Proposition 7.5

$$\begin{aligned}
\langle X(r_R(G)), X(r_m)\rangle &= \frac{\sum_{g\in G}\overline{X(r_R(G))(g)}X(r_m)(g)}{|G|}\\
&= \frac{\sum_{g\in G}\overline{\sum_{j=1}^{k} a_j X(r_j)(g)}X(r_m)(g)}{|G|}\\
&= \frac{\sum_{j=1}^{k}\overline{a_j X(r_j)(g)}X(r_m)(g)}{|G|}\\
&= \sum_{j=1}^{k} a_j\langle x(r_j), x(r_m)\rangle\\
&= a_m\langle x(r_m), x(r_m)\rangle\\
&= a_m.
\end{aligned}$$

The last equation is true by Proposition 7.10. Note that this is true for all representations, not just regular representations.

For the regular representation, we can write (by Proposition 7.5)

$$\begin{aligned}
a_m &= \langle X(r_R(G)), X(r_m)\rangle\\
&= \frac{\sum_{g\in G}\overline{X(r_R(G))(g)}X(r_m)(g)}{|G|}\\
&= \frac{\overline{X(r_R(G))(1_G)}X(r_m)(1_G)}{|G|}\\
&= X(r_m)(1_G)\\
&= \dim V_m.
\end{aligned}$$

■

8. Algebraic Integers

In this section, we will discuss a special subset of the complex numbers called the algebraic integers. This concept might sound unrelated to group theory and representation theory, but it is crucial to our proof of Burnside's Theorem.

Definition 8.1. A complex number z is called an *algebraic integer* if there exists a monic polynomial P with integer coefficients such that $P(z) = 0$.

Example 8.1. All integers are algebraic integers (as they are roots of monic, linear polynomials with integer coefficients). Other algebraic integers include $\sqrt{2}$, $196883^{1/196883}$, and $\frac{-1+i\sqrt{3}}{2}$. Complex numbers that aren't algebraic integers include $\frac{1}{2}$, $\frac{-2+i\sqrt{5}}{7}$, and π.

Definition 8.2. The *minimal polynomial* of an algebraic integer z, denoted by $P_{min}(z)$, is the unique monic polynomial P of minimal degree such that $P(z) = 0$.

Example 8.2. $P_{min}(1) = x - 1$ and $P_{min}(196883^{1/196883}) = x^{196883} - 196883$.

Definition 8.3. Let z be an algebraic integer. Then the roots of $P_{min}(z)$ (including z) are called the *algebraic conjugates* of z.

Example 8.3. $P_{min}(\frac{-1+i\sqrt{3}}{2}) = x^2 + x + 1$. The only other root of this polynomial is $\frac{-1-i\sqrt{3}}{2}$, so the algebraic conjugates of $\frac{-1+i\sqrt{3}}{2}$ are $\frac{-1+i\sqrt{3}}{2}$ itself and $\frac{-1-i\sqrt{3}}{2}$.

Theorem 8.4. *The sum of any* 2 *algebraic integers is an algebraic integer.*

Proof. Let a and b be arbitrary algebraic integers such that $A = P_{min}(a)$ and $deg(A) = n \geq deg(P_{min}(b))$. Let $a_1 = a, a_2, a_3 \dots a_n$ be the algebraic conjugates of a, and let B, with degree n and possibly repeated roots $b_1 = b, b_2, b_3 \dots b_n$, be a monic polynomial with integer coefficients that is divisible by $P_{min}(b)$.

Now, take the multiset $S = a_j + b_k | 1 \leq j, k \leq n$ of possible sums that you get when a root of A is added to a root of B. Note that this set includes $a + b$. Let C be the monic polynomial whose roots are the elements of S. Since the elementary symmetric sums of S don't change when $a_1, a_2 \dots a_n, b_1, b_2 \dots b_n$ is permuted, these elementary symmetric sums are symmetric polynomials in the elements of S, with integer coefficients. Thus, by the Fundamental Theorem of Symmetric Polynomials, these can be expressed as polynomials in the union of the set of elementary symmetric sums of $a_1, a_2 \dots a_n$ and the set of elementary symmetric sums of $b_1, b_2 \dots b_n$. with integer coefficients. Since $a_1, a_2 \dots a_n$ is the multiset of roots of a monic polynomial with integer coefficients, all the elementary symmetric sums of these roots are integers (by Vieta's Formulas). Similarly, all the elementary symmetric sums of $b_1, b_2 \dots b_n$ are integers. Thus the elementary symmetric sums of S are all integers, so, by Vieta's Formulas, C is a monic polynomial with integer coefficients and $a + b$ as a root. Hence $a + b$ is an algebraic integer. ■

Theorem 8.5. *The product of any* 2 *algebraic integers is an algebraic integer.*

Proof. The proof of this theorem is the same as the proof of Theorem 8.4, except that S, which is the multiset of possible sums you get when a root of A is added to a root of B,is changed to the multiset of possible products you get when a root of A is multiplied by a root of B. ■

9. Proof of Burnside's Theorem

In this section we will prove Burnside's p^aq^b theorem. The theorem can easily be understood by a student who knows group theory but has never learned representation theory,

but it's very hard to prove with only group theory. (It has been done, but by then it was more than half a century since Burnside found his proof, and the group-theoretic proofs are extremely complicated)

Theorem 9.1. *(Burnside's Theorem) Let G be a group whose order has at most 2 prime factors. Then G is solvable.*

Proof. We first prove 2 lemmas.

Lemma 9.2. *Let r be an irreducible representation sending a finite group G to $GL(V)$ for some complex vector space V. Let K be a conjugacy class of G such that $\gcd(|K|, X_r(1)) = 1$. Then either $X_r(g) = 0$ or r_g is a scalar multiple of I, for all $g \in K$.*

Proof. We know that $|K|$ and $X_r(1)$ are both integers (one is the dimension of a subspace and the other is the size of a finite set). Since their gcd is 1, there exist integers a, b such that $a|K| + bX_r(1) = 1$, by Bezout's Identity. Now let $g \in K$, and multiply both sides of this equation by $\frac{X_r(g)}{X_r(1)}$. We get

$$\frac{a|K|X_r(g)}{X_r(1)} + bX_r(g) = \frac{X_r(g)}{X_r(1)}.$$

The set of $\mathbb{C}[G]$ of linear combinations of elements of G with complex coefficients is an abelian group with respect to addition, and the center of that group is spanned by the set of sums of elements of conjugacy classes of G (because the sums of elements of a conjugacy class of G are fixed under conjugation by any element of G). Let $c \in \mathbb{C}[G]$. Since V is irreducible, by Schur's Lemma, every vector in V. is an eigenvector of c. Let $\mathbf{v} \in V$ and let the corresponding eigenvalue be z, which is a sum of roots of unity, and thus an algebraic integer. Taking traces gives $zx_r(1) = |K|x_g$, so $\frac{Kx_g}{x_r(1)} = z$, an algebraic integer. Since

$$\frac{a|K|X_r(g)}{X_r(1)} + bX_r(g) = \frac{X_r(g)}{X_r(1)},$$

$\frac{X_r(g)}{X_r(1)}$ is an algebraic integer.

Since $X_r(g)$ is the sum of $X_r(1)$ roots of unity , $\frac{X_r(g)}{X_r(1)}$ is a sum of $X_r(1)$ roots of unity divided by $X_r(1)$. From this, we can see that all the algebraic conjugates of $\frac{X_r(g)}{X_r(1)}$ are sums of $X_r(1)$ roots of unity divided by $X_r(1)$, so their magnitudes are all at most 1. Since $\frac{X_r(g)}{X_r(1)}$ is an algebraic integer, the product of the algebraic conjugates of $\frac{X_r(g)}{X_r(1)}$ is an integer, by Vieta's Formulas. Since the magnitudes of the conjugates are all at most 1, the product of the conjugates is either 0 or ± 1.

If the product of the conjugates is 0, at least 1 of the conjugates is 0, and since the polynomial whose roots are the conjugates is $P_{min}(\frac{X_r(g)}{X_r(1)})$, $\frac{X_r(g)}{X_r(1)} = 0$. Thus, $X_r(g) = 0$. In the other case, the product of the conjugates is ± 1. Then the roots of unity which $X_r(g)$ is a sum of are all the same, so $|X_r(g)| = X_r(1)$. By Lemma 7.4, r_g has a minimal polynomial with distinct roots that are $|g|$th roots of unity, so it's diagonalizable, and its eigenvalues $d_1, d_2 \dots d_{X_r(1)}$ are $|g|$th roots of unity.Then

$$|Tr(r_g)| = |X_r(g)| = |\sum_{j=1}^{X_r(1)} d_j| \leq \sum_{j=1}^{X_r(1)} |d_j| = X_r(1),$$

with equality if and only if the d_j are all equal. Thus r_g is a diagonalizable linear operator with equal eigenvalues, so it's a scalar multiple of I. ■

Lemma 9.3. *Let G be a simple group with a conjugacy class whose size is a positive integer power of a prime p. Then G is cyclic of prime order.*

Proof. Let r_1 be the trivial representation of G and let $r_2 \dots r_n$ be the irreducible representations of G, where r_j sends G to $GL(V_j)$. Consider the regular representation r_R of G and a nonidentity element g of G. Then, by Proposition 7.12,

$$0 = X_{r_R}(g) = \sum_{j=1}^{n} \dim(V_j) X_{r_j}(g) = \sum_{j=1}^{n} X_{r_j}(1) X_{r_j}(g).$$

Since $X_{r_1}(g)X_{r_1}(1) = 1$, $\sum_{j=2}^{n} X_{r_j}(1)X_{r_j}(g) = -1$.

Now, we take 2 cases. In the first case, for all k such that $1 \leq k \leq n$, either $X_{r_k}(1)$ is divisible by p or $X_{r_k}(g) = 0$. Since either $X_{r_k}(1)$ is divisible by p or $X_{r_k}(g) = 0$ for all k such that $1 \leq k \leq n$,

$$\frac{\sum_{j=2}^{n} X_{r_j}(1)X_{r_j}(g)}{p} = \frac{-1}{p}$$

is a sum of integer multiples of values of $X_{r_k}(g)$. By Theorem 7.3, these values are sums of roots of unity, and thus algebraic integers, making $\frac{-1}{p}$ an algebraic integer, contradiction. In the second case, there exists a value of m such that $1 \leq m \leq n$ such that $X_{r_m}(1)$ is not divisible by p and $X_{r_m}(g)$ is nonzero. Then $\gcd(|K|, X_r(1)) = 1$, so by Lemma 9.2, $r_m(k)$ is a nonzero scalar multiple of I for all $k \in K$, so r_g is therefore central in $Rng(r)$. Letting $h \in G$ we can thus write $r_h r_g = r_g r_h$, so $r_{hg} = r_{gh}$. By Lemma 4.6, $Null(r)$ is a normal subgroup of G, but since G is a nontrivial simple group, $Null(r) = C_1$, so r is injective. Therefore, $r_{hg} = r_{gh}$ implies $hg = gh$ for all $h \in G$, so $g \in C_G(G)$. Since $g \neq 1_G$ and $C_G(G)$ is a normal subgroup of G (by Corollary 3.15), $C_G(G) = G$, so G is abelian. Hence every subgroup of G is normal, but G also has no nontrivial normal subgroups except for G itself. Thus G has no nontrivial proper subgroups, so it's a cyclic group of prime order. ■

We now use our Lemmas to prove Burnside's Theorem. Let x and y be fixed prime numbers, and assume for the sake of contradiction that there exists an unsolvable group H of order $x^a y^b$ where a and b are nonnegative integers, such that every group whose order is a proper

divisor of $x^a y^b$ is solvable. Then, since all abelian groups are solvable, H is not abelian. We take 2 cases:

CASE 1: H is not simple. Then it has a proper, nontrivial, normal subgroup J whose order is a proper divisor of $x^a y^b$, and the order of the quotient group, $|H/J| = \frac{|H|}{|J|}$ is also a proper divisor of $x^a y^b$. Since these groups have order that are proper divisors of $x^a y^b$, they are both solvable. Then by Proposition 4.14, H is solvable, contradiction.

CASE 2: H is simple. Then, by the 1st Sylow Theorem, there exists a Sylow x-subgroup X of G. Then $|X| = x^a$, so, by Lemma 4.16, there exists a nonidentity element $c \in C_X(X)$. Thus X is a subgroup of $C_H(c)$, so $|C_H(c)|$ is divisible by $|X| = p^a$. By Theorem 3.16, $K_H(c) = \frac{|H|}{|C_H(c)|}$ is not divisible by p, so it's a power of q. By Lemma 9.3, H is either trivial or cyclic of prime order, so it is abelian, contradiction.

Thus no group H satisfying our assumption exists. Hence, by the well-ordering principle, any group G whose order has at most 2 prime factors is solvable. ■

10. Conclusion

Other than its practical applications, Burnside's Theorem has a lot of applications in pure math. For instance, Burnside's Theorem was used to prove that every acyclic finite simple group has even order (known as the Feit-Thompson Theorem) [Linman, 2010] and, eventually, the following groundbreaking theorem that, in effect, creates a "periodic table" of all the finite simple groups.

Theorem 10.1. *(Enormous Theorem) Every nontrivial finite simple group, up to isomorphism, is either cyclic of prime order, alternating (i.e. the group of n-permutations with an even number of inversions for some integer $n \geq 5$), in one of 16 "Lie-type" infinite families, or one of 26 or 27 "sporadic" groups (Some texts consider one group, the Tits group, to be of Lie-type and other texts consider it to be sporadic).*

11. Acknowledgements

Thanks to Simon Rubinstein-Salzedo, Ari Krishna, and Euler Circle for teaching me how to write and present mathematical research papers and for helping me fill any gaps in my knowledge of representation theory prior to writing this paper.
Thanks also to Julie Linman for writing a reference that is especially easy to understand and learn the theory of group representations from.

References

[Magnell, 2022] Ian Magnell *https://math.uchicago.edu/∼may/REU2022/REUPapers/Magnell.pdf* Linear Representations of Finite Groups

[Linman, 2010] Julie Linman *https://www.ms.uky.edu/∼sohum/ma561/notes/workspace/burnside_theorem.pdf* Burnside's Theorem

[Bouchard] Vincent Bouchard *https://sites.ualberta.ca/∼vbouchar/MAPH464/chapter−applications.html* MA PH 464 - Group Theory In Physics: Lecture notes

Euler Circle, Mountain View, CA 94040
Email address: emailpranavshankar@gmail.com